Méthodes de Monte Carlo multi-niveaux améliorées pour l'évaluation des options à barrière en finance

Mouna Ben Derouich

UNIVERSITÉ SORBONNE PARIS NORD

UNIVERSITÉ PARIS XIII - SORBONNE PARIS NORD

École Doctorale Sciences, Technologies, Santé Galilée

Improved Multi Level Monte Carlo Methods for Pricing Barrier Options in Finance

THÈSE DE DOCTORAT

Présentée par

Mouna BEN DEROUICH

Laboratoire Analyse, Géométrie et Applications

Pour l'obtention du grade de

DOCTEUR EN MATHÉMATIQUES APPLIQUÉES

Soutenue le 15 Décembre 2023 devant le jury d'examen composé de :

M. Aurélien ALFONSI	Professeur à l'École Nationale des Ponts et Chaussées	Président du jury
M. Christian BAYER	Professeur à Weierstrass Institute	Rapporteur
Mme. Caroline HILLAIRET	Professeure à l'ENSAE Paris	Rapportrice
M. Yueyun HU	Professeur à l'Université Sorbonne Paris Nord	Examinateur
M. Ahmed KEBAIER	Professeur à l'Université d'Évry	Directeur de thèse
M. Mohamed MNIF	Professeur à l'École Nationale d'Ingenieurs de Tunis	Examinateur
M. Mohamed MRAD	MCF HDR à l'Université Sorbonne Paris Nord	Examinateur

Remerciements

Je tiens tout d'abord à exprimer un remerciement profond et sincère à mon directeur de thèse Ahmed KEBAIER qui m'a accueillie au sein du Laboratoire Analyse, Géométrie et Applications, fait confiance et donnée envie de réaliser cette thèse. Sa disponibilité constante pour me guider et répondre à mes questions a été inestimable et ses conseils avisés ont contribué à alimenter ma réflexion. Si ces années de thèse demeurent un souvenir mémorable, c'est en grande partie grâce à son soutien indéfectible. Ses qualités humaines, marquées par sa gentillesse, sa patience et sa compréhension à mon égard ont été d'une grande valeur pour moi .

Je tiens aussi à remercier Caroline HILLAIRET et Christian BAYER qui ont accepté de rapporter cette thèse. Merci pour votre lecture attentive, et le temps que vous m'avez consacrée. Je suis également très reconnaissante à l'encontre de Mohamed MNIF, Yueyun HU, Aurélien ALFONSI et Mohamed MRAD qui ont accepté d'examiner mon travail.

Le temps que j'ai passé au LAGA a été rendu encore plus agréable grâce à mes collègues: Elie, Neige, Safa, Hugo, Samar, Moussa, Wassim, Mohamed, Ilyes, Arthur, Tristan, Vincent, Amine, Quang, Oisin et Ghina. J'adresse aussi mes remerciements au personnel de l'administration du LAGA, Yolande et Leila, pour leur gentillesse et leur aide tout au long de mes années de thèse.

Je tiens également à exprimer toute ma gratitude envers Cyril pour son aide, ses précieux conseils, sa patience et surtout pour toutes les discussions intéressantes qu'on a pu avoir.

Je remercie également tous mes amis qui rendent ma vie si agréable, et qui ont toujours eu une place importante dans celle-ci. Je remercie particulièrement Arij, Amel, Marwa, Chefia et Tram. Leur bonne humeur, les moments conviviaux partagés, ainsi que leur soutien durant ma thèse, me sont très précieux.

Je suis à court de mots pour exprimer ma gratitude envers mes chers parents. A ma mère, pour son rôle inestimable dans la personne que je suis devenue aujourd'hui. Je te remercie infiniment pour les années de vie, les sacrifices et le support que tu m'a donnée. Je t'aime très fort. A mon père, pour son écoute et ses encouragements constants. Tu as joué non seulement le rôle de parent mais tu as été un vrai soutien pour moi. Je remercie du fond du cœur ma chère sœur Mariem, qui m'a aimée, aidée, soutenue et encouragée tout au long de cette thèse, dans les bons moments comme dans les moments difficiles. Une pensée particulière à toutes les personnes qui m'ont quittées durant cette période de thèse. Ma grand-mère Baya, ta perte a été la plus grande tragédie que j'ai pu rencontrée. Tu resteras la personne à qui je porte le plus d'affection et d'amour. A ma tante Bahija, tu as toujours été à mon écoute, je ne t'oublierai jamais...A Raja, je ressens un grand vide depuis nos discussions, souvent intenses, me manquent tellement. Vous êtes parties trop tôt. Vous resterez gravées dans mon cœur à vie.

Enfin, je souhaite exprimer ma gratitude envers mon partenaire Zakaria, à qui je dédie cette thèse. Je te remercie pour ta patience, ton soutien et ton réconfort infaillibles. Merci pour ton amour et ta présence constante à mes côtés. Je n'aurais jamais pu accomplir cela sans toi.

Résumé

Dans cette thèse, nous démontrons que les techniques de "Multilevel Monte Carlo" (MLMC) peuvent être utilisées pour évaluer les prix des options barrières dans des modèles avec des coefficients de diffusion non nécessairement globalement Lipschitz. Dans la première partie, nous considérons un cadre général de modèles unidimensionnels avec des coefficients de diffusion qui ne sont pas nécessairement globalement lipschitziens. Nous introduisons ensuite un schéma d'Euler implicite interpolé pour lequel nous démontrons une convergence forte d'ordre un. Nous analysons également les trajectoires extrêmes du processus de diffusion et de son approximation, et prouvons que la méthode MLMC atteint son régime optimal en $O(\varepsilon^{-2})$ pour une précision donnée ε. Nous appliquons ces résultats à l'évaluation des options barrières dans le modèle CIR et le modèle de volatilité locale CEV, et développons des formules semi-analytiques pour les densités du minimum et du maximum associés à ces deux processus. Dans la deuxième partie, nous étendons cette approche au problème plus complexe de l'évaluation des options barrières dans le modèle bidimensionnel log-Heston. Nous développons un schéma d'approximation permettant d'analyser la variance de la méthode MLMC combinée à des techniques de pont Brownien. Nous montrons que la méthode MLMC obtenue a une complexité d'ordre inférieur à une méthode de Monte Carlo classique, mais sans atteindre le régime optimal. Enfin, dans la dernière partie nous étudions l'efficacité de la méthode "Gaussian Process Regression" (GPR) pour calculer les prix et les sensibilités d'une option barrière "Down and Out" dans le modèle de Heston. Nous obtenons que la précision de la méthode GPR reste acceptable et fonctionne raisonnablement bien pour l'évaluation des options barrières. Nous développons ensuite des formules analytiques pour calculer les sensibilités Delta (Δ) et Vega (ν) de l'option barrière "Down and Out" pour trois schémas d'approximation différents du modèle log-Heston. Ces procédures nous permettent de quantifier l'incertitude associée à l'utilisation de formules de calcul des sensibilités fournies automatiquement par la méthode GPR entraînée.

Abstract

In this thesis, we demonstrate how Multi-Level Monte Carlo (MLMC) techniques can be used to price barrier options for models possibly with non-globally Lipschitz diffusion coefficients. In the first part, we consider a general framework of one-dimensional models with diffusion coefficients that are not necessarily globally Lipschitz. We then introduce an interpolated implicit Euler scheme for which we prove a strong convergence result of order one. We also analyse the extreme trajectories of the diffusion process and its approximation, and prove that the MLMC method reaches its optimal regime $O(\varepsilon^{-2})$ for a given total precision ε. We apply these results to the pricing of barrier options in the CIR model and the CEV local volatility model, and develop semi-analytical formulas for the densities of the running minimum and maximum of these two processes. In the second part, we extend this approach to the more challenging problem of pricing barrier options under the two-dimensional log-Heston model. We develop an approximation scheme allowing us to analyse the variance of the MLMC method when combined with Brownian bridge techniques. We show that the aforementioned MLMC method has a lower order complexity than the standard Monte Carlo method, however it does not reach the optimal regime. In the last part, we analyse the Gaussian Process Regression (GPR) method to calculate the prices and sensitivities of a Down and Out barrier option in the log-Heston model. We show that the accuracy of the GPR method is acceptable and remains a reasonable approach for pricing Barrier options. We then develop analytical formulae for computing the sensitivities Delta (Δ) and Vega (ν) of the Down and Out barrier option for different approximation schemes of the log-Heston model using the "Pathwise Monte Carlo" method. These procedures enable us to quantify the uncertainty associated with the use of sensitivity calculation formulae provided automatically by a trained GPR method.

Contents

Chapter 1

Introduction

L'objectif de cette thèse est d'étudier les méthodes Multi-Level Monte Carlo (MLMC) pour le calcul des espérances intervenant dans la valorisation et la gestion de risques des produits dérivés. On s'intéresse particulièrement à des produits dépendant de la trajectoire ("Path-dependant") dans le cadre de modèle à coefficients non-Lipschitz. Des schémas améliorant le temps de calcul par rapport aux méthodes Monte-Carlo classiques seront présentés pour les modèles CIR, CEV et Heston. On finira par l'étude de la méthode d'apprentissage statistique Gaussian Processes Regression pour le calcul des sensibilités des options barrières dans le modèle de Heston.

1.1 Modèles d'évaluation d'options

La théorie de l'évaluation des options a été développée en 1973 par Black et Scholes [74]. Ils modélisent la dynamique des prix des actifs sous les hypothèses suivantes:

- Absence d'opportunité d'arbitrage.
- Les actifs ne paient pas de dividendes.
- Les frais de transaction et taxes sont négligeables.
- Il est possible d'acheter et de vendre des fractions d'actions.
- Le taux d'intérêt sans risque est constant et connu.
- Le prix du sous-jacent $(S_t)_{t\geq 0}$ suit un mouvement brownien géométrique de drift μ et volatilité σ constantes, i.e. $(S_t)_{t\geq 0}$ est une solution de l'E.D.S

$$dS_t = \mu S_t dt + \sigma S_t dW_t.$$

où W_t est un mouvement Brownien.

L'avantage principal du modèle Black et Scholes (BS) est qu'il permet d'évaluer le prix des produits financiers les plus traités sur le marché et de calculer les indicateurs de risques avec des formules fermées, ce qui permet d'optimiser le temps de calcul. Cependant, bien que considérée comme l'une des innovations les plus remarquables de la finance de marché, le modèle BS est souvent critiqué pour ses hypothèses ne reflétant pas la réalité des marchés financiers. D'abord, le rendement des prix des actions ne suit pas une distribution log-normale. De plus, le modèle de BS suppose que la volatilité du sous-jacent est constante pendant toute la durée de vie de l'option indépendamment du niveau de prix du sous-jacent, ce qui n'est pas réaliste. En effet, pour chaque option d'achat de maturité T et de strike K dont le prix est observable sur le marché,

on peut calculer sa *volatilité implicite* $\sigma(T,K)$ comme la volatilité à utiliser dans le modèle BS pour que le prix calculé dans le modèle coïncide avec le prix observé sur le marché. Ainsi, étant donné une famille de prix $C(T_i,K_i)$ observés sur le marché, on en déduit une nappe de volatilité implicite $\sigma(T_i,K_i)$. En pratique, cette nappe n'est pas constante comme le suppose le modéle de BS, et a la forme d'un sourire "Smile". Le modèle de BS n'est donc pas adapté à l'évaluation et la couverture des options exotiques dépendant fortement de la volatilité. Pour prendre en compte ces observations, il est nécessaire d'utiliser des modèles plus sophistiqués. Diverses approches de modélisation de la volatilité ont été introduites au cours du temps. On distingue principalement trois types d'approches : les modèles à volatilité locale, les modèles à volatilité stochastique et les modèles hybrides. Dans la suite, on s'intéressera aux deux premières classes de modèles.

Les modèles à volatilité locale supposent que la volatilité est une fonction déterministe du temps et des paramètres du marché. En d'autre termes, on suppose que $(S_t)_{t\geq 0}$ est solution de l'équation différentielle stochastique (E.D.S) suivante :

$$dS_t = b(t,S_t)dt + \sigma(t,S_t)dW_t,$$

où $\sigma(t,x)$ est une fonction déterministe en t et x. L'un des modèles de volatilité locale qui nous intéresse dans cette thèse est le modèle d'élasticité constante de la variance (CEV) introduit par Cox [20] en 1975. Sa dynamique est définie par :

$$\begin{aligned} &dS_t = \mu S_t dt + \sigma S_t^{\alpha} dW_t, \;\; t \geq 0, \\ &S_0 \in \mathbb{R}_+, \end{aligned} \tag{1.1}$$

avec $\mu \in \mathbb{R}$, $\sigma > 0$ et $\alpha \in \mathbb{R}$ des constantes. Ici α est le paramètre d'élasticité de la volatilité locale. Selon la valeur de α, les propriétés du modèle changent et ont été étudiées par différents auteurs. Notamment, le cas $\alpha < 1$ a été étudié par Cox et Ross [22]. Dans ce cas, la volatilité locale augmente lorsque le prix de l'action diminue. De plus, Cox et Ross [22] ont obtenu une solution analytique pour le calcul du prix d'une option d'achat. Le cas $\alpha > 1$ a été introduit par Emanuel et Macbeth [30] en 1982. Dans ce cas, la volatilité locale augmente au fur et à mesure que le prix de l'action augmente. Cela crée une distribution de probabilité avec une queue de distribution plus épaisse à droite qu'à gauche. Cela correspond à un "Smile" de volatilité où la volatilité implicite est une fonction croissante du prix d'exercice. Également, une solution analytique pour le prix d'une option d'achat vanille a été établie dans [75] et [27]. Le cas $\alpha = 1$ correspond au modèle de BS. Le cas où $\alpha = \frac{1}{2}$ a été étudié en détail. En considérant de plus un terme de retour à la moyenne, on obtient le modèle de Cox-Ingersoll-Ross (CIR) [21]. Ce dernier a été conçu pour modéliser les taux courts. Cependant, il ne permet pas de modéliser le cas ou les taux d'intérêts sont négatifs. La dynamique du modèle CIR s'écrit alors :

$$\left\{ \begin{array}{rcl} dX_t &=& \kappa(\theta - V_t)dt + \sigma\sqrt{X_t}dW_t, \\ X_0 &=& x > 0, \end{array} \right. \tag{1.2}$$

avec $W = (W)_{t\geq 0}$ est un mouvement Brownien standard, $\kappa, \theta > 0$, $\sigma > 0$, $X_0 = x > 0$.

D'après le test de Feller (voir [76, Chapter 5]), (1.2) admet une unique solution forte strictement positive losque la condition $a \geq \sigma^2/2$ est satisfaite. Dans ces modèles à volatilité locale le "Smile" de volatilité peut être parfaitement généré en calibrant le modèle aux volatilités implicites des prix d'options européennes. En outre, cette calibration est rapide car elle est faite à l'aide de formule fermées. Cependant, ces modèles unidimensionnels ne tiennent pas compte du bruit exogène à celui de l'actif sous-jacent. Afin de pallier ce problème, les modèles à volatilité stochastique supposent que la volatilité est elle même un processus stochastique. Parmi les modèle à volatilité

stochastique, l'un des plus utilisés pour évaluer le prix d'une option est le modèle d'Heston [49]. Il suppose que la volatilité $(V_t)_{t\geq 0}$ suit une dynamique de CIR et que le prix $(S_t)_{t\geq 0}$ de l'actif associée suit une loi log-normale. Ainsi, le couple (S, V) est solution de l'EDS suivante:

$$\begin{aligned} dS_t &= rS_t dt + \sqrt{V_t} S_t \;\; d\left(\rho W_t^v + \sqrt{1-\rho^2} W_t^s\right) \\ dV_t &= \kappa(\theta - V_t)dt + \sigma\sqrt{V_t} dW_t^v, \end{aligned} \tag{1.3}$$

avec $\kappa, \theta, \sigma, S_0, V_0 > 0$, $r \in \mathbb{R}$, $\rho \in [-1, 1]$, $T > 0$, et où W_t^s, W_t^v sont deux mouvements browniens indépendants. On a les interprétations financiers suivantes:

- r est le taux de rendement espéré instantané de l'action,
- θ est la moyenne de la volatilité à long terme,
- κ correspond à la vitesse de retour à la moyenne de la volatilité,
- σ représente la volatilité de la volatilité,
- ρ est le coefficient de corrélation entre le prix de l'actif et sa volatilité.

Sous les conditions de Feller, la condition $2\kappa\theta > \sigma^2$ assure que la volatilité reste positive et ainsi que le processus V est bien défini (voir e.g. [55]). Ce type de modèle permet de mieux capturer la surface de volatilité "forward" ce qui permet une évaluation pertinente des produits "path-dependant" comme les "option on option" ou bien "forward start option ", mais l'absence de formules fermées permettant de calculer le prix des options notamment de type exotique rend la calibration de ces modèles lentes et donc leur utilisation plus difficile en pratique.

Parmi ces options exotiques on note les options barrières qui dépendent de toute la trajectoire du prix du sous-jacent. On distingue deux types d'options barrières:

- les options "Knock-Out" de barrière $B > 0$ et de maturité $T > 0$, qui se désactivent automatiquement lorsque le prix du sous-jacent atteint au moins une fois la barrière B avant T;
- les options "Knock-in" de barrière B et de maturité $T > 0$, qui s'activent automatiquement lorsque le prix du sous-jacent atteint au moins une fois la barrière B avant T.

Prenons l'exemple de l'option "Down-and-Out" qui est de type "Knock-Out". Elle est définie par sa barrière $B < S_0$ et se désactive si le prix S du sous-jacent passe sous la barrière B avant la maturité $T > 0$. Le payoff d'une option barrière "Down-and-Out" s'écrit alors:

$$f(S_T)\mathbb{1}_{\{\tau_D > T\}},$$

avec $\tau_D := \inf\{t \in [0, T] | S_t \leq B\}$ et $f : \mathbb{R} \to \mathbb{R}_+$ la fonction de payoff. Si $\tau_D < T$, c'est à dire il existe $t < T$ tel que $S_t \leq B$, on dit que la barrière a été franchie avant la maturité et l'option Down-and-Out est désactivée, son acheteur obtient 0 à la maturité. Par contre si $\tau_D > T$, alors $S_t > B$ pour tout $t \in [0, T]$. Dans ce cas, la barrière n'a pas été franchie et l'acheteur de l'option obtient alors $f(S_T)$ à la maturité. Le payoff d'une option barrière dépend donc de la valeur terminale S_T du prix du sous-jacent, mais également de toute la trajectoire $(S_t)_{t\in[0,T]}$. Leur valorisation est donc particulièrement non-triviale.

Plusieurs approches sont disponibles pour le calcul effectif de prix d'options. On peut citer en particulier :

- La résolution analytique de l'EDS définissant le prix. Cependant, l'obtention de formules analytiques ou semi-analytiques est connue seulement pour des modèles simples tels que le modèle de Black-Scholes, et seulement pour certaines options particulièrement simples, typiquement des options européennes d'achat ou de vente.
- L'utilisation de formule de type Feynman-Kac. On peut souvent démontrer que la fonction prix d'une option est la solution d'une équations aux dérivées partielles (EDP), en utilisant une formule de type Feynman-Kac. On peut ensuite résoudre cette EDP en utilisant des méthodes numériques, par exemple la méthode des différences finies. Ces méthodes permettent en général d'avoir des résultats précis, mais sont coûteuses en terme de temps de calcul, en particulier lorsque la dimension des processus stochastiques sous-jacents est grande. De plus, ces méthodes sont difficiles à appliquer dans le cadre des modèles à volatilité stochastique.
- Les méthodes de type Monte-Carlo. Elle sont applicables aux différents types de modèles, et permettent d'avoir des résultats en un temps raisonnable. Ces méthodes sont en général plus simples à utiliser pour les options exotiques que les méthodes par résolution d'EDP, et peuvent facilement être utilisées en grande dimension. Elles sont donc utilisées pour la valorisation d'options complexes pour lesquelles il n'y pas des formules fermées.

Il existe plusieurs méthodes de type Monte-Carlo, que nous détaillons maintenant.

1.2 Méthode de Monte Carlo standard

On considère que X, le processus de prix du sous-jacent, est solution de l'équation différentielle stochastique (EDS) suivante sous la mesure risque-neutre:

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t, \;\; t \in [0, T], \tag{1.4}$$

avec $X_0 = x \in \mathbb{R}^d$, W un mouvement Brownien de dimension $q \geq 1$ et $a : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$, $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\times q}$ deux fonctions globalement Lipschitz en espace uniformément en temps.

Soit $f : \mathbb{R}^d \to \mathbb{R}$ le payoff d'une option européenne. On cherche à calculer $\mathbb{E}\big[f(X_T)\big]$, correspondant au prix de l'option. La méthode de Monte-Carlo standard consiste à simuler $N \geq 1$ trajectoires du processus X sous la probabilité risque-neutre. On calcule ensuite l'espérance comme moyenne empirique sur les trajectoires simulées:

$$\mathbb{E}\left[f(X_T)\right] \simeq \frac{1}{N}\sum_{i=1}^{N} f(X_T^i).$$

Par la loi forte des grands nombres, si $X_T \in L^1(\mathbb{P})$ et si $(X_T^i)_{i\geq 1}$ est un échantillon indépendant et identiquement distribué selon la loi de X_T, on a

$$\frac{1}{n}\sum_{i=1}^{n} f(X_T^i) \to_{n\to\infty} \mathbb{E}\left[f(X_T)\right], \mathbb{P}\text{-p.s.}$$

De plus, on obtient une vitesse de convergence d'ordre $n^{-\frac{1}{2}}$ grâce au Théorème de la Limite Centrale. Cependant, pour des modèles sophistiqués, la loi de X_T est en général inconnue, et il n'est donc pas possible de construire un échantillon i.i.d $(X_T^i)_{i\geq 1}$ suivant la loi de X_T. En pratique, on est amené à construire des schémas de discrétisation sur l'EDS (1.4).

Pour $n \geq 1$, on considère une subdivision régulière $\{t_i := \frac{i}{n}T, 0 \leq i \leq n\}$ de l'intervalle $[0, T]$ de

pas $h = \frac{T}{n}$. Le *schéma d'Euler* consiste à approcher la diffusion (1.4) en remplaçant a et b par les fonctions constantes par morceaux

$$\sum_{i=0}^{n-1} a(t_{i-1}, x)\mathbb{1}_{[t_{i-1},t_i)}(t) \text{ and } \sum_{i=0}^{n-1} b(t_{i-1}, x)\mathbb{1}_{[t_{i-1},t_i)}(t), \quad t \in [0, T].$$

On obtient alors le schéma suivant :

$$\begin{aligned} X^n_{t_{i+1}} =& X^n_{t_i} + a(t_i, X^n_{t_i})h + b(t_i, X^n_{t_i})(W_{t_{i+1}} - W_{t_i}), \quad 0 \leq i \leq n-1, \\ X^n_0 =& x. \end{aligned} \tag{1.5}$$

Si de plus on suppose que la fonction b est de classe C^2. En appliquant la formule d'Itô à la fonction σ, on obtient le schéma de Milstein, défini pour $0 \leq i \leq n-1$ par

$$X^n_{t_{i+1}} = X^n_{t_i} + a(t_i, X^n_{t_i})h + b(t_i, X^n_{t_i})(W_{t_{i+1}} - W_{t_i}) + \frac{1}{2}b(t_i, X^n_{t_i})b'(t_i, X^n_{t_i})((W_{t_{i+1}} - W_{t_i})^2 - h). \tag{1.6}$$

Cependant, l'utilisation de ce schéma n'est pas simple en dimension supérieure à cause des aires de Lévy difficiles à simuler en pratique (pour plus de détails voir [69, Section 7.5]).

Afin de quantifier la convergence d'un schéma vers la solution de l'EDS, nous pouvons étudier deux erreurs. La première est appelée erreur de convergence forte en moyenne d'ordre $p > 0$, et est définie par :

$$\mathbb{E}\Big[\, |X^n_{t_n} - X_T|^p \,\Big]^{\frac{1}{p}}.$$

On dit que l'ordre de convergence forte en moyenne p du schéma vers la solution de l'EDS est $\beta > 0$ si c'est le plus grand réel $\gamma > 0$ tel que $\mathbb{E}\Big[\, |X^n_{t_n} - X_T|^p \,\Big]^{\frac{1}{p}} = O(h^\gamma)$.
La seconde est appelée erreur de convergence faible, et est définie par:

$$|\mathbb{E}[X_T] - \mathbb{E}[X^n_{t_n}]|$$

De façon similaire, l'ordre de convergence faible est $\alpha > 0$ si c'est le plus grand réel $\gamma > 0$ tel que $|\mathbb{E}[X_T] - \mathbb{E}[X^n_{t_n}]| = O(h^\gamma)$, et correspond au biais introduit par le schéma de discrétisation du processus X. On peut aussi étudier la convergence forte et faible de $f(X^n_T)$ vers $f(X_T)$ lorsque $n \to \infty$, pour une fonction mesurable f.

Sous certaines conditions de régularité des coefficients a et b de l'EDS (1.4), l'ordre de convergence faible du schéma d'Euler est de $\alpha_{Euler} = 1$ pour des fonctions f à croissance polynomiale et régulières en dehors d'un nombre fini de potentielles discontinuités (voir [77]). L'ordre de convergence forte pour un schéma d'Euler est de $\beta_{Euler} = \frac{1}{2}$ (voir [59]).
Sous d'autres conditions de régularité des coefficients a et b de l'EDS (1.4) (voir les hypothèses ci dessous 1.3.2), l'ordre de convergence du schéma de Milstein est de $\alpha_{Milstein} = 1$ et $\beta_{Milstein} = 1$.

Dans le cadre du modèle CIR, différentes approches ont été proposées concernant la simulation telles que le schéma d'Euler. Cependant, le schéma d'Euler associé au processus X solution de l'E.D.S. (1.2) peut devenir négatif avec une probabilité non nulle. Pour résoudre ce problème en pratique, on peut fixer le processus égal à zéro lorsqu'il devient négatif ("absorption fix") ou le faire réfléchir à l'origine ("reflection fix"). Un aperçu des schémas d'Euler explicites considérés jusqu'à présent dans la littérature peut être trouvé dans [63]. Les preuves classiques de la convergence ne s'appliquent pas au processus CIR puisque le coefficient de diffusion en racine carré n'est pas globalement Lipschitz. Par conséquent, des approches alternatives ont été utilisées pour prouver la convergence forte et faible de différentes discrétisations du processus CIR.

Berkaoui, Bossy et Diop [13] ont montré que la convergence forte est d'ordre $\frac{1}{2}$ pour le schéma symétrisé dans un régime de paramètres restreint. De même, Dereich, Neuenkirch, et Szpruch [28] ont établi la convergence forte d'ordre $\frac{1}{2}$ pour le schéma *"Drift Implicit Euler scheme"* (ou "Backward Euler Scheme" (BEM)). Plus récemment, Alfonsi [5] et Neuenkirch and Szpruch [68] utilisent la transformation de Lamperti afin d'obtenir une EDS pour laquelle le schéma d'Euler converge fortement avec ordre 1.

Concernant le modèle de Heston (1.3), beaucoup d'efforts ont été déployés pour améliorer les méthodes numériques utilisées dans la discrétisation de l'E.D.S. du processus de la volatilité $(V_t)_{t\geq 0}$ et celle du processus d'actif sous-jacent $(S_t)_{t\geq 0}$. Dans le cas où on a des fonctions régulières Altmayer et Neuenkirch [6] ont montré une convergence faible d'ordre 1 pour un schéma numerique sous la condition $\kappa\theta > \sigma^2$ alors que Zheng [81] arrive à montrer que la vitesse de convergence faible est d'ordre 2 pour tous les régimes de paramètres. En revanche, peu de résultats de convergence forte ont été établis pour ce modèle. Nous citons Cozma et al. [23] qui ont prouvé une convergence d'un schéma d'Euler pour un certain régime de paramètres, mais aucun taux de convergence n'a été calculé.

1.2.1 Complexité algorithmique

Dans cette partie, nous rappelons la complexité de calcul de prix d'une option via la méthode de Monte Carlo. Supposons que nous voulons estimer $\mathbb{E}\left[f(X_T)\right]$ le prix d'une option avec une précision ε. Soit $\hat{P} = f(\hat{X}_T^n)$ une approximation de $f(X_T)$ de pas $h = T/n$ et $\hat{Y} := \frac{1}{N}\sum_{i=1}^{N}\hat{P}_i$ la moyenne empirique avec N copies indépendantes de $\hat{P}$. L'erreur quadratique moyenne (MSE) associée est définie par: $MSE = \mathbb{E}\left[\left(\hat{Y} - P\right)^2\right]$.On peut décomposer la MSE comme suit:

$$\begin{aligned}
\mathbb{E}\left[\left(\hat{Y} - \mathbb{E}\left[f(X_T)\right]\right)^2\right] &= \mathbb{E}\left[\left(\hat{Y} - \mathbb{E}(\hat{Y}) + \mathbb{E}(\hat{Y}) - \mathbb{E}\left[f(X_T)\right]\right)^2\right] \\
&= \mathbb{E}\left[\left(\hat{Y} - \mathbb{E}(\hat{Y})\right)^2\right] + \left(\mathbb{E}(\hat{Y}) - \mathbb{E}\left[f(X_T)\right]\right)^2 \\
&\quad - 2\mathbb{E}\left[\left(\hat{Y} - \mathbb{E}(\hat{Y})\right)\left(\mathbb{E}(\hat{Y}) - \mathbb{E}\left[f(X_T)\right]\right)\right] \\
&= \mathbb{E}\left[\left(\hat{Y} - \mathbb{E}(\hat{Y})\right)^2\right] + \left(\mathbb{E}(\hat{Y}) - \mathbb{E}\left[f(X_T)\right]\right)^2 \\
&= \frac{1}{N}\mathrm{Var}(\hat{P}) + \left(\mathbb{E}(\hat{Y}) - \mathbb{E}\left[f(X_T)\right]\right)^2.
\end{aligned}$$

Le premier terme correspond à la variance de l'estimateur et le second terme au biais dû à la discrétisation. Si on a que

$$\left(\mathbb{E}(\hat{Y}) - \mathbb{E}\left[f(X_T)\right]\right) = O(h)$$

alors asymptotiquement, $MSE \approx c_1 N^{-1} + c_2 h^2$, où c_1, c_2 sont des constantes positives, i.e $MSE = O(N^{-1} + h^2)$. Pour avoir une précision totale d'ordre ε, il suffit d'avoir une erreur quadratique moyenne en $O(\varepsilon^2)$, c'est à dire $N = O(\varepsilon^{-2})$ et $h = O(\varepsilon)$. Ainsi, la complexité algorithmique par la méthode de Monte-Carlo en utilisant un schéma d'approximation est $C = O(N \times n) = O(\varepsilon^{-3})$.

1.3 Multilevel Monte Carlo Method

Nous présentons maintenant la méthode Multilevel Monte Carlo (MLMC) proposée par Giles [36], qui est une technique permettant de réduire la complexité du calcul d'espérance tout en obtenant la même précision finale. Inspiré de la méthode à deux niveaux ou méthode de

Romberg statistique introduite par Kebaier [56], Giles [36] propose une méthode d'approximation multi-niveaux.
Soient M et L deux entiers non-nuls quelconques, on note $P := f(X_T)$ la variable aléatoire d'espérance inconnue. Pour tout $\ell = 0, \ldots, L$, on définit un pas de discrétisation $h_\ell = M^{-\ell}$ et on définit également $\hat{X}^{\ell,M}$ le schéma d'Euler avec pas h_l, et $\hat{P}_l := f(\hat{X}_T^{\ell,M})$. Par linéarité de l'espérance, on considère la décomposition télescopique suivante

$$\mathbb{E}[\hat{P}_L] = \mathbb{E}[\hat{P}_0] + \sum_{\ell=1}^{L} \mathbb{E}[\hat{P}_\ell - \hat{P}_{\ell-1}],$$

puis chacune des espérances du membre de droite de la somme ci-dessous est estimée par la méthode de Monte Carlo standard utilisant N_ℓ simulations indépendantes. Pour $1 \leq \ell \leq L$ fixé, on a donc pour $\mathbb{E}[\hat{P}_\ell - \hat{P}_{\ell-1}]$ l'estimateur suivant

$$\hat{Y}_\ell = \frac{1}{N_\ell} \sum_{i=1}^{N_\ell} \left(\hat{P}_\ell^{(i)} - \hat{P}_{\ell-1}^{(i)} \right),$$

où pour chaque $1 \leq i \leq N_\ell$, $\hat{P}_\ell^{(i)} = f(\hat{X}_T^{\ell,M,(i)})$ et $\hat{P}_{\ell-1}^{(i)} = f(\hat{X}_T^{\ell-1,M,(i)})$, où la trajectoire grossière $\hat{X}_T^{\ell-1,M,(i)}$ et la trajectoire fine $\hat{X}_T^{\ell,M,(i)}$ sont simulées en utilisant la même trajectoire Brownienne sous-jacente. De même, l'estimateur de $\hat{P}_0$ s'écrit

$$\hat{Y}_0 = \frac{1}{N_0} \sum_{i=1}^{N_0} \hat{P}_0^{(i)}.$$

Les termes $\hat{P}_\ell^{(i)} - \hat{P}_{\ell-1}^{(i)}$ sont des termes correctifs représentant la différence entre la discrétisation fines de pas h_ℓ, et grossière de pas $h_{\ell-1}$. Cette différence est faible par convergence forte du schéma d'Euler lorsque le pas de discrétisation tend vers 0. De plus, la variance des estimateurs est égale à

$$\mathrm{Var}[\hat{Y}_\ell] = \frac{V_\ell}{N_\ell},$$

où $V_\ell = \mathrm{Var}[\hat{P}_\ell - \hat{P}_{\ell-1}]$. Enfin, on note $\hat{Y} = \sum_{\ell=0}^{L} \hat{Y}_\ell$ l'estimateur du prix du produit financier de payoff $f(X_T^L)$, de sorte que comme les estimateurs $\{\hat{Y}_\ell, l = 0, 1, ..., L\}$ sont indépendants, la variance de l'estimateur global s'écrit $\mathrm{Var}[\hat{Y}] = \sum_{\ell=0}^{L} \frac{V_\ell}{N_\ell}$.

1.3.1 Théorème de complexité

Nous allons maintenant énoncer le théorème de complexité introduit par Giles [36] pour la méthode MLMC.

Theorem 1.3.1. *Soit P le payoff d'un produit financier en fonction de $X_T^{\ell,M}$. On note $\hat{P}_l$ son approximation par un schéma numérique de pas $h_\ell = \frac{T}{M^\ell}$. On suppose qu'il existe des estimateurs indépendants $\{\hat{Y}_\ell, \ell = 0, 1, \ldots, L\}$ issus d'échantillons de tailles respectives $\{N_\ell, \ell = 0, 1, \ldots, L\}$ et des constantes positives β, c_1, c_2, c_3 et $\alpha \geq \frac{1}{2}$ tels que:*

1. $\mathbb{E}[\hat{P}_\ell - P] \leq c_1 h_\ell^\alpha$
2. $\mathbb{E}[\hat{Y}_\ell] = \begin{cases} \mathbb{E}[\hat{P}_0], \\ \mathbb{E}[\hat{P}_\ell - \hat{P}_{\ell-1}] \end{cases}$

3. $Var[\hat{Y}_\ell] \leq c_2 N_\ell^{-1} h_\ell^\beta$

4. C_l, *le coût calculatoire de l'estimateur* $\hat{Y}_\ell$, *vérifie:*

$$C_\ell \leq c_3 N_\ell h_\ell^{-1}$$

Alors, il existe une constante positive c_4, *telle que pour tout* $\varepsilon < e^{-1}$ *il existe des valeurs* L *et* N_ℓ *pour lesquelles l'estimateur Multi-Level*

$$\hat{Y} = \sum_{\ell=0}^{L} \hat{Y}_\ell = \frac{1}{N_0} \sum_{i=1}^{N_0} \hat{P}_0^{(i)} + \sum_{\ell=1}^{L} \frac{1}{N_\ell} \sum_{i=1}^{N_\ell} \left(\hat{P}_\ell^{(i)} - \hat{P}_{\ell-1}^{(i)} \right)$$

a une erreur quadratique moyenne bornée par ϵ^2, *i.e*

$$MSE \equiv \mathbb{E}\left[(\hat{Y} - \mathbb{E}[P])^2 \right] < \varepsilon^2,$$

avec une complexité calculatoire bornée par

$$C \leq \begin{cases} c_4 \varepsilon^{-2}, & \beta > 1 \\ c_4 \varepsilon^{-2} (\log \epsilon)^2, & \beta = 1 \\ c_4 \varepsilon^{-2-(1-\beta)/\alpha}, & 0 < \beta < 1. \end{cases}$$

La complexité C de la méthode MLMC dépend du taux de convergence de la variance de la méthode et du taux de convergence faible. Le paramètre α est un détail mineur, car le Théorème 1.3.1 nécessite que $\alpha > \frac{1}{2}$, ce qui est généralement le cas pour les schémas discrets en temps. Alors que le paramètre β joue un rôle important pour l'efficacité de la méthode MLMC, en effet si $\beta > 1$ la variance des estimateurs $\mathrm{Var}[\hat{Y}_\ell]$ décroît plus rapidement que $O(h_\ell)$ lorsque $h_\ell \to 0$ ce qui réduit la complexité à $O(\epsilon^{-2})$. On note dans le cas où $0 < \beta < 1$, la méthode de "Richardson–Romberg extrapolation" MLMC introduite par Lemaire et Pagès [60] améliore la complexité C en atteignant l'ordre $\varepsilon^{-2}\, \mathrm{e}^{\frac{1-\beta}{\sqrt{\alpha}} \sqrt{2 \log\left(\frac{1}{\varepsilon}\right) \log(M)}} = o(\varepsilon^{-\eta})$ pour tout $\eta > 0$.

1.3.2 Pricing des options exotiques pour des diffusions avec coefficients Lipschitz

Dans la littérature, pour des diffusions avec des coefficients Lipshitz, Giles, Higham and Mao [38] montrent que pour les options barrières la variance V_ℓ d'un bloc ℓ donné dans la méthode Euler MLMC avec un payoff contenant l'indicatrice liée à l'option barrière vérifie

$$V_\ell = O(h_\ell^{\frac{1}{2}-\delta})$$

pour tout $\delta > 0$. Puis, Giles, Debrabant et Rößler [37] démontrent que pour un schéma de Milstein combiné avec une technique de pont Brownien et sous certaines hypothèses de régularité, la variance V_ℓ d'un bloc ℓ donné dans l'estimateur MLMC est d'ordre

$$V_\ell = O(h_\ell^{\frac{3}{2}-\delta}), \quad \forall \delta > 0.$$

Dans la suite, nous détaillons les résultats obtenus dans [37] pour les options barrières. Pour ce faire, on considère l'option "Down-and-Out" de payoff

$$P = (X_T - K)_+ \mathbf{1}_{\{\inf_{t \in [0,T]} X_t > B\}},$$

où X solution de l'E.D.S. suivante

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t, X_0 > B > 0, \tag{1.7}$$

avec W est un mouvement Brownien standard uni-dimensionnel. Soit

$$L^0 \equiv \frac{\partial}{\partial t} + a\frac{\partial}{\partial x},$$
$$L^1 \equiv b\frac{\partial}{\partial x}.$$

Dans ce contexte, Giles et al. [37] ont considéré les hypothèses suivantes :

- A1 $a \in C^{2,1}(\mathbb{R}^+ \times \mathbb{R})$, $b \in C^{3,1}(\mathbb{R}^+ \times \mathbb{R})$ et b ne s'annule pas.
- A2 (Condition de Lipschitz uniforme): il existe une constante K_1 telle que pour tout $x, y \in \mathbb{R}$,

$$|a(t,x) - a(t,y)| + |b(t,x) - b(t,y)| + |L^1b(t,x) - L^1b(t,y)| \leq K_1|x - y|$$

- A3 (Limite de croissance linéaire): il existe une constante K_2 telle que pour tout $x \in \mathbb{R}$ on a

$$\begin{aligned}|a(t,x)|+|L^0a(t,x)| + |L^1a(t,x)| + |b(t,x)| \\ + |L^0b(t,x)| + |L^1b(t,x)| + |L^0L^1b(x,t)| + |L^1L^1b(t,x)| \leq K_2(1 + |x|)\end{aligned}$$

- A4 (Condition de Lipschitz par rapport au temps t): il existe une constante K_3 telle que pour tout $x \in \mathbb{R}$ et $s, t \in \mathbb{R}^+$ on a

$$|b(t,x) - b(s,x)| \leq K_3(1 + |x|)\sqrt{|t - s|}.$$

Puis, ils ont introduit le schéma de Milstein interpolé donné par :

$$\hat{X}_t = \hat{X}_{t_i} + \frac{(t - t_i)}{h}(\hat{X}_{t_{i+1}} - \hat{X}_{t_i}) + b(t_i, X_{t_i})\Big(W_t - W_{t_i} - \frac{(t - t_i)}{h}(W_{t_{i+1}} - W_{t_i})\Big), t \in [t_i, t_{i+1}] \tag{1.8}$$

avec $(\hat{X}_{t_i})_{1\leq i\leq n}$ le schéma de Milstein discret (1.6). Afin d'éviter le biais introduit dans l'approximation de l'indicatrice du payoff P, ils utilisent la technique du pont Brownien introduite dans [46] afin de déterminer explicitement $q_i := 1 - p_i$ la probabilité que le schéma interpolé $\hat{X}_t$ ne franchisse pas la barrière dans chaque petit intervalle de temps $[t_i, t_{i+1}]$ sachant les points X_{t_i} et $X_{t_{i+1}}$ où

$$p_i = \exp\Big(\frac{-2(\hat{X}_{t_i} - B)_+(\hat{X}_{t_{i+1}} - B)_+}{b(t_i, X_{t_i})^2 h}\Big).$$

Ainsi, le prix $\mathbb{E}[P]$ est approché par

$$\begin{aligned}&\mathbb{E}\left[(\hat{X}_T - K)_+\mathbf{1}_{\{\inf_{t\in[0,T]}\hat{X}_t>B\}}\Big|\hat{X}_0, \hat{X}_{t_1}, ..., \hat{X}_{t_n}\right] \\ &= (\hat{X}_T - K)_+\mathbb{E}\left[\mathbf{1}_{\{\inf_{t\in[0,T]}\hat{X}_t>B\}}\Big|\hat{X}_0, \hat{X}_{t_1}, ..., \hat{X}_{t_n}\right] \\ &= (\hat{X}_T - K)_+\mathbb{E}\left[\prod_{i=0}^{n-1}\mathbf{1}_{\{\inf_{t\in[t_i,t_{i+1}]}\hat{X}_t>B\}}\Big|\hat{X}_0, \hat{X}_{t_1}, ..., \hat{X}_{t_n}\right] \\ &= (\hat{X}_T - K)_+\prod_{i=0}^{n-1}\mathbb{E}\left[\mathbf{1}_{\{\inf_{t\in[t_i,t_{i+1}]}\hat{X}_t>B\}}\Big|\hat{X}_{t_i}, \hat{X}_{t_{i+1}}\right] \\ &= (\hat{X}_T - K)_+\prod_{i=0}^{n-1}(1 - p_i).\end{aligned}$$

Pour mettre en place la méthode MLMC, Giles et al. [37] considèrent $\hat{X}_t^{2^\ell}$ le schéma interpolé défini en (1.8) avec un pas de discrétisation $h_\ell = 2^{-\ell}T$ pour $\ell \in \{0, ..., L\}$, où $L = \dfrac{\log n}{\log 2}$ et n désigne le nombre de pas de temps le plus fin. D'abord, on commence par le niveau fin, on simule les trajectoires $\hat{X}_0, \hat{X}_{t_1^\ell}^{2^\ell}, ..., \hat{X}_{t_n^\ell}^{2^\ell}$ et on utilise la technique du pont Brownien pour le schéma interpolé $\hat{X}_t^{2^\ell}$ en appliquant simplement la formule précédente pour définir

$$\hat{P}_\ell^f = (\hat{X}_T^{2^\ell} - K)_+ \prod_{i=0}^{2^\ell - 1} (1 - \hat{p}_i^{2^\ell}) \text{ avec}$$

$$p_i^{2^\ell} = \exp\Big(\frac{-2(\hat{X}_{t_i^\ell}^{2^\ell} - B)_+ (\hat{X}_{t_{i+1}^\ell}^{2^\ell} - B)_+}{b(t_i^\ell, \hat{X}_{t_i^\ell}^{2^\ell})^2 h_\ell}\Big),$$

où $t_i^\ell = \frac{iT}{2^\ell}$ pour $\ell \in \{0, ..., L\}$. Puis on passe aux pas grossiers. Après avoir simulé les $\hat{X}_0, \hat{X}_{t_1^\ell}^{2^\ell}, ..., \hat{X}_{t_n^\ell}^{2^\ell}$ à chaque pas du temps grossier $[t_i^{\ell-1}, t_{i+1}^{\ell-1}]$, on simule une valeur intermédiaire supplémentaire qui correspond à la valeur du schéma interpolé en t_{2i+1}^ℓ. Par conséquent, la probabilité de dépasser la barrière dans l'intervalle $[t_i^{\ell-1}, t_{i+1}^{\ell-1}]$ est $p_i^{2^{\ell-1}}$ est égale à la probabilité conditionnelle que le schéma interpolé $\hat{X}_t^{2^{\ell-1}}$ n'atteigne pas la barrière sur les pas de temps fins $[t_i^{\ell-1}, t_{2i+1}^\ell]$ et $[t_{2i+1}^\ell, t_{i+1}^{\ell-1}]$, conditionnellement à t_{2i+1}^ℓ. Donc on peut écrire

$$\hat{P}_{\ell-1}^c = (\hat{X}_T^{2^{\ell-1}} - K)_+ \prod_{i=0}^{2^\ell - 1} (1 - \overline{p}_i^{2^{\ell-1}}) \tag{1.9}$$

$$= (\hat{X}_T^{2^{\ell-1}} - K)_+ \prod_{i=0}^{2^{\ell-1} - 1} (1 - \overline{p}_{i,1}^{2^{\ell-1}})(1 - \overline{p}_{i,2}^{2^{\ell-1}}) \text{ with} \tag{1.10}$$

$$\hat{p}_{i,1}^{2^{\ell-1}} = \exp\Big(\frac{-2(\hat{X}_{t_i^{\ell-1}}^{2^{\ell-1}} - B)_+ (\hat{X}_{t_{2i+1}^\ell}^{2^{\ell-1}} - B)_+}{b(t_i^{\ell-1}, \hat{X}_{t_i^{\ell-1}}^{2^{\ell-1}})^2 h_\ell}\Big),$$

$$\hat{p}_{i,2}^{2^{\ell-1}} = \exp\Big(\frac{-2(\hat{X}_{t_{2i+1}^\ell}^{2^{\ell-1}} - B)_+ (\hat{X}_{t_i^{\ell-1}}^{2^{\ell-1}} - B)_+}{b(t_i^{\ell-1}, \hat{X}_{t_i^{\ell-1}}^{2^{\ell-1}})^2 h_\ell}\Big).$$

L'estimateur MLMC proposé est alors donné par

$$\frac{1}{N_0}\sum_{i=1}^{N_0} \hat{P}_0^{c,(i)} + \sum_{\ell=1}^{L} \frac{1}{N_\ell} \sum_{i=1}^{N_\ell} \Big(\hat{P}_\ell^{f,(i)} - \hat{P}_{\ell-1}^{c,(i)}\Big)$$

démontrent que $\text{Var}\Big(\hat{P}_\ell^f - \hat{P}_{\ell-1}^c\Big) = O(h_\ell^{\frac{3}{2}-\delta})$ pour tout $\delta > 0$. Ainsi, par le Théorème 1.3.1 cet estimateur MLMC permet d'obtenir la complexité optimale $O(\varepsilon^{-2})$.

1.4 Méthode GPR

En pratique, l'évaluation des prix d'options barrières par la méthode MC ou la méthode MLMC peut peut s'avérer coûteuses en temps. Récemment, des techniques de "Machine Learning" (ML) ont fait leurs preuves pour l'accélération de certaines procédures numériques tels que l'évaluation des prix d'options exotiques (voir e.g. [26]) ou la calibration de modèles (voir e.g.

[71], [2]). Plus précisément, les techniques d'apprentissage automatique peuvent être en mesure de résoudre ces problèmes et de récupérer plus rapidement les prix des options ainsi que les ratios de couverture, tout en conservant une précision raisonnable. Ces dernières années, de nombreuses recherches ont été menées pour améliorer les méthodes traditionnelles de modélisation la plupart de ses applications concernant la modélisation des risques et la construction de portefeuille [31].

Dans cette thèse, nous nous intéressons plus particulièrement à la méthode d'apprentissage automatique bayésienne : "Gaussian Process Regression" (GPR) pour la tarification et la couverture des options barrières. Cette méthode s'appuie sur une estimation non paramétrique pour évaluer les prix et calculer les ratios de couverture lorsque les paramètres (sous-jacents) au modèle ne sont pas spécifiés. De nombreux travaux ont été développés dans le cadre d'application de la méthode GPR pour la finance quantitative. Ludkovski [64] applique la méthode GPR pour étudier le problème du temps d'arrêt optimal d'une option de vente bermudienne. En utilisant différents type de noyaux : Matern $\frac{5}{2}$, Matern $\frac{3}{2}$ et des noyaux exponentiels, au lieu des noyaux fixes pour estimer les valeurs de continuation de l'option bermudienne. Il conclut que cette approche permet une réduction importante du budget de calcul. En outre, l'un des principaux avantages de la méthode GPR est que les sensibilités associées aux calculs menés sont automatiquement obtenu après la phase d'apprentissage. De Spiegeleer et al. [26] ont appliqué la méthode GPR pour estimer et ajuster des produits dérivées. Ils illustrent sur des exemples numérique l'efficacité de la méthode GPR comparée aux méthodes de Monte Carlo lors de l'évaluation des prix d'options exotiques et de calculs des sensibilités associées tout en obtenant une perte de précision raisonnable d'un point de vue pratique. Également, Crépey and Dixon [24] ont étendu le travail de De Spiegeleer et al. [26] en appliquant le GPR a un portefeuille d'options, ce qui leur a notamment permis d'optimiser le calculs des CVA. Goudenège et al. [48]. En outre, [48] appliquent la méthode GPR de manière similaire à Ludkovski [64] en mettant l'accent sur les options "Americain Basket put options". Ils concluent que la méthode de GPR est plus performante que la méthode de Monte Carlo.

1.5 Contributions de la thèse

Dans le Chapitre 2 de la thèse, nous nous intéressons à l'étude de la méthode MLMC pour la valorisation d'options barrières pour des modèles avec des coefficients de diffusion non-Lipschitz. Dans ce cadre et à l'aide de techniques de ponts Browniens combinées astucieusement avec la méthode MLMC, nous démontrons que la complexité de cette dernière peut atteindre le régime optimal d'ordre ϵ^{-2} pour une tolérance ϵ donnée et ce en dépit du coefficient de diffusion non-Lipschitz pouvant ralentir significativement la vitesse du schéma de discrétisation associé. Ces nouveaux résultats ont été appliqués dans le cadre du modèle CIR et du modèle à volatilité locale CEV. Plus précisément, on considère X la solution de l'E.D.S. définie sur $I = (c, +\infty)$, $c \geq 0$ par

$$\begin{cases} dX_t &= b(X_t)dt + \sigma(X_t)dW_t, \quad t \geq 0, \\ X_0 &= x \in I, \end{cases} \tag{1.11}$$

avec W un mouvement Brownien standard, $b : \mathbb{R} \to \mathbb{R}$ mesurable et $\sigma : \mathbb{R} \to (0, \infty)$ continue mais non globalement Lipschitz. En utilisant la transformation de Lamperti

$$\psi(x) = \int_0^x \frac{1}{\sigma(z)} dz,$$

on obtient l'E.D.S. suivante pour $Y = \psi(X)$

$$\left\{ \begin{array}{rcl} dY_t & = & L(Y_t)dt + \gamma dW_t, \quad t \geq 0, \\ Y_0 & = & y \in I, \text{for} \quad \gamma > 0. \end{array} \right. \tag{1.12}$$

On suppose que le coefficient de dérive L satisfait la condition de monotonie suivante :

$$L : I \longrightarrow \mathbb{R} \text{ est } C^2 \text{ et telle que } \exists \kappa > 0, \ \ \forall y, y' \in I, y \leq y', L(y') - L(y) \leq \kappa(y' - y). \tag{1.13}$$

De plus, on suppose que pour un point arbitraire $d \in I$,

$$v(x) = \int_d^x \int_d^y \exp\Big(- \frac{2}{\gamma^2} \int_z^y L(\xi) d\xi \Big) dz dy \ \text{ satisfait } \lim_{x \to 0^+} v(x) = +\infty. \tag{H1}$$

D'une part, les conditions de Feller (voir e.g. [76]) assurent que l'E.D.S. (1.12) admet une unique solution $(Y_t)_{t \geq 0}$ sur I. D'autre part, sous ces deux conditions, le schéma continu implicite introduit par Alfonsi [5]

$$\begin{aligned} &\widehat{Y}_t^n = \widehat{Y}_{t_i}^n + L(\widehat{Y}_t^n)(t - t_i) + \gamma(W_t - W_{t_i}), \text{ with } t \in (t_i, t_{i+1}], t_i = \frac{iT}{n}, \ 0 \leq i \leq n-1, \\ &\widehat{Y}_0^n = y, \end{aligned} \tag{1.14}$$

est bien défini pour tout $n \geq 1$ et pour tout $t \in [0, T]$, $Y_t^n \in I$. Dans la littérature, Dereich, Neuenkirch et Szpruch [28] ont démontré que la convergence forte est d'ordre $\frac{1}{2}$ pour ce schéma. Leur résultat n'est toutefois pas optimal. Alfonsi [5] a obtenu une vitesse de convergence forte d'ordre 1 en supposant que pour $p \geq 1$, on a

$$\mathbb{E}\Big[\Big(\int_0^T |L'(Y_u)L(Y_u) + \frac{\gamma^2}{2} L''(Y_u)| du \Big)^p \Big] < \infty \ \ et \ \ \mathbb{E}\Big[\Big(\int_0^T (L'(Y_u))^2 du \Big)^{\frac{p}{2}} \Big] < \infty. \tag{H2}$$

Nous introduisons une interpolation légèrement différente du schéma implicite (1.14), défini pour tout $n \geq 1$ par

$$\overline{Y}_t^n = \widehat{Y}_{t_i}^n + L(\widehat{Y}_{t_{i+1}}^n)(t - t_i) + \gamma(W_t - W_{t_i}), \text{ pour } \ t \in [t_i, t_{i+1}] \text{ et pour } 0 \leq i \leq n-1. \tag{1.15}$$

L'avantage principal de cette interpolation est qu'elle est compatible avec la technique du pont Brownien permettant d'atténuer le biais d'approximation pour la valorisation des options barrières. De plus, nous introduisons une hypothèse renforcée sur le coefficient de dérive L comme suit

$$\begin{aligned} &L : I = (0, \infty) \to \mathbb{R} \in \mathcal{C}^2(I) \text{ est décroissante sur } (0, A) \text{ pour un } A > 0, \\ &\qquad \text{et est globalement Lipschitz avec constante de Lipschitz } \kappa > 0 \text{ sur } (A, \infty). \end{aligned} \tag{H3}$$

Sous les hypothèses, (H2) et (H3) nous montrons que le schéma (1.15) a une vitesse de convergence d'ordre 1 également (voir Théorème 2.2.1).

Dans ce contexte, en vue de l'étude de l'estimateur MLMC combiné avec la technique du pont Brownien sur le schéma (1.15) pour l'évaluation des options barrières, nous développons des estimés permettant de contrôler les trajectoires extrêmes (voir Lemme 2.3.2). En se basant sur ce lemme, et en s'inspirant de Giles et al. [37], nous partageons l'ensemble des trajectoires en deux catégories. La première sera la partie des trajectoires qui se comporte "bien", c'est-à-dire qu'entre deux pas de temps de discrétisations consécutifs dans la grille "multilevel", les incréments restent relativement petits, ce sont les trajectoires les plus courantes. Contrairement aux trajectoires

"extrêmes" qui peuvent avoir de plus grands accroissements mais qui sont plus beaucoup rares. Une fois cette disjonction terminée, le principe est d'étudier le comportement de chaque quantité concernant les évènements non-extrêmes et d'évaluer l'erreur forte résultant de la discrétisation. Plus précisément, en supposant de plus que le coefficient de dérive vérifie

$$\exists\, \alpha > 0 \text{ tel que } \forall\, y \in I,\ \ yL(y) \leq \alpha(1 + |y|^2), \tag{H4}$$

nous démontrons que la variance V_ℓ d'un bloc ℓ donné dans l'estimateur MLMC est d'ordre

$$V_\ell = O(h_\ell^{1+\delta}),\ \ \forall \delta \in (0, \frac{1}{2}),\ \text{pour } h_\ell \text{ suffisamment petit},$$

à condition que $\inf_{t\in[0,T]} Y_t$ (resp. $\sup_{t\in[0,T]} Y_t$) admette une densité continue au voisinage de la barrière (voir Théorème 2.3.3). Ce dernier résultat conduit à une méthode MLMC de complexité optimale $O(\varepsilon^{-2})$. Ensuite, nous appliquons ces résultats pour les modèles X^{CIR} solution de (1.2) et X^{CEV} solution de (1.1). Pour ce faire nous obtenons des formules semi-analytiques pour la densité de :

- $\inf_{t\in[0,T]} X_t^{CIR}$, voir Théorème 2.4.2,
- $\sup_{t\in[0,T]} X_t^{CIR}$, voir Théorème 2.4.1,
- $\inf_{t\in[0,T]} X_t^{CEV}$, voir Théorème 2.5.3,
- $\sup_{t\in[0,T]} X_t^{CEV}$, voir Théorème 2.5.2.

En appliquant, le Théorème 2.3.3 aux modèles CIR et CEV nous obtenons des restrictions sur le choix des paramètres du modèles (voir sections 2.4 et 2.5). Cependant, les illustrations numériques montrent que la variance V_ℓ d'un bloc ℓ donné dans l'estimateur MLMC est d'ordre $O(h_\ell^{1+\delta})$ avec $\delta > 0$ proche de zéro même pour un choix des paramètres qui ne respecte pas forcément nos conditions théoriques.

Dans le chapitre 3 de la thèse, nous nous intéressons à l'extension du premier travail au modèle de log-Heston $(X_t)_{t\geq 0}$ défini par

$$\begin{aligned} dX_t &= (r - \frac{1}{2}V_t)dt + \sqrt{V_t}\ \ d\left(\rho W_t^v + \sqrt{1-\rho^2}W_t^s\right), \\ dV_t &= \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dW_t^v. \end{aligned} \tag{1.16}$$

Le but de ce chapitre est de définir et d'étudier une approximation numérique de type MLMC du prix d'options barrières. Plus particulièrement, les options "Down-and-Out" et "Up-and -out".

$$\pi_D = \mathbb{E}\Big[f(X_T)\mathbb{1}_{\{\tau_D > T\}}\Big] \text{ et } \pi_U = \mathbb{E}\Big[f(X_T)\mathbb{1}_{\{\tau_U > T\}}\Big], \tag{1.17}$$

où f désigne une fonction payoff et τ_D et τ_U les premiers temps de passage définis par :

- $\tau_D = \inf\{t \in [0,T], X_t \leq D\}$ avec $y > D > 0$, pour l'option "Down-and-Out" (D-O),

  et

- $\tau_U = \inf\{t \in [0,T], X_t \geq U\}$ avec $0 < y < U$, pour l'option "Up-and-Out" (U-O).

Afin de pouvoir utiliser la technique du pont Brownien, nous avons introduit un schéma d'approximation interpolé $\overline{X}_t^n$ du processus log-Heston X basé sur une simulation exacte du processus de variance V et donné par

$$\begin{aligned}\overline{X}_t^n = \overline{X}_{t_i}^n + (r - \frac{1}{2}V_{t_i})(t-t_i) + \frac{\rho}{\sigma}\Big(\frac{n}{T}(V_{t_{i+1}} - V_{t_i})(t-t_i) - \kappa\theta(t-t_i) + \kappa V_{t_i}(t-t_i)\Big) \\ + \sqrt{1-\rho^2}\sqrt{V_{t_i}}(W_t^s - W_{t_i}^s), \text{ pour } t \in [t_i, t_{i+1}[, \end{aligned} \tag{1.18}$$

en utilisant que conditionnellement à V_u, le processus de variance V_t à la date $t > u$, a la même densité qu'une loi de Chi-2 non-centrée définie par

$$V_t \overset{d}{=} \frac{\sigma^2(1-\mathrm{e}^{-\kappa(t-u)})}{4\kappa}\chi_d^2\Big(\frac{4\kappa\,\mathrm{e}^{-\kappa(t-u)}}{\sigma^2(1-\mathrm{e}^{-\kappa(t-u)})}V_u\Big),$$

où $d := \frac{4\theta\kappa}{\sigma^2}$ et $\chi_d^2(\lambda)$ désigne une variable aléatoire de loi de Chi-2 non-centrée avec degrés de liberté d et paramètre de décentralisation λ. Cette approximation est ensuite couplée aux techniques MLMC présentées précédemment, dont nous étudions le comportement de la variance ainsi que de la complexité dans la section 3.5. Pour ce faire, nous avons tout d'abord montré que la convergence forte du schéma $\overline{X}^n$ est d'ordre $h^{\frac{1}{2}}$ avec $h = \frac{T}{n}$, ensuite nous avons développé un contrôle fin des trajectoires des processus X, $\overline{X}^n$ et V dans le but d'analyser les contributions des trajectoires extrêmes et non-extrêmes. Sous la condition que $\inf_{t\in[0,T]} X_t$ (resp. $\sup_{t\in[0,T]} X_t$) admette une densité continue au voisinage de la barrière D pour l'option D-O (resp. U pour l'option U-O) et pour une fonction de payoff Lipschitz f, nous obtenons que la variance V_ℓ d'un bloc ℓ donné dans l'estimateur MLMC est d'ordre $O(h_\ell^\beta)$ où $\beta \simeq \frac{1}{2}$. Pour une précision donnée, $\varepsilon > 0$ cela entraîne une complexité $\varepsilon^{-2-(1-\beta)/\alpha}$ où α représente l'ordre de convergence faible, ce qui est loin d'être le régime optimal. La principale raison est que la convergence forte du schéma $\overline{X}^n$ avec $n^{\frac{1}{2}}$ ralentit significativement la méthode MLMC et ceci est illustré par les tests numériques effectués pour le modèle de Heston (voir Section 3.7). Cependant, à notre connaissance il n'y a pas d'étude dans la littérature traitant du comportement asymptotique de la méthode MLMC combinée avec la technique du pont Brownien pour l'évaluation des options barrières pour les modèles d-dimensionnels avec $d > 1$. De plus, nous n'avons pas trouvé dans la littérature d'autres schémas compatibles avec la technique du pont Brownien et pour lesquels l'analyse fine en trajectoires extrêmes et non-extrêmes soit possible. Nous avons également testé la méthode de MLMC antithétique proposée par Giles et Szpruch [40] dans le modèle Clarck-Cameron, mais les tests numériques montrent que l'ordre de la variance de la méthode n'est pas amélioré (Voir sections 3.6 et 3.7). Par ailleurs, nous étudions le problème de l'explosion des moments pour le modèle de Heston, en prolongement des résultats obtenus par Cozma et al.[23].

Dans le chapitre 4 de la thèse, nous nous intéressons à l'apprentissage automatique particulièrement de la méthode GPR pour accélérer l'évaluation des prix des options barrières tout en conservant une précision raisonnable, ce qui constitue de nos jours un enjeu important en finance quantitative. Dans un premier temps, nous utilisons la méthode Monte Carlo pour l'évaluation de prix d'options barrières de type DO pour différents schémas d'approximation dans le modèle de Heston avec/sans la technique de pont Brownien (voir Section 4.4). Le modèle est entraîné sur un ensemble d'entraînement $X := (X_i)_{i=1,\dots,n} = (K^{(i)}, V_0^{(i)}, \sigma, T^{(i)}, r^{(i)}, \theta^{(i)}, \kappa^{(i)}, \rho^{(i)})_{i=1,\dots,n}$ tiré uniformément selon des plages de paramètres prédéfinies. Les prix Monte Carlo des options DO sont calculés pour chaque ligne de X, ce qui conduit au vecteur de sortie Y qui représente les prix des options DO cibles. Une fois la procédure d'apprentissage terminée, le modèle est prêt pour la prédiction. La performance de la méthode GPR est mesurée en terme de AAE, MAE et RMSE.

Les tests numériques confirment que la méthode GPR entraînée accélère considérablement la procédure d'évaluation des prix d'options barrières sans perte significative de précision. Des résultats similaires ont été observés dans [26] mais pour une méthode GPR entrainée par une méthode Monte Carlo sans technique de pont Brownien. Dans nos tests numériques, nous avons observé une amélioration dans la précision de la méthode GPR quand elle est entraînée par des méthodes Monte Carlo combinées avec des techniques de pont Brownien (voir la Section 4.4 pour plus de précisions). En outre, nous développons de nouvelles formules analytiques pour calculer les sensibilités associées aux options barrières. Ces formules nous ont permis de développer de nouveaux estimateurs "Pathwise MC" pour chacun des schémas d'approximation du modèle de log-Heston étudiés (voir par exemple [43] ou [69] pour plus de détails sur la méthode "Pathwise MC"), afin de quantifier l'incertitude liée à l'utilisation des dérivés automatiques fournies systématiquement par la méthode GPR (voir par exemple [29, section 5.1]). Il s'avère que lorsqu'on utilise nos estimateurs "Pathwise MC" comme valeurs de références, la sensibilité Δ d'une option D-O calculée par la méthode GPR est relativement précise. Mais en revanche, la sensibilité ν de la même option évaluée par cette dernière méthode semble beaucoup moins précise (voir 4.4 pour plus de détails).

1.6 Conclusion et Perspectives

Les produits dérivés jouent un rôle clé dans les marchés financiers car ils permettent de financer et de couvrir les risques financiers des acteurs des marchés financiers. Le calcul des prix des produits dérivés ainsi que leurs indicateurs de risque est un enjeux majeur pour les différents intervenants dans les marchés financiers. Afin de calculer les prix des produits financiers et leurs indicateurs de risques, des méthodes numériques, comme la méthode Monte-Carlo, sont utilisés. Cependant, ces méthodes sont coûteuses en terme de temps de calcul. Sa complexité est d'ordre $O(\varepsilon^{-3})$. La méthode Multilevel permet de réduire le temps de calcul dans divers contextes. Notamment, en ce qui concerne les prix des options dépendant de la trajectoire, ceux-ci peuvent atteindre une complexité de $O(\varepsilon^{-2})$ dans les cas optimaux. Une autre alternative permettant d'accélérer le temps de calcul des prix d'options consiste à considérer des techniques d'intelligence artificielle telle que la méthode de régression par processus gaussiens.

Dans cette thèse, nous montrons que les techniques de MLMC peuvent être utilisées pour l'évaluation des prix d'options barrières dans des modèles avec des coefficients de diffusion non nécessairement globalement Lipschitz et cela en les combinant minutieusement avec les techniques de pont Brownien. Dans ce contexte, nous considérons un cadre général de modèles avec des coefficients de diffusion pas forcément globalement Lipschitz et nous utilisons une transformation de Lamperti pour se ramener à un nouveau processus avec un bruit additif. Ensuite, nous introduisons un schéma d'Euler implicite interpolé pour lequel nous prouvons un résultat de convergence forte d'ordre un. Puis, en analysant les trajectoires extrêmes du processus de diffusion et de son approximation, nous prouvons que la méthode MLMC atteint son régime optimal. Nous appliquons ces résultats pour l'évaluation des prix d'options barrières dans le modèle CIR et le modèle à volatilité locale CEV et développons des formules semi-analytiques pour les densités du minimum et du maximum de ces deux processus.

Ensuite, nous étendons cette approche pour le problème plus difficile d'évaluation de prix d'options barrières dans le modèle log-Heston. Nous concevons un schéma d'approximation pour lequel on est en mesure d'analyser la variance de la méthode MLMC combinée avec des techniques de ponts Brownien. Cette analyse nécessite le contrôle des trajectoires extrêmes

du modèle log-Heston, du processus de variance associé ainsi que du schéma d'approximation considéré. Nous montrons que la méthode MLMC obtenu a une complexité d'ordre inférieur à celle d'une méthode Monte Carlo mais sans atteindre le régime optimal $O(\varepsilon^{-}2)$. Une comparaison numérique avec la méthode MLMC antithétique montre que cette dernière n'atteint pas non plus le régime optimal pour l'évaluation des options barrières par des techniques de pont Brownien. Une étude de problème de l'explosion des moments pour le modèle de Heston, en prolongeant des résultats obtenus par Cozma et al. [23].

Dans la dernière partie de la thèse, nous nous intéressons à l'utilisation de la méthode GPR pour le calcul des prix et des sensibilités d'une option barrière D-O dans le modèle de Heston. Pour ce faire, nous introduisons trois différents schémas d'approximation du modèle de Heston. En utilisant la technique du pont Borwnien, nous observons que les prix obtenus par la méthode GPR entraînée par ces schémas sont plus précis en terme de RMSE que ceux d'une méthode GPR entraînée par une méthode de Monte Carlo classique en gardant l'indicatrice dans le payoff de l'option. Ensuite, nous développons des formules analytiques permettant de calculer les sensibilités Δ et ν de l'option barrière DO pour chacun des trois schémas d'approximation du modèle de Heston par la méthode "Pathwise Monte Carlo". Ces procédures nous permettent de quantifier l'incertitude liée à l'utilisation des formules de calcul de sensibilités fournies automatiquement par une méthode GPR entraînée. Nos tests numériques montrent que pour le calcul du Δ la méthode GPR reste assez précise alors que pour le calcul du ν elle l'est beaucoup moins.

Par ailleurs, ce travail de recherche a soulevé de nouvelles perspectives qui devront être explorées à l'avenir.

- La conception d'un schéma d'approximation du modèle log-Heston ou Heston pour lequel on est capable démonter que la variance de la méthode MLMC associée par des techniques de pont Brownien atteint le régime optimal $O(\varepsilon^{-2})$.
- L'étude de la méthode de "Multilevel Richardson-Romberg extrapolation" combinée avec les techniques de pont Brownien pour l'évaluation des options barrières dans le modèle de Heston.
- Extension des résultats obtenus pour d'autres options exotique telles que les options Lookback, les options digitales, et les options auto-collables [34].
- Améliorer la précision des méthodes GPR pour le calcul de la sensibilité ν pour les options barrières et plus généralement les options dépendant de la trajectoire.

Chapter 2

The interpolated drift implicit Euler scheme Multilevel Monte Carlo method for pricing Barrier options and applications to the CIR and CEV models

Abstract

Recently, Giles et al. [37] proved that the efficiency of the Multilevel Monte Carlo (MLMC) method for evaluating Down-and-Out barrier options for a diffusion process $(X_t)_{t\in[0,T]}$ with globally Lipschitz coefficients, can be improved by combining a Brownian bridge technique and a conditional Monte Carlo method provided that the running minimum $\inf_{t\in[0,T]} X_t$ has a bounded density in the vicinity of the barrier. In the present work, thanks to the Lamperti transformation technique and using a Brownian interpolation of the drift implicit Euler scheme of Alfonsi [5], we show that the efficiency of the MLMC can be also improved for the evaluation of barrier options for models with non-Lipschitz diffusion coefficients under certain moment constraints. We study two example models: the Cox-Ingersoll-Ross (CIR) and the Constant of Elasticity of Variance (CEV) processes for which we show that the conditions of our theoretical framework are satisfied under certain restrictions on the models parameters. In particular, we develop semi-explicit formulas for the densities of the running minimum and running maximum of both CIR and CEV processes which are of independent interest. Finally, numerical tests are processed to illustrate our results.

2.1 Introduction

Barrier options are one of the most widely traded exotic options in the financial markets. Pricing and hedging such path-dependent option can quickly become very challenging especially when we need to achieve a good precision for the approximation. Evaluating barrier options by a classic Monte Carlo method introduces a systematic bias when approximating the continuous running maximum (resp. minimum) in the crossing-barrier indicator by a discrete running maximum (resp. minimum). To overcome this difficulty, several numerical strategies exist in the literature among them the popular Brownian bridge technique introduced in [8] well known for its efficiency and ease of use (see also [47] for related refinements). The Brownian bridge technic uses an analytic expression for the probability of hitting the barrier between two known values in a simulated path of the underlying asset. More recently, a combination of the Multilevel Monte Carlo (MLMC) method with the Brownian bridge technique has been developed in [37] for pricing barrier options. The Multilevel Monte Carlo method introduced in Giles [36] as an extension of the two-level Monte Carlo method of [56], significantly reduces the time complexity of the classical Monte Carlo method. More precisely, for a given precision $\varepsilon > 0$ and a Lipschitz payoff function, if the underlying asset $(X_t)_{t\in[0,T]}$ is approximated using a discretization scheme $(\bar{X}_t)_{t\in[0,T]}$ with time step $h > 0$ satisfying $\mathbb{E}|X_t - \bar{X}_t|^2 = O(h^\beta)$ and $|\mathbb{E}[X_t - \bar{X}_t]| = O(h^\alpha)$ with $\alpha \geq \frac{1}{2}$, then the time complexity of the MLMC methods is: $O(\varepsilon^{-2})$ when $\beta > 1$, $O(\varepsilon^{-2}(\log \varepsilon)^2)$ when $\beta = 1$ and $O(\varepsilon^{-2-\frac{1-\beta}{\alpha}})$ when $\beta \in (0, 1)$. However, for the same precision $\varepsilon > 0$ the optimal time complexity of a classic Monte Carlo method is $O(\varepsilon^{-3})$. As the payoff function of a barrier option is not Lipschitz, Giles et al. [37] take advantage of the Brownian bridge to run the MLMC method for pricing such options, since this technique substitutes the barrier-crossing indicators by the probabilities that the approximation scheme $(\bar{X}_t)_{t\in[0,T]}$ hits the barrier between each two consecutive discretization times t_i and t_{i+1} and which are represented as smooth functions of the realized points $\bar{X}_{t_i}$ and $\bar{X}_{t_{i+1}}$. More precisely, Giles et al. [37] consider an underlying asset solution to a one-dimensional stochastic differential equation (SDE) with globally Lipschitz smooth coefficients that is approximated by a high order strong approximation scheme namely the Milstein scheme $(\bar{X}_t^{\text{Milstein}})_{t\in[0,T]}$ that satisfies $\mathbb{E}|X_t - \bar{X}_t^{\text{Milstein}}|^2 = O(h^2)$. For this case, they prove that the MLMC method reaches its optimal time complexity $O(\varepsilon^{-2})$ for pricing a Down-and-Out barrier option [1] provided that $\inf_{t\in[0,T]} X_t$ has a bounded density in the neighborhood of the barrier. This latter condition cannot be easily checked even when the SDE coefficients are Lipschitz except for very specific cases.

In the current paper, we are interested in studying the MLMC method for pricing barrier options when the underlying asset is solution to a SDE with a non-Lipschitz diffusion coefficient such as the popular Cox-Ingersoll-Ross (CIR) and the Constant of Elasticity of Variance (CEV) processes. Only few works exist in the literature that studied the problem of pricing path-dependent options under such singular models (see e.g. [25]). To analyze the performance of the MLMC method, we consider in Section 2.2 a general framework of models with non-Lipschitz diffusion coefficients and use a Lamperti transformation to focus our study on a new process $(Y_t)_{t\in[0,T]}$ with an additive noise diffusion but in counterpart with a possibly singular drift coefficient L. Then, we introduce a Brownian interpolation scheme $(\bar{Y}_t)_{t\in[0,T]}$ associated to the drift implicit Euler scheme of Alfonsi [5] for which we prove a strong convergence result with order one (see Theorem 2.2.1). In Section 2.3, we use the Brownian bridge technic that substitutes the crossing-indicators with smooth functions of realized points in the path of the scheme $(\bar{Y}_t)_{t\in[0,T]}$ to build the corresponding MLMC estimator. Next, under suitable assumptions on the drift L,

[1] A Down-and-Out barrier Call (resp. Put) is the option to buy (resp. sell), at maturity T, the underlying with a fixed strike if the underlying value never falls below the barrier before time T.

we prove that the obtained MLMC method for pricing Down-and-Out (resp. Up-and-Out) reaches its optimal time complexity $O(\varepsilon^{-2})$ provided that $\inf_{t\in[0,T]} Y_t$ (resp. $\sup_{t\in[0,T]} Y_t$) has a bounded density in the neighborhood of the barrier (see Theorem 2.3.3 and Remark 2.3.4). In Sections 2.4 and 2.5, we provide two examples of processes satisfying Theorem 2.3.3 conditions, namely the CIR and the CEV models. It turns out that under additional constraints on the parameters of these two models ensuring the existence of finite negative moments up to a certain order, the MLMC method behaves exactly like a classical unbiased Monte Carlo estimator despite the use of approximation schemes. To show that the conditions of our theoretical framework are satisfied for these two models, we develop using fine asymptotic properties of confluent hypergeometric type functions, semi-explicit formulas for the densities of the running minimum and running maximum of both CIR and CEV processes which are of independent interest (see Theorems 2.4.1, 2.4.2, 2.5.2 and 2.5.3). Finally, we proceed to several numerical tests illustrating our results.

2.2 General framework

Let us consider a process $(X_t)_{t\in[0,T]}$ solution to

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x, \tag{2.1}$$

where $(W_t)_{t\geq 0}$ is a standard Brownian motion, $b : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to \mathbb{R}_+^*$ are locally Lipschitz-functions such that $\frac{1}{\sigma}$ is locally integrable. For $\phi(y) = \int_{y_0}^{y} \frac{1}{\sigma(x)}dx$, if $\sigma \in \mathcal{C}^1$ then by the Lamperti transform $Y_t = \phi(X_t)$ solves the stochastic differential equation

$$dY_t = L(X_t)dt + dW_t, \quad Y_0 = \phi(x),$$

with $L(x) = \left(\frac{b}{\sigma} - \frac{\sigma'}{2}\right)(\phi^{-1}(x))$. In this work, we are interested in approximating barrier option prices such as the Down-and-Out (D-O) and the Up-and-Out (U-O) barrier options

$$\pi_{\mathcal{B}_D} = \mathbb{E}\Big[f(X_T)\mathbb{1}_{\{\inf_{t\in[0,T]} X_t > \mathcal{B}_D\}}\Big] \text{ and } \pi_{\mathcal{B}_U} = \mathbb{E}\Big[f(X_T)\mathbb{1}_{\{\sup_{t\in[0,T]} X_t < \mathcal{B}_U\}}\Big].$$

The other types of barrier options such as the Down-and-In and the Up-and-In can be easily deduced from the price of the vanilla option $\mathbb{E}[f(X_T)]$. As the function ϕ is monotonic, by the Lamperti transformation we reduce ourselves to a pricing problem with the process $(Y_t)_{t\in[0,T]}$. More precisely, we get $\pi_{\mathcal{B}_D} = \pi_{\mathcal{D}}$ and $\pi_{\mathcal{B}_U} = \pi_{\mathcal{U}}$ where

$$\pi_{\mathcal{D}} = \mathbb{E}\Big[g(Y_T)\mathbb{1}_{\{\inf_{t\in[0,T]} Y_t > \mathcal{D}\}}\Big], \ \pi_{\mathcal{U}} = \mathbb{E}\Big[g(Y_T)\mathbb{1}_{\{\sup_{t\in[0,T]} Y_t < \mathcal{U}\}}\Big],$$

$g(x) = f \circ \phi^{-1}(x)$, $\mathcal{D} = \phi(\mathcal{B}_D)$ and $\mathcal{U} = \phi(\mathcal{B}_U)$. In the sequel, we consider the general setting given in [5] and let $(Y_t)_{t\geq 0}$ denote the SDE defined on $I = (0, +\infty)$ solution to

$$dY_t = L(Y_t)dt + \gamma dW_t, \quad t \geq 0,\ Y_0 = y \in I, \text{ with } \gamma \in \mathbb{R}^*, \tag{2.2}$$

where the drift coefficient L is supposed to satisfy the following monotonicity assumption:

$$L : I \longrightarrow \mathbb{R} \ \ is \ \ C^2, \ \text{ such that } \ \exists\ \kappa > 0, \ \ \forall y, y' \in I, y \leq y', L(y') - L(y) \leq \kappa(y' - y). \tag{2.3}$$

In addition, for an arbitrary point $d \in I$, we assume that

$$v(x) = \int_d^x \int_d^y \exp\Big(-\frac{2}{\gamma^2}\int_z^y L(\xi)d\xi\Big)dzdy \ \text{ satisfies } \lim_{x\to 0^+} v(x) = +\infty. \tag{H1}$$

On the one hand, by the Feller's test (see e.g. [55]), (2.3) and (H1) ensure that the SDE (2.2) admits a unique strong solution $(Y_t)_{t\geq 0}$ on I that never reaches the boundaries 0 and $+\infty$. On the other hand, under these two conditions the below drift implicit continuous scheme introduced in [5],

$$\begin{aligned}&\widehat{Y}_t^n = \widehat{Y}_{t_i}^n + L(\widehat{Y}_t^n)(t-t_i) + \gamma(W_t - W_{t_i}), \text{ with } t \in (t_i, t_{i+1}], t_i = \frac{iT}{n},\ 0 \leq i \leq n-1, \\ &\widehat{Y}_0^n = y.\end{aligned} \tag{2.4}$$

is well defined and for all $t \in [0,T]$, $Y_t^n \in I$. Besides, if in addition we assume that for $p \geq 1$, we have

$$\mathbb{E}\Big[\Big(\int_0^T |L'(Y_u)L(Y_u) + \frac{\gamma^2}{2}L''(Y_u)|du\Big)^p\Big] < \infty \ \text{ and } \ \mathbb{E}\Big[\Big(\int_0^T (L'(Y_u))^2 du\Big)^{\frac{p}{2}}\Big] < \infty, \tag{H2}$$

then by [5], there exists a positive constant K_p such that

$$\mathbb{E}^{\frac{1}{p}}\left[\sup_{t\in[0,T]} |\widehat{Y}_t^n - Y_t|^p\right] \leq K_p \frac{T}{n}.$$

For our purpose, we rather focus on a slightly different interpolated version of the drift implicit scheme. More precisely, we first introduce the discrete version of the drift implicit scheme given by

$$\left\{\begin{array}{rcl} \overline{Y}_{t_{i+1}}^n &=& \overline{Y}_{t_i}^n + L(\overline{Y}_{t_{i+1}}^n)\frac{T}{n} + \gamma(W_{t_{i+1}} - W_{t_i}), \text{ with } t_i = \frac{iT}{n},\ 0 \leq i \leq n-1, \\ \overline{Y}_0^n &=& y. \end{array}\right. \tag{2.5}$$

and introduce the following interpolated drift implicit scheme

$$\overline{Y}_t^n = \overline{Y}_{t_i}^n + L(\overline{Y}_{t_{i+1}}^n)(t-t_i) + \gamma(W_t - W_{t_i}), \qquad \text{for } t \in [t_i, t_{i+1}[,\ \ 0 \leq i \leq n-1. \tag{2.6}$$

The main advantages of this Brownian interpolation is that it preserves the rate of strong convergence of the original drift implicit scheme (2.4) and allows at the same time the use of the Brownian bridge technique for pricing Barrier options (see Section 2.3.1 below). In what follows, we strengthen our assumption on the drift coefficient L as follows:

$$\begin{aligned}&L : I \to \mathbb{R} \text{ is } \mathcal{C}^2 \text{ such that: } L \text{ is decreasing on } (0,A) \text{ for } A > 0, \\ &\quad \text{and } L' \text{ the first derivative of } L \text{ satisfies } \exists L'_A > 0 \text{ s.t. } \forall y \in (A,\infty),\ |L'(y)| \leq L'_A.\end{aligned} \tag{H3}$$

Theorem 2.2.1. *Assume that conditions* (H2) *and* (H3) *hold true for a given $p > 1$ and with $L'_A < \frac{n}{2T}$. Then, there exists a constant $K_p > 0$ such that*

$$\mathbb{E}^{\frac{1}{p}}\Big[\sup_{t\in[0,T]} |\overline{Y}_t^n - Y_t|^p\Big] \leq K_p \frac{T}{n}.$$

Proof. At first, for $p \geq 1$ and $t \in [0,T]$, we denote $e_t = \overline{Y}_t^n - Y_t$. By (2.2), we have for $0 \leq i \leq n-1$,

$$e_{t_{i+1}} = e_{t_i} + L(\overline{Y}_{t_{i+1}}^n)(t_{i+1} - t_i) - \int_{t_i}^{t_{i+1}} L(Y_s)ds$$

since for all $0 \leq i \leq n-1$, $\overline{Y}^n_{t_i} \in I$. As L is of class $\mathcal{C}^2$ there exists a point $\xi_{t_{i+1}}$ lying between $Y_{t_{i+1}}$ and $\overline{Y}^n_{t_{i+1}}$ such that $L(\overline{Y}^n_{t_{i+1}}) - L(Y_{t_{i+1}}) = \beta_{t_{i+1}}(\overline{Y}^n_{t_{i+1}} - Y_{t_{i+1}})$ with $\beta_{t_{i+1}} = L'(\xi_{t_{i+1}})$. Besides, according to the proof [5, Proposition 3], we know that

$$\mathbb{E}\Big[\sup_{1\leq i\leq n}|e_{t_i}|^p\Big] \leq K\Big(\frac{T}{n}\Big)^p\Big[\mathbb{E}\Big[\Big(\int_0^T |L'(Y_u)L(Y_u) + \frac{\gamma^2}{2}L''(Y_u)|du\Big)^p\Big] + |\gamma|^p\mathbb{E}\Big[\Big(\int_0^T (L'(Y_u))^2du\Big)^{\frac{p}{2}}\Big]\Big], \quad (2.7)$$

where K is a positive constant that depends on T and p. On the one hand, we first use (2.6) to write

$$\begin{aligned} e_{t_{i+1}} &= e_{t_i} + L(\overline{Y}^n_{t_{i+1}})(t_{i+1}-t_i) - \int_{t_i}^{t_{i+1}} L(Y_s)ds \\ &= e_{t_i} + \Big[L(\overline{Y}^n_{t_{i+1}}) - L(Y_{t_{i+1}})\Big](t_{i+1}-t_i) + \int_{t_i}^{t_{i+1}} (L(Y_{t_{i+1}}) - L(Y_s))ds \\ &= e_{t_i} + \beta_{t_{i+1}}e_{t_{i+1}}(t_{i+1}-t_i) + \int_{t_i}^{t_{i+1}} (L(Y_{t_{i+1}}) - L(Y_s))ds. \end{aligned}$$

It follows that

$$(1-\beta_{t_{i+1}}(t_{i+1}-t_i))e_{t_{i+1}} = e_{t_i} + \int_{t_i}^{t_{i+1}} (L(Y_{t_{i+1}}) - L(Y_s))ds. \quad (2.8)$$

On the other hand, we have for all $t \in [t_i, t_{i+1})$

$$\begin{aligned} \overline{Y}^n_t &= \overline{Y}^n_{t_i} + L(\overline{Y}^n_{t_{i+1}})(t-t_i) + \gamma(W_t - W_{t_i}) \\ &= \overline{Y}^n_{t_i} + L(\overline{Y}^n_{t_{i+1}})(t_{i+1}-t_i) + \gamma(W_{t_{i+1}} - W_{t_i}) - L(\overline{Y}^n_{t_{i+1}})(t_{i+1}-t) - \gamma(W_{t_{i+1}} - W_t) \\ &= \overline{Y}^n_{t_{i+1}} - L(\overline{Y}^n_{t_{i+1}})(t_{i+1}-t) - \gamma(W_{t_{i+1}} - W_t). \end{aligned}$$

Then, it follows that for all $t \in [t_i, t_{i+1})$

$$\begin{aligned} \overline{Y}^n_t - Y_t &= \overline{Y}^n_{t_{i+1}} - Y_{t_{i+1}} + Y_{t_{i+1}} - Y_t - L(\overline{Y}^n_{t_{i+1}})(t_{i+1}-t) - \gamma(W_{t_{i+1}} - W_t) \\ e_t &= e_{t_{i+1}} + \int_t^{t_{i+1}} L(Y_s)ds + \gamma(W_{t_{i+1}} - W_t) - L(\overline{Y}^n_{t_{i+1}})(t_{i+1}-t) - \gamma(W_{t_{i+1}} - W_t) \\ &= e_{t_{i+1}} - (L(\overline{Y}^n_{t_{i+1}}) - L(Y_{t_{i+1}}))(t_{i+1}-t) + \int_t^{t_{i+1}} L(Y_s) - L(Y_{t_{i+1}})ds \\ &= e_{t_{i+1}} - \beta_{t_{i+1}}(\overline{Y}^n_{t_{i+1}} - Y_{t_{i+1}})(t_{i+1}-t) + \int_t^{t_{i+1}} L(Y_s) - L(Y_{t_{i+1}})ds. \end{aligned}$$

So, we deduce that for all $t \in [t_i, t_{i+1})$

$$e_t = (1-\beta_{t_{i+1}}(t_{i+1}-t))e_{t_{i+1}} + \int_t^{t_{i+1}} L(Y_s) - L(Y_{t_{i+1}})ds. \quad (2.9)$$

By assumption (H3), on $(0, A)$ L is decreasing, so it is easy to see that $1 < 1-\beta_{t_{i+1}}(t_{i+1}-t) < 1-\beta_{t_{i+1}}(t_{i+1}-t_i)$. On (A, ∞), as L' is bounded and since $n > 2L'_A T$, we have $|1-\beta_{t_{i+1}}(t_{i+1}-t)| \leq \frac{3}{2}$ and $1-\beta_{t_{i+1}}(t_{i+1}-t_i) > \frac{1}{2}$. Then, it follows that $\Big|\frac{1-\beta_{t_{i+1}}(t_{i+1}-t)}{1-\beta_{t_{i+1}}(t_{i+1}-t_i)}\Big| \leq 3$. Now, combining (2.8) and (2.9) we easily get

$$e_t = \frac{1-\beta_{t_{i+1}}(t_{i+1}-t)}{1-\beta_{t_{i+1}}(t_{i+1}-t_i)}\Big(e_{t_i} + \int_{t_i}^{t_{i+1}} (L(Y_{t_{i+1}}) - L(Y_s))ds\Big) + \int_t^{t_{i+1}} L(Y_s) - L(Y_{t_{i+1}})ds. \quad (2.10)$$

Then, by Itô's formula and Fubini theorem we get

$$\begin{aligned}|e_t| &\leq 3\left(|e_{t_i}| + \Big|\int_{t_i}^{t_{i+1}} (L(Y_{t_{i+1}}) - L(Y_s))ds\Big|\right) + \Big|\int_t^{t_{i+1}} L(Y_s) - L(Y_{t_{i+1}})ds\Big| \\ &\leq 3\left(|e_{t_i}| + \frac{T}{n}\int_{t_i}^{t_{i+1}} \Big|L'(Y_u)L(Y_u) + \frac{\gamma^2}{2}L''(Y_u)\Big|du + |\gamma|\Big|\int_{t_i}^{t_{i+1}} (u-t_i)L'(Y_u)dW_u\Big|\right) \\ &+ \frac{T}{n}\int_t^{t_{i+1}} \Big|L'(Y_u)L(Y_u) + \frac{\gamma^2}{2}L''(Y_u)\Big|du + |\gamma|\Big|\int_t^{t_{i+1}} (u-t)L'(Y_u)dW_u\Big|.\end{aligned}$$

Therefore, there exists a positive constant C_p such that

$$\begin{aligned}|e_t|^p \leq C_p\Bigg[&|e_{t_i}|^p + 2\Big(\frac{T}{n}\Big)^p\Big(\int_{t_i}^{t_{i+1}} \Big|L'(Y_u)L(Y_u) + \frac{\gamma^2}{2}L''(Y_u)\Big|du\Big)^p + |\gamma|^p\Big|\int_{t_i}^{t_{i+1}} (u-t_i)L'(Y_u)dW_u\Big|^p \\ &+ |\gamma|^p\Big|\int_t^{t_{i+1}} uL'(Y_u)dW_u\Big|^p + |\gamma|^p T^p\Big|\int_t^{t_{i+1}} L'(Y_u)dW_u\Big|^p\Bigg]\end{aligned}$$

and thus,

$$\begin{aligned}\sup_{t\in[0,T]}|e_t|^p &\leq C_p\Bigg[\sup_{0\leq i\leq n}|e_{t_i}|^p + 2\Big(\frac{T}{n}\Big)^p\Big(\int_0^T \Big|L'(Y_u)L(Y_u) + \frac{\gamma^2}{2}L''(Y_u)\Big|du\Big)^p \\ &+ |\gamma|^p \sup_{0\leq s\leq t\leq T}\Big|\int_s^t (u-t_{\eta(u)})L'(Y_u)dW_u\Big|^p + |\gamma|^p \sup_{0\leq s\leq t\leq T}\Big|\int_s^t uL'(Y_u)dW_u\Big|^p \\ &+ |\gamma|^p T^p \sup_{0\leq s\leq t\leq T}\Big|\int_s^t L'(Y_u)dW_u\Big|^p\Bigg] \\ &\leq C_p\Bigg[\sup_{0\leq i\leq n}|e_{t_i}|^p + 2\Big(\frac{T}{n}\Big)^p\Big(\int_0^T \Big|L'(Y_u)L(Y_u) + \frac{\gamma^2}{2}L''(Y_u)\Big|du\Big)^p \\ &+ 2^{p-1}|\gamma|^p \sup_{0\leq t\leq T}\Big|\int_0^t (u-t_{\eta(u)})L'(Y_u)dW_u\Big|^p + 2^{p-1}|\gamma|^p \sup_{0\leq t\leq T}\Big|\int_0^t uL'(Y_u)dW_u\Big|^p \\ &+ 2^{p-1}|\gamma|^p T^p \sup_{0\leq t\leq T}\Big|\int_0^t L'(Y_u)dW_u\Big|^p\Bigg].\end{aligned}$$

The result follows using (2.7) and the Burkholder-Davis-Gundy inequality. □

Corollary 2.2.2. *Assume that conditions of Theorem 2.2.1 hold true for a given $p > 1$ and $0 < L'_A < \frac{n}{2T}$. If in addition the drift coefficient L satisfies the following one-sided linear growth assumption:*

$$\exists\, \alpha > 0 \text{ such that } \forall\, y \in I,\ \ yL(y) \leq \alpha(1 + |y|^2) \tag{H4}$$

then $\mathbb{E}[\sup_{0\leq t\leq T}|\overline{Y}_t^n|^p] < \infty$.

Proof. Under assumption (H4), [51, Lemma 3.2] ensures that for all $q > 0$ $\mathbb{E}[\sup_{0\leq t\leq T}|Y_t|^q] < \infty$. Thus, by Theorem 2.2.1 we get

$$\mathbb{E}[\sup_{0\leq t\leq T}|\overline{Y}_t^n|^p] \leq 2^{p-1}\left(\mathbb{E}[\sup_{0\leq t\leq T}|Y_t|^p + \mathbb{E}[\sup_{0\leq t\leq T}|\overline{Y}_t^n - \overline{Y}_t|^p]\right) < \infty.$$

□

2.3 The Multilevel Monte Carlo method for pricing Barrier options with the interpolated drift implicit scheme.

We are interested to approximate the following quantities of interest of the form

$$\pi_{\mathcal{D}} = \mathbb{E}\Big[g(Y_T)\mathbb{1}_{\{\tau_{\mathcal{D}}>T\}}\Big] \text{ and } \pi_{\mathcal{U}} = \mathbb{E}\Big[g(Y_T)\mathbb{1}_{\{\tau_{\mathcal{U}}>T\}}\Big],$$

where τ denotes the first passage time given by:

- $\tau_{\mathcal{D}} = \inf\{t \in [0,T], Y_t \leq \mathcal{D}\}$ with $y > \mathcal{D} > 0$, for a Down-Out (D-O) option,

  and

- $\tau_{\mathcal{U}} = \inf\{t \in [0,T], Y_t \geq \mathcal{U}\}$ with $0 < y < \mathcal{U}$, for an Up-Out (U-O) option.

2.3.1 Brownian bridge and drift implicit scheme for pricing Barrier options

For a given time grid $t_i = \frac{iT}{n}$, $i \in \{1,\dots,n\}$, we consider the Brownian interpolation of the drift implicit scheme $(\overline{Y}^n_t)_{t\in[0,T]}$ defined in (2.6). Then, the above option prices can be approximated respectively by

$$\overline{\pi}_{\mathcal{D}} := \mathbb{E}\Big[g(\overline{Y}^n_T)\prod_{i=0}^{n-1}\mathbf{1}_{\{\inf_{t\in[t_i,t_{i+1}]}\overline{Y}^n_t>\mathcal{D}\}}\Big] \text{ and } \overline{\pi}_{\mathcal{U}} := \mathbb{E}\Big[g(\overline{Y}^n_T)\prod_{i=0}^{n-1}\mathbf{1}_{\{\sup_{t\in[t_i,t_{i+1}]}\overline{Y}^n_t<\mathcal{U}\}}\Big].$$

To get more accurate approximations, we use the Brownian bridge technique to substitute the barrier-crossing indicators by the probabilities that the approximation scheme $(\overline{Y}^n_t)_{t\in[0,T]}$ do not cross the barrier in each time interval $[t_i,t_{i+1}]$, $i \in \{1,\dots,n\}$. In what follows, for $x \in \mathbb{R}$, $(x)_+$ stands for $\sup(x,0)$.

Proposition 2.3.1. *Under the above notation, for $h = \frac{T}{n}$, we have*

$$\overline{\pi}_{\mathcal{D}} = \mathbb{E}\Big[g(\overline{Y}^n_T)\prod_{i=0}^{n-1}(1-\overline{q}_i)\Big], \text{ where } \overline{q}_i := \exp\Big(\frac{-2(\overline{Y}^n_{t_i}-\mathcal{D})_+(\overline{Y}^n_{t_{i+1}}-\mathcal{D})_+}{\gamma^2 h}\Big)$$

and

$$\overline{\pi}_U = \mathbb{E}\Big[g(\overline{Y}^n_T)\prod_{i=0}^{n-1}(1-\overline{p}_i)\Big], \text{ where } \overline{p}_i = \exp\Big(\frac{-2(\mathcal{U}-\overline{Y}^n_{t_i})_+(\mathcal{U}-\overline{Y}^n_{t_{i+1}})_+}{\gamma^2 h}\Big).$$

Proof. For the U-O barrier option, we first notice that conditionally on $(\overline{Y}^n_0, \overline{Y}^n_{t_1},\dots,\overline{Y}^n_T)$, the barrier-crossing indicators $(\mathbf{1}_{\{\sup_{t\in[t_i,t_{i+1}]}\overline{Y}^n_t<\mathcal{U}\}}, i \in \{1,\dots,n\})$ are independent, we write

$$\begin{aligned}
\overline{\pi}_{\mathcal{U}} &= \mathbb{E}\Big[g(\overline{Y}^n_T)\mathbb{E}\big[\prod_{i=0}^{n-1}\mathbb{1}_{\{\sup_{t\in[t_i,t_{i+1}]}\overline{Y}^n_t<\mathcal{U}\}}|\overline{Y}^n_0,\overline{Y}^n_{t_1},\dots,\overline{Y}^n_T\big]\Big]\\
&= \mathbb{E}\Big[g(\overline{Y}^n_T)\prod_{i=0}^{n-1}\mathbb{E}\big[\mathbb{1}_{\{\sup_{t\in[t_i,t_{i+1}]}\overline{Y}^n_t<\mathcal{U}\}}|\overline{Y}^n_{t_i},\overline{Y}^n_{t_{i+1}}\big]\Big]\\
&= \mathbb{E}\Big[g(\overline{Y}^n_T)\prod_{i=0}^{n-1}(1-\varphi(\overline{Y}^n_{t_i},\overline{Y}^n_{t_{i+1}}))\Big],
\end{aligned}$$

where, for $y_i, y_{i+1} \in I$, $\varphi(y_i, y_{i+1}) = \mathbb{P}(\sup_{t\in[t_i,t_{i+1}]} \overline{Y}^n_t \geq \mathcal{U} | \overline{Y}^n_{t_i} = y_i, \overline{Y}^n_{t_{i+1}} = y_{i+1})$. Without loss of generality, we may assume $\gamma > 0$, the same arguments below work for $\gamma < 0$ using that $(W_t)_{t\geq 0}$ and $(-W_t)_{t\geq 0}$ have the same law. By(2.6), we write

$$\sup_{t\in[t_i,t_{i+1}]} \overline{Y}^n_t = \overline{Y}^n_{t_i} + \gamma \sup_{t\in[t_i,t_{i+1}]} \Big[W_t - W_{t_i} + \frac{1}{\gamma} L(\overline{Y}^n_{t_{i+1}})(t - t_i) \Big], \text{ with}$$
$$W_{t_{i+1}} - W_{t_i} + \frac{1}{\gamma} L(\overline{Y}^n_{t_{i+1}})(t_{i+1} - t_i) = \frac{1}{\gamma}(\overline{Y}^n_{t_{i+1}} - \overline{Y}^n_{t_i}).$$

By the stationarity property of the brownian increments and using a change of probability measure, we easily get that the law of $\sup_{t\in[t_i,t_{i+1}]} \Big[W_t - W_{t_i} + \frac{1}{\gamma} L(y_{i+1})(t - t_i) \Big]$ given $W_{t_{i+1}} - W_{t_i} + \frac{1}{\gamma} L(y_{i+1})(t_{i+1} - t_i) = \frac{1}{\gamma}(y_{i+1} - y_i)$ is equal to the law of $\sup_{t\in[0,t_1]} W_t$ given $W_{t_1} = \frac{1}{\gamma}(y_{i+1} - y_i)$ which is given by $\mathbb{P}\Big(\sup_{t\in[0,t_1]} W_{t_1} \geq y \Big| W_{t_1} = x \Big) = \mathrm{e}^{\frac{-2(y)_+(y-x)_+}{h}}$ (see e.g. [55, p. 265]). Thus, we get

$$\varphi(y_i, y_{i+1}) = \mathbb{P}\Big(\sup_{t\in[0,t_1]} W_t \geq \frac{1}{\gamma}(\mathcal{U} - y_i) | W_{t_1} = \frac{1}{\gamma}(y_{i+1} - y_i) \Big)$$
$$= \exp\Big(\frac{-2(\mathcal{U} - y_i)_+(\mathcal{U} - y_{i+1})_+}{\gamma^2 h} \Big).$$

The same arguments applied to $(-W_t)_{t\geq 0}$ work for the Down-Out barrier option. □

2.3.2 The interpolated drift implicit Euler scheme MLMC method analysis.

We consider the drift implicit scheme $(\overline{Y}^{2^\ell}_{t_i})_{0\leq i\leq 2^\ell}$ given in (2.5) that approximates $(Y_t)_{0\leq t\leq T}$ solution to (2.2) using a time step $h_\ell = 2^{-\ell} T$ for $\ell \in \{0, ..., L\}$, with $L = \log n / \log 2$, where n denotes the finest time step number. Let $(\overline{Y}^{2^\ell}_t)_{0\leq t\leq T}$ denote the Brownian interpolation of the drift implicit scheme defined in (2.6) with time step h_ℓ. As the same arguments work for both Down-Out and Up-Out barrier options, we give details only for the latter one. To do so, we introduce

$$\overline{P}_\ell := g(\overline{Y}^{2^\ell}_T) \prod_{i=0}^{2^\ell - 1} \mathbf{1}_{\{\sup_{t\in[t^\ell_i, t^\ell_{i+1}]} \overline{Y}^{2^\ell}_t < \mathcal{U}\}}, \quad \text{where } t^\ell_i = \frac{iT}{2^\ell} \quad \text{for } \ell \in \{0, ..., L\}, \tag{2.11}$$

and write

$$\overline{\pi}_{\mathcal{U}} = \mathbb{E}[\overline{P}_L] = \mathbb{E}[\overline{P}_0] + \sum_{\ell=1}^{L} \mathbb{E}[\overline{P}_\ell - \overline{P}_{\ell-1}], \tag{2.12}$$

where $\overline{\pi}_{\mathcal{U}}$ is introduced in subsection 2.3.1. On the one hand, applying Proposition 3.4.1 yields

$$\mathbb{E}[\overline{P}_\ell] = \mathbb{E}[\overline{P}^f_\ell], \text{ where } \overline{P}^f_\ell := g(\overline{Y}^{2^\ell}_T) \prod_{i=0}^{2^\ell - 1} (1 - \overline{p}^{2^\ell}_i) \text{ with} \tag{2.13}$$
$$\overline{p}^{2^\ell}_i = \exp\Big(\frac{-2(\mathcal{U} - \overline{Y}^{2^\ell}_{t^\ell_i})_+(\mathcal{U} - \overline{Y}^{2^\ell}_{t^\ell_{i+1}})_+}{\gamma^2 h_\ell} \Big).$$

On the other hand, we write

$$\mathbb{E}[\overline{P}_{\ell-1}] = \mathbb{E}[g(\overline{Y}_T^{2^{\ell-1}}) \prod_{i=0}^{2^{\ell-1}-1} \mathbb{E}[\mathbf{1}_{\{\sup_{t\in[t_i^{\ell-1},t_{i+1}^{\ell-1}]} \overline{Y}_t^{2^{\ell-1}} < \mathcal{U}\}} | \overline{Y}_{t_i^{\ell-1}}^{2^{\ell-1}}, \overline{Y}_{t_{2i+1}^{\ell}}^{2^{\ell-1}}, \overline{Y}_{t_{i+1}^{\ell-1}}^{2^{\ell-1}}]]$$

$$= \mathbb{E}[g(\overline{Y}_T^{2^{\ell-1}}) \prod_{i=0}^{2^{\ell-1}-1} \mathbb{E}[\mathbf{1}_{\{\sup_{t\in[t_i^{\ell-1},t_{2i+1}^{\ell}]} \overline{Y}_t^{2^{\ell-1}} < \mathcal{U}\}} \mathbf{1}_{\{\sup_{t\in[t_{2i+1}^{\ell},t_{i+1}^{\ell-1}]} \overline{Y}_t^{2^{\ell-1}} < \mathcal{U}\}} | \overline{Y}_{t_i^{\ell-1}}^{2^{\ell-1}}, \overline{Y}_{t_{2i+1}^{\ell}}^{2^{\ell-1}}, \overline{Y}_{t_{i+1}^{\ell-1}}^{2^{\ell-1}}]],$$

where the coarse scheme $\overline{Y}_{t_{2i+1}^{\ell}}^{2^{\ell-1}}$ is computed using our Brownian interpolation scheme (2.6) that is

$$\overline{Y}_{t_{2i+1}^{\ell}}^{2^{\ell-1}} = \overline{Y}_{t_i^{\ell-1}}^{2^{\ell-1}} + L(\overline{Y}_{t_{i+1}^{\ell-1}}^{2^{\ell-1}})(t_{2i+1}^{\ell} - t_i^{\ell-1}) + \gamma(W_{t_{2i+1}^{\ell}} - W_{t_i^{\ell-1}}).$$

Thus, we rewrite $\sup_{t\in[t_i^{\ell-1},t_{2i+1}^{\ell}]} \overline{Y}_t^{2^{\ell-1}}$ and $\sup_{t\in[t_{2i+1}^{\ell},t_{i+1}^{\ell-1}]} \overline{Y}_t^{2^{\ell-1}}$ as follows

$$\sup_{t\in[t_i^{\ell-1},t_{2i+1}^{\ell}]} \overline{Y}_t^{2^{\ell-1}} = \overline{Y}_{t_i^{\ell-1}}^{2^{\ell-1}} + \gamma \sup_{t\in[t_i^{\ell-1},t_{2i+1}^{\ell}]} \left(W_t - W_{t_i^{\ell-1}} + \frac{1}{\gamma} L(\overline{Y}_{t_{i+1}^{\ell-1}}^{2^{\ell-1}})(t - t_i^{\ell-1})\right), \text{ with}$$

$$W_{t_{2i+1}^{\ell}} - W_{t_i^{\ell-1}} + \frac{1}{\gamma} L(\overline{Y}_{t_{i+1}^{\ell}}^{2^{\ell-1}})(t_{2i+1}^{\ell} - t_i^{\ell-1}) = \frac{1}{\gamma}\left(\overline{Y}_{t_{2i+1}^{\ell}}^{2^{\ell-1}} - \overline{Y}_{t_i^{\ell-1}}^{2^{\ell-1}}\right)$$

and

$$\sup_{t\in[t_{2i+1}^{\ell},t_{i+1}^{\ell-1}]} \overline{Y}_t^{2^{\ell-1}} = \overline{Y}_{t_{2i+1}^{\ell}}^{2^{\ell-1}} + \gamma \sup_{t\in[t_i^{\ell-1},t_{2i+1}^{\ell}]} \left(W_t - W_{t_{2i+1}^{\ell}} + \frac{1}{\gamma} L(\overline{Y}_{t_{i+1}^{\ell-1}}^{2^{\ell-1}})(t - t_{2i+1}^{\ell})\right), \text{ with}$$

$$W_{t_{i+1}^{\ell-1}} - W_{t_{2i+1}^{\ell}} + \frac{1}{\gamma} L(\overline{Y}_{t_{i+1}^{\ell-1}}^{2^{\ell-1}})(t_{i+1}^{\ell-1} - t_{2i+1}^{\ell}) = \frac{1}{\gamma}\left(\overline{Y}_{t_{i+1}^{\ell-1}}^{2^{\ell-1}} - \overline{Y}_{t_{2i+1}^{\ell}}^{2^{\ell-1}}\right).$$

Then, using the independence of the Brownian increments and the same arguments as in the proof of Proposition 3.4.1, we get

$$\mathbb{E}[\overline{P}_{\ell-1}] = \mathbb{E}[\overline{P}_{\ell-1}^c], \text{ where } \overline{P}_{\ell-1}^c := g(\overline{Y}_T^{2^{\ell-1}}) \prod_{i=0}^{2^{\ell-1}-1} (1 - \overline{p}_{i,1}^{2^{\ell-1}})(1 - \overline{p}_{i,2}^{2^{\ell-1}}) \text{ with} \tag{2.14}$$

$$\overline{p}_{i,1}^{2^{\ell-1}} = \exp\Big(\frac{-2(\mathcal{U} - \overline{Y}_{t_i^{\ell-1}}^{2^{\ell-1}})_+(\mathcal{U} - \overline{Y}_{t_{2i+1}^{\ell}}^{2^{\ell-1}})_+}{\gamma^2 h_\ell}\Big),$$

$$\overline{p}_{i,2}^{2^{\ell-1}} = \exp\Big(\frac{-2(\mathcal{U} - \overline{Y}_{t_{2i+1}^{\ell}}^{2^{\ell-1}})_+(\mathcal{U} - \overline{Y}_{t_{i+1}^{\ell-1}}^{2^{\ell-1}})_+}{\gamma^2 h_\ell}\Big),$$

which can be rewritten as

$$\overline{P}_{\ell-1}^c := g(\overline{Y}_T^{2^{\ell-1}}) \prod_{i=0}^{2^{\ell}-1} (1 - \overline{p}_i^{2^{\ell-1}}) \text{ with } \overline{p}_i^{2^{\ell-1}} = \exp\Big(\frac{-2(\mathcal{U} - \overline{Y}_{t_i^{\ell}}^{2^{\ell-1}})_+(\mathcal{U} - \overline{Y}_{t_{i+1}^{\ell}}^{2^{\ell-1}})_+}{\gamma^2 h_\ell}\Big), \tag{2.15}$$

where the coarse approximation scheme $\overline{Y}_{t_i^{\ell}}^{2^{\ell-1}}$ for odd index $i \in \{0, ..., 2^\ell - 1\}$ is computed using the Brownian interpolation scheme (2.6). Thus, the improved MLMC method approximates $\bar{\pi}_U$ by

$$\bar{P}_U := \frac{1}{N_0} \sum_{k=1}^{N_0} \overline{P}_{0,k}^f + \sum_{\ell=1}^{L} \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \left(\overline{P}_{\ell,k}^f - \overline{P}_{\ell-1,k}^c\right), \tag{2.16}$$

where the condition $\mathbb{E}[\overline{P}^f_{\ell-1}] = \mathbb{E}[\overline{P}^c_{\ell-1}]$ is satisfied according to (3.9) and (3.10). Here, $(\overline{P}^f_{\ell,k})_{1\le k\le N_\ell}$ and $(\overline{P}^c_{\ell-1,k})_{1\le k\le N_\ell}$ are respectively independent copies of $\overline{P}^f_\ell$ and $\overline{P}^c_{\ell-1}$ given by (3.9) and (3.11). Similarly, the improved MLMC method approximates $\bar{\pi}_{\mathcal{D}}$ by

$$\bar{Q}_{\mathcal{D}} := \frac{1}{N_0}\sum_{k=1}^{N_0}\overline{Q}^f_{0,k} + \sum_{\ell=1}^{L}\frac{1}{N_\ell}\sum_{k=1}^{N_\ell}\Big(\overline{Q}^f_{\ell,k} - \overline{Q}^c_{\ell-1,k}\Big), \tag{2.17}$$

where $(\overline{Q}^f_{\ell,k})_{1\le k\le N_\ell}$ and $(\overline{Q}^c_{\ell-1,k})_{1\le k\le N_\ell}$ are respectively independent copies of $\overline{Q}^f_\ell$ and $\overline{Q}^c_{\ell-1}$ given by

$$\overline{Q}^f_\ell := g(\overline{Y}^{2^\ell}_T)\prod_{i=0}^{2^\ell-1}(1-\overline{q}^{2^\ell}_i) \text{ with } \overline{q}^{2^\ell}_i = \exp\Big(\frac{-2(\overline{Y}^{2^\ell}_{t^\ell_i}-\mathcal{D})_+(\overline{Y}^{2^\ell}_{t^\ell_{i+1}}-\mathcal{D})_+}{\gamma^2 h_\ell}\Big)$$

$$\overline{Q}^c_{\ell-1} := g(\overline{Y}^{2^{\ell-1}}_T)\prod_{i=0}^{2^\ell-1}(1-\overline{q}^{2^{\ell-1}}_i) \text{ with } \overline{q}^{2^{\ell-1}}_i = \exp\Big(\frac{-2(\overline{Y}^{2^{\ell-1}}_{t^\ell_i}-\mathcal{D})_+(\overline{Y}^{2^{\ell-1}}_{t^\ell_{i+1}}-\mathcal{D})_+}{\gamma^2 h_\ell}\Big).$$

In what follows, we need to strengthen our assumption (H2) as below.

$$\text{For } p\ge 1, \text{ assumption (H2) is valid and } \sup_{t\in[0,T]}\mathbb{E}\Big[|L(Y_t)|^p\Big] < \infty. \tag{$\tilde{\text{H2}}$}$$

Lemma 2.3.2. *Assume that conditions* ($\tilde{\text{H2}}$), (H3) *and* (H4) *are satisfied for a given* $p>1$ *and* $0<L'_A<\frac{1}{2h_\ell}$, *with* $h_\ell = 2^{-\ell}T$ *sufficiently small. Let* $\eta\in(0,1)$, *the following extreme path events satisfy*

$$\mathbb{P}\left(\max\Big(\sup_{0\le i\le 2^\ell}(|Y_{t^\ell_i}|,|\overline{Y}^{2^\ell}_{t^\ell_i}|,|\overline{Y}^{2^{\ell-1}}_{t^\ell_i}|)\Big) > h_\ell^{-\eta}\right) = o(h_\ell^q) \tag{2.18}$$

$$\mathbb{P}\left(\max\Big(\sup_{0\le i\le 2^\ell}(|Y_{t^\ell_i}-\overline{Y}^{2^\ell}_{t^\ell_i}|,|Y_{t^\ell_i}-\overline{Y}^{2^{\ell-1}}_{t^\ell_i}|,|\overline{Y}^{2^\ell}_{t^\ell_i}-\overline{Y}^{2^{\ell-1}}_{t^\ell_i}|)\Big) > h_\ell^{1-\eta}\right) = o(h_\ell^q) \tag{2.19}$$

$$\sup_{0\le i\le 2^\ell}\mathbb{P}\left(\int_{t^\ell_i}^{t^\ell_{i+1}}|L(Y_s)|ds > h_\ell^{1-\eta}\right) = o(h_\ell^q) \tag{2.20}$$

$$\mathbb{P}\left(\sup_{t\in[0,T]}|\overline{Y}^{2^\ell}_t - Y_t| > h_\ell^{1-\eta}\right) = o(h_\ell^q) \tag{2.21}$$

for all $0<q<p\eta$, *and*

$$\sup_{0\le i\le 2^\ell}\mathbb{P}\left(\sup_{t\in[t^\ell_i,t^\ell_{i+1}]}|W_t - W_{t^\ell_i}| > h_\ell^{\frac{1}{2}-\eta}\right) = o(h_\ell^q), \quad \textit{for all } q>0. \tag{2.22}$$

Proof. For the first extreme path property, we have

$$\mathbb{P}\big(\max(\sup_{0\le i\le 2^\ell}(|Y_{t^\ell_i}|,|\overline{Y}^{2^\ell}_{t^\ell_i}|,|\overline{Y}^{2^{\ell-1}}_{t^\ell_i}|)) > h_\ell^{-\eta}\big)$$
$$\le \mathbb{P}\big(\sup_{0\le i\le 2^\ell}|Y_{t^\ell_i}| > h_\ell^{-\eta}\big) + \mathbb{P}\big(\sup_{0\le i\le 2^\ell}|\overline{Y}^{2^\ell}_{t^\ell_i}| > h_\ell^{-\eta}\big) + \mathbb{P}\big(\sup_{0\le i\le 2^\ell}|\overline{Y}^{2^{\ell-1}}_{t^\ell_i}| > h_\ell^{-\eta}\big).$$

Then, by Markov's inequality we get for $m\ge 1$

$$\mathbb{P}\big(\max(\sup_{0\le i\le 2^\ell}(|Y_{t^\ell_i}|,|\overline{Y}^{2^\ell}_{t^\ell_i}|,|\overline{Y}^{2^{\ell-1}}_{t^\ell_i}|)) > h_\ell^{-\eta}\big) \le$$
$$h_\ell^{m\eta}\Big(\mathbb{E}[\sup_{0\le t\le T}|Y_t|^m] + \mathbb{E}[\sup_{0\le t\le T}|\overline{Y}^{2^\ell}_t|^m] + \mathbb{E}[\sup_{0\le t\le T}|\overline{Y}^{2^{\ell-1}}_t|^m]\Big).$$

The result follows by Corollary 2.2.2 for h_ℓ sufficiently small with choosing m such that $0 < \frac{q}{\eta} < m \leq p$. For the second extreme path property, we proceed in the same way to get for $m \geq 1$

$$\mathbb{P}\left(\max\Big(\sup_{0\leq i\leq 2^\ell}(|Y_{t_i^\ell}-\overline{Y}^{2^\ell}_{t_i^\ell}|,|Y_{t_i^\ell}-\overline{Y}^{2^{\ell-1}}_{t_i^\ell}|,|\overline{Y}^{2^\ell}_{t_i^\ell}-\overline{Y}^{2^{\ell-1}}_{t_i^\ell}|)\Big)>h_\ell^{1-\eta}\right)$$
$$\leq h_\ell^{-m(1-\eta)}\Big(\mathbb{E}[\sup_{0\leq t\leq T}|Y_t-\overline{Y}^{2^\ell}_t|^m]+\mathbb{E}[\sup_{0\leq t\leq T}|Y_t-\overline{Y}^{2^{\ell-1}}_t|^m]+\mathbb{E}[\sup_{0\leq t\leq T}|\overline{Y}^{2^\ell}_t-\overline{Y}^{2^{\ell-1}}_t|^m]\Big).$$

Thus, we deduce the result using Theorem 2.2.1 with choosing m such that $0 < \frac{q}{\eta} < m \leq p$, for h_ℓ sufficiently small. For the third extreme path, we proceed in the same way to get that for all $0 \leq i \leq 2^\ell$, using Jensen's inequality

$$\begin{aligned}\mathbb{P}\left(\int_{t_i^\ell}^{t_{i+1}^\ell}|L(Y_s)|ds>h_\ell^{1-\eta}\right)&\leq h_\ell^{m(\eta-1)}\mathbb{E}\left(\left|\int_{t_i^\ell}^{t_{i+1}^\ell}|L(Y_s)|ds\right|^m\right)\\&\leq h_\ell^{m\eta-1}\mathbb{E}\left(\int_{t_i^\ell}^{t_{i+1}^\ell}|L(Y_s)|^m ds\right)\\&\leq h_\ell^{m\eta}\sup_{t\in[0,T]}\mathbb{E}\left(|L(Y_s)|^m\right).\end{aligned}$$

Then we conclude using ($\tilde{\text{H}}2$) by choosing m such that $0 < \frac{q}{\eta} < m \leq p$, for h_ℓ sufficiently small. The fourth extreme path property follows in the same way as the second one. Finally, the last property follows using that $\mathbb{E}\left[\sup_{t\in[0,T]}|W_t|^m\right]$ is finite for any positive power m. □

Now, we are able to state our main theorem for the MLMC method to price barrier options when the underlying asset has possibly non-Lipschitz coefficients.

Theorem 2.3.3. *Let g denote a payoff function satisfying : $\exists C > 0$ s.t. $\forall x, y > 0$,*

$$|g(x)-g(y)| \leq C|x-y|(1+|x|^\nu+|y|^\nu) \text{ and } |g(x)| \leq C(1+|x|^{\nu+1}), \text{ with } \nu \in \mathbb{R}_+. \tag{2.23}$$

Moreover, assume that conditions ($\tilde{\text{H}}2$), (H3) *and* (H4) *are satisfied for $p > \frac{(1+\delta)(1+\gamma)[7(1+\varepsilon)+2\nu]}{\frac{1}{2}-\delta}$, with $\varepsilon, \gamma > 0$, $\delta \in (0, 1/2)$ and $0 < L'_A < \frac{1}{2h_\ell}$, with $h_\ell = 2^{-\ell}T$ sufficiently small. If in addition $\inf_{t\in[0,T]} Y_t$ (resp. $\sup_{t\in[0,T]} Y_t$) has a bounded density in the neighborhood of the barrier $\mathcal{D}$ (resp. $\mathcal{U}$), then the multilevel estimator $\bar{Q}_\mathcal{D}$ given by* (3.13) *(resp. $\bar{P}_\mathcal{U}$ given by* (3.12) *) for the D-O (resp. U-O) barrier option satisfies* $\text{Var}(\overline{\text{Q}}^{\text{f}}_\ell - \overline{\text{Q}}^{\text{c}}_\ell) = \text{O}(\text{h}_\ell^{1+\delta})$ *(resp.* $\text{Var}(\overline{\text{P}}^{\text{f}}_\ell - \overline{\text{P}}^{\text{c}}_\ell) = \text{O}(\text{h}_\ell^{1+\delta})$*).*

Remark 2.3.4. • *Combining the complexity theorem in [36, Theorem 3.1] with the above result, we deduce that for any $\delta \in (0, \frac{1}{2})$ the MLMC estimators $\bar{Q}_\mathcal{D}$ and $\bar{P}_\mathcal{U}$ reach the optimal time complexity $O(\varepsilon^{-2})$, for a given precision $\varepsilon > 0$, and behave like an unbiased Monte Carlo estimator.*

• *Taking δ close to $\frac{1}{2}$ achieves a smaller variance of the difference between the finer and coarse approximations which is of order $O(h_\ell^\beta)$ with β close to $\frac{3}{2}$ similar to the case of diffusion with Lipschitz coefficients studied in [37, Theorem 3.15], but clearly leads to very restrictive conditions on the finiteness of the moments of $(Y_t)_{t\in[0,T]}$ and $(\bar{Y}^n_t)_{t\in[0,T]}$.*

Proof. We only give a proof for the D-O barrier option since the proof for the U-O barrier option is quite similar. At first, following the extreme path approach given in [39, 37], we write

$$\begin{aligned}\text{Var}[\overline{\text{Q}}^{\text{f}}_\ell - \overline{\text{Q}}^{\text{c}}_\ell] &\leq \mathbb{E}[(\overline{Q}^f_\ell - \overline{Q}^c_\ell)^2]\\&= \mathbb{E}[(\overline{Q}^f_\ell - \overline{Q}^c_\ell)^2\mathbb{1}_{A_1}] + \mathbb{E}[(\overline{Q}^f_\ell - \overline{Q}^c_\ell)^2\mathbb{1}_{A_2}] + \mathbb{E}[(\overline{Q}^f_\ell - \overline{Q}^c_\ell)^2\mathbb{1}_{A_3}]\end{aligned}$$

where we split the paths into the following three events.

First event A_1 We consider any of the extreme path events given in Lemma 2.3.2 that satisfy (2.18)-(2.22), with some $\eta > 0$ to be fixed later on. For $\gamma > 0$, we use Hölder's inequality to get

$$\begin{aligned}\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_1}] &\leq \mathbb{E}^{\frac{\gamma}{1+\gamma}}\Big[|\overline{Q}_\ell^f - \overline{Q}_\ell^c|^{\frac{2(1+\gamma)}{\gamma}}\Big]\mathbb{E}^{\frac{1}{1+\gamma}}[\mathbb{1}_{A_1}] \\ &\leq 2^{\frac{2+\gamma}{1+\gamma}}\left(\mathbb{E}^{\frac{\gamma}{1+\gamma}}[|\overline{Q}_\ell^f|^{\frac{2(1+\gamma)}{\gamma}}] + \mathbb{E}^{\frac{\gamma}{1+\gamma}}[|\overline{Q}_\ell^c|^{\frac{2(1+\gamma)}{\gamma}}]\right)\Big(\mathbb{P}[A_1]\Big)^{\frac{1}{1+\gamma}}.\end{aligned}$$

As the payoff function g satisfies assumption (2.23), we deduce using Corollary 2.2.2, that for h_ℓ sufficiently small $\mathbb{E}[(\overline{Q}_\ell^f)^{\frac{2(1+\gamma)}{\gamma}}]$ and $\mathbb{E}[(\overline{Q}_\ell^c)^{\frac{2(1+\gamma)}{\gamma}}]$ are finite. By Lemma 2.3.2, we get that

$$\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_1}] = o(h_\ell^{\frac{q}{1+\gamma}}) \text{ for all } q \text{ such that } 0 < \frac{q}{\eta} \leq p. \tag{2.24}$$

Second event A_2 This event corresponds to the non-extreme paths satisfying

$$|\inf_{t\in[0,T]} Y_t - \mathcal{D}| > h_\ell{}^{\frac{1}{2}-\eta(1+\varepsilon)} \text{ for } \eta \in (0, 1/2(1+\varepsilon)) \text{ with } \varepsilon > 0.$$

Let us assume that $\inf_{t\in[0,T]} Y_t = Y_\tau$ with $\tau \in [t_i^\ell, t_{i+1}^\ell]$ for a given $i \in \{0, \dots, 2^\ell\}$.

- First case: for $Y_\tau < \mathcal{D} - h_\ell{}^{\frac{1}{2}-\eta(1+\varepsilon)}$, we write

$$\begin{aligned}|\overline{Y}_{t_i^\ell}^{2^\ell} - Y_\tau| &\leq |\overline{Y}_{t_i^\ell}^{2^\ell} - Y_{t_i^\ell}| + |Y_{t_i^\ell} - Y_\tau| \\ &\leq |\overline{Y}_{t_i^\ell}^{2^\ell} - Y_{t_i^\ell}| + \int_{t_i^\ell}^{t_{i+1}^\ell} |L(Y_s)|ds + \gamma \sup_{t\in[t_i^\ell, t_{i+1}^\ell]} |W_t - W_{t_i^\ell}|.\end{aligned}$$

Then, as we work on the non-extreme paths events (see Lemma 2.3.2 for the extreme paths events), for h_ℓ sufficiently small we have that $\left|\overline{Y}_{t_i^\ell}^{2^\ell} - Y_\tau\right| = O(h_\ell{}^{\frac{1}{2}-\eta})$ and then $\left|\overline{Y}_{t_i^\ell}^{2^\ell} - Y_\tau\right| < h_\ell{}^{\frac{1}{2}-\eta(1+\varepsilon)}$, which yields that $\overline{Y}_{t_i^\ell}^{2^\ell} < \mathcal{D}$. Once again, as we are in the case of non extreme paths we have $|\overline{Y}_{t_i^\ell}^{2^\ell} - \overline{Y}_{t_i^\ell}^{2^{\ell-1}}| < h_\ell{}^{1-\eta}$ and so $\left|\overline{Y}_{t_i^\ell}^{2^{\ell-1}} - Y_\tau\right| = O(h_\ell{}^{\frac{1}{2}-\eta})$. Hence, we also have $\overline{Y}_{t_i^\ell}^{2^{\ell-1}} < \mathcal{D}$, for sufficiently small h_ℓ which yields that $\overline{Q}_\ell^f - \overline{Q}_\ell^c = 0$.

- Second case: for $Y_\tau > \mathcal{D} + h_\ell{}^{\frac{1}{2}-\eta(1+\varepsilon)}$, we proceed in the same way and we easily check that for h_ℓ sufficiently small $\prod_{i=0}^{2^\ell-1}(1-\overline{q}_i^{2^\ell})$ and $\prod_{i=0}^{2^\ell-1}(1-\overline{q}_i^{2^{\ell-1}})$ are both equal to $1 + o(h_\ell{}^a)$ for all $a > 0$. Consequently, as the payoff function g satisfies condition (2.23) and as we work with the non-extreme paths events, we deduce that $\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_2}] = O(h_\ell{}^{2(1-\eta)-2\eta\nu})$.

Third event A_3 This last event corresponds to the rest of the non extreme paths. At first, let us note that

$$\begin{aligned}|\overline{Y}_{t_{i+1}^\ell}^{2^\ell} - \overline{Y}_{t_i^\ell}^{2^\ell}| &\leq |\overline{Y}_{t_{i+1}^\ell}^{2^\ell} - Y_{t_{i+1}^\ell}| + |Y_{t_{i+1}^\ell} - Y_{t_i^\ell}| + |\overline{Y}_{t_i^\ell}^{2^\ell} - Y_{t_i^\ell}| \\ &\leq |\overline{Y}_{t_i^\ell}^{2^\ell} - Y_{t_i^\ell}| + |\overline{Y}_{t_i^\ell}^{2^\ell} - Y_{t_i^\ell}| + \int_{t_i^\ell}^{t_{i+1}^\ell} |L(Y_s)|ds + |\gamma| \sup_{t\in[t_i^\ell, t_{i+1}^\ell]} |W_t - W_{t_i^\ell}|.\end{aligned}$$

Then, similarly as above we deduce that $\left|\overline{Y}_{t_{i+1}^\ell}^{2^\ell} - \overline{Y}_{t_i^\ell}^{2^\ell}\right| = O(h_\ell{}^{\frac{1}{2}-\eta})$ since we work on the non-extreme paths events. Thus, it is clear that if any one of $\overline{Y}_{t_i^\ell}^{2^\ell}, \overline{Y}_{t_{i+1}^\ell}^{2^\ell}, \overline{Y}_{t_i^\ell}^{2^{\ell-1}}, \overline{Y}_{t_{i+1}^\ell}^{2^{\ell-1}}$ is greater than $\mathcal{D} + h_\ell{}^{\frac{1}{2}-\eta(1+\varepsilon)}$, then the others will be greater than $\mathcal{D} + \frac{1}{2}h_\ell{}^{\frac{1}{2}-\eta(1+\varepsilon)}$, for h_ℓ sufficiently

small. In this case, if R denotes the set of indices for which none of $\overline{Y}^{2^\ell}_{t_i^\ell}, \overline{Y}^{2^\ell}_{t_{i+1}^\ell}, \overline{Y}^{2^{\ell-1}}_{t_i^\ell}, \overline{Y}^{2^{\ell-1}}_{t_{i+1}^\ell}$ is greater than $\mathcal{D} + {h_\ell}^{\frac{1}{2}-\eta(1+\varepsilon)}$ and R^c its complementary set, then similarly as above we get $\prod_{i\in R^c}(1-\overline{q}_i^{2^\ell}) = 1 + o({h_\ell}^a)$ and $\prod_{i\in R^c}(1-\overline{q}_i^{2^{\ell-1}}) = 1 + o({h_\ell}^a)$ for all $a > 0$. Thus, we have

$$\prod_{i=0}^{2^\ell-1}(1-\overline{q}_i^{2^\ell}) = \prod_{i\in R}(1-\overline{q}_i^{2^\ell}) + o({h_\ell}^a) \text{ and } \prod_{i=0}^{2^\ell-1}(1-\overline{q}_i^{2^{\ell-1}}) = \prod_{i\in R}(1-\overline{q}_i^{2^{\ell-1}}) + o({h_\ell}^a), \text{ for all } a > 0. \tag{2.25}$$

Now, for $i \in R$ we have

$$\begin{aligned}
|\log\overline{q}_i^{2^\ell} - \log\overline{q}_i^{2^{\ell-1}}| &= \frac{2}{\gamma^2 h_\ell}\Big|(\overline{Y}^{2^\ell}_{t_i^\ell} - \mathcal{D})_+(\overline{Y}^{2^\ell}_{t_{i+1}^\ell} - \mathcal{D})_+ - (\overline{Y}^{2^{\ell-1}}_{t_i^\ell} - \mathcal{D})_+(\overline{Y}^{2^{\ell-1}}_{t_{i+1}^\ell} - \mathcal{D})_+\Big| \\
&\le \frac{1}{\gamma^2 h_\ell}\Big|(\overline{Y}^{2^\ell}_{t_i^\ell} - \mathcal{D})_+ - (\overline{Y}^{2^{\ell-1}}_{t_i^\ell} - \mathcal{D})_+\Big|\Big|(\overline{Y}^{2^\ell}_{t_{i+1}^\ell} - \mathcal{D})_+ + (\overline{Y}^{2^{\ell-1}}_{t_{i+1}^\ell} - \mathcal{D})_+\Big| \\
&+ \frac{1}{\gamma^2 h_\ell}\Big|(\overline{Y}^{2^\ell}_{t_i^\ell} - \mathcal{D})_+ + (\overline{Y}^{2^{\ell-1}}_{t_i^\ell} - \mathcal{D})_+\Big|\Big|(\overline{Y}^{2^\ell}_{t_{i+1}^\ell} - \mathcal{D})_+ - (\overline{Y}^{2^{\ell-1}}_{t_{i+1}^\ell} - \mathcal{D})_+\Big|,
\end{aligned}$$

where we used the relation $f_1g_1 - f_2g_2 = \frac{1}{2}(f_1 - f_2)(g_1 + g_2) + \frac{1}{2}(f_1 + f_2)(g_1 - g_2)$. By the Lipschitz property of the map $x \mapsto (x-\mathcal{D})_+$ and as $i \in R$, there exists a positive constant C that may vary from line to line such that

$$\begin{aligned}
|\log\overline{q}_i^{2^\ell} - \log\overline{q}_i^{2^{\ell-1}}| &\le C{h_\ell}^{-\frac{1}{2}-\eta(1+\varepsilon)}\Big(|\overline{Y}^{2^\ell}_{t_i^\ell} - \overline{Y}^{2^{\ell-1}}_{t_i^\ell}| + |\overline{Y}^{2^\ell}_{t_{i+1}^\ell} - \overline{Y}^{2^{\ell-1}}_{t_{i+1}^\ell}|\Big) \\
&\le C{h_\ell}^{\frac{1}{2}-\eta(2+\varepsilon)}.
\end{aligned}$$

So, for h_ℓ sufficiently small we have $|\log\overline{q}_i^{2^\ell} - \log\overline{q}_i^{2^{\ell-1}}| < {h_\ell}^{\frac{1}{2}-2\eta(1+\varepsilon)}$. Thus, we have

$$\begin{aligned}
1-\overline{q}_i^{2^{\ell-1}} &= (1-\overline{q}_i^{2^\ell}) + \overline{q}_i^{2^\ell}\big(1-\exp\Big(\log\overline{q}_i^{2^{\ell-1}} - \log\overline{q}_i^{2^\ell}\Big)\big) \\
&\le (1-\overline{q}_i^{2^\ell}) + \overline{q}_i^{2^\ell}\big(1-\exp\Big(-{h_\ell}^{\frac{1}{2}-2\eta(1+\varepsilon)}\Big)\big).
\end{aligned}$$

Consequently, we get

$$\begin{aligned}
\prod_{i\in R}(1-\overline{q}_i^{2^{\ell-1}}) &\le \prod_{i\in R}\Big((1-\overline{q}_i^{2^\ell}) + \overline{q}_i^{2^\ell}\big(1-\exp\Big(-{h_\ell}^{\frac{1}{2}-2\eta(1+\varepsilon)}\Big)\big)\Big) \\
&\le \prod_{i\in R}(1-\overline{q}_i^{2^\ell}) + \big(1-\exp\Big(-{h_\ell}^{\frac{1}{2}-2\eta(1+\varepsilon)}\Big)\big),
\end{aligned}$$

where we used the convexity property of the function $\phi : x \in [0,1] \mapsto \prod_{i\in R}\Big((1-\overline{q}_i^{2^\ell}) + \overline{q}_i^{2^\ell}x\Big) - \prod_{i\in R}(1-\overline{q}_i^{2^\ell}) - x$ and that $\phi(0)$ and $\phi(1)$ are both non-positive. In the same way, we get $\prod_{i\in R}(1-\overline{q}_i^{2^\ell}) \le \prod_{i\in R}(1-\overline{q}_i^{2^{\ell-1}}) + \big(1-\exp\Big(-{h_\ell}^{\frac{1}{2}-2\eta(1+\varepsilon)}\Big)\big)$ and we deduce that

$$\Big|\prod_{i\in R}(1-\overline{q}_i^{2^\ell}) - \prod_{i\in R}(1-\overline{q}_i^{2^{\ell-1}})\Big| \le \big(1-\exp\Big(-{h_\ell}^{\frac{1}{2}-2\eta(1+\varepsilon)}\Big)\big) \le {h_\ell}^{\frac{1}{2}-2\eta(1+\varepsilon)}.$$

Then, by (2.25) we get

$$\Big|\prod_{i=0}^{2^\ell-1}(1-\overline{q}_i^{2^\ell}) - \prod_{i=0}^{2^\ell-1}(1-\overline{q}_i^{2^{\ell-1}})\Big| = O({h_\ell}^{\frac{1}{2}-2\eta(1+\varepsilon)}).$$

Therefore, as we work on the non-extreme paths events, we deduce using condition (2.23) on the payoff function g that

$$\begin{aligned}
|\overline{Q}_\ell^f - \overline{Q}_\ell^c|^2 &= \left| g(\overline{Y}_T^{2^\ell}) \prod_{i=0}^{2^\ell-1} (1-\overline{q}_i^{2^\ell}) - g(\overline{Y}_T^{2^{\ell-1}}) \prod_{i=0}^{2^\ell-1} (1-\overline{q}_i^{2^{\ell-1}}) \right|^2 \\
&\leq C|\overline{Y}_T^{2^\ell} - \overline{Y}_T^{2^{\ell-1}}|^2 \left(1 + |\overline{Y}_T^{2^\ell}|^{2\nu} + |\overline{Y}_T^{2^{\ell-1}}|^{2\nu}\right) + {h_\ell}^{1-4\eta(1+\varepsilon)} |g(\overline{Y}_T^{2^\ell})|^2 \\
&\leq C\left({h_\ell}^{2(1-\eta)-2\eta\nu} + {h_\ell}^{1-4\eta(1+\varepsilon)-2\eta-2\eta\nu}\right) \\
&\leq C{h_\ell}^{1-6\eta(1+\varepsilon)-2\eta\nu},
\end{aligned}$$

where C is a positive constant that may varies from line to line. Thus,

$$\begin{aligned}
\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_3}] &= O({h_\ell}^{1-6\eta(1+\varepsilon)-2\eta\nu} \times \mathbb{P}(|\inf_{t\in[0,T]} Y_t - \mathcal{D}| \leq {h_\ell}^{\frac{1}{2}-\eta(1+\varepsilon)})) \\
&= O({h_\ell}^{\frac{3}{2}-7\eta(1+\varepsilon)-2\eta\nu})
\end{aligned}$$

since the random variable $\inf_{t\in[0,T]} Y_t$ has a bounded density on the neighbuorhood of D. To complete the proof, we choose

$$\eta = \frac{\frac{1}{2}-\delta}{7(1+\varepsilon)+2\nu},$$

which yields $\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_3}] = O(h_\ell^{1+\delta})$. Now concerning the second event noticing that $\eta \leq \frac{\frac{1}{2}-\delta}{7+2\nu}$ we easily see that $2 - 2\eta(1+2\nu) > 1+\delta$, for $\delta \in (0, 1/2)$, which yields $\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_2}] = o(h_\ell^{1+\delta})$. Finally, for the first event, we choose $q = (1+\gamma)(1+\delta)$ to guarantee that $\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_1}] = O(h_\ell^{1+\delta})$ which is satisfied as soon as $p > \frac{(1+\delta)(1+\gamma)[7(1+\varepsilon)+2\nu]}{\frac{1}{2}-\delta}$. □

The following result considers a different assumption on the payoff function, covering an application for the local volatility CEV model.

Theorem 2.3.5. *Let g denotes a bounded payoff function satisfying : $\exists C > 0$ s.t. $\forall x, y > 0$,*

$$|g(x) - g(y)| \leq C|x-y|(1 + |x|^{-\nu} + |x|^{-\nu}), \text{ for } \nu > 0. \tag{2.26}$$

Moreover, assume that conditions ($\tilde{\text{H}}$2), (H3) *and* (H4) *are satisfied for $p > \frac{(1+\delta)(1+\gamma)5(1+\varepsilon)}{\frac{1}{2}-\delta}$, with $\varepsilon, \gamma > 0$, $\delta \in (0, 1/2)$ and $0 < L'_A < \frac{1}{2h_\ell}$, with $h_\ell = 2^{-\ell}T$ sufficiently small. If in addition $\inf_{t\in[0,T]} Y_t$ has a bounded density in the neighborhood of the barrier, then the multilevel estimator $\bar{Q}_{\mathcal{D}}$ given by* (3.13) *for the D-O barrier option satisfies* $\mathrm{Var}(\overline{\mathrm{Q}}_\ell^{\mathrm{f}} - \overline{\mathrm{Q}}_\ell^{\mathrm{c}}) = \mathrm{O}(\mathrm{h}_\ell^{1+\delta})$.

Proof. We use a similar decomposition as in the proof of Theorem 2.3.3

$$\begin{aligned}
\mathrm{Var}[\overline{\mathrm{Q}}_\ell^{\mathrm{f}} - \overline{\mathrm{Q}}_\ell^{\mathrm{c}}] &\leq \mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2] \\
&= \mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_1}] + \mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_2}] + \mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_3}]
\end{aligned}$$

where we split the paths into the following three events.

First event A_1 We consider any of the extreme path events given in Lemma 2.3.2 that satisfy (2.18)-(2.22), with some $\eta > 0$ to be fixed later on. For $\gamma > 0$, we use Hölder's inequality to get

$$\begin{aligned}\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_1}] &\leq \mathbb{E}^{\frac{\gamma}{1+\gamma}}\left[|\overline{Q}_\ell^f - \overline{Q}_\ell^c|^{\frac{2(1+\gamma)}{\gamma}}\right]\mathbb{E}^{\frac{1}{1+\gamma}}[\mathbb{1}_{A_1}] \\ &\leq 2^{\frac{2+\gamma}{1+\gamma}}\left(\mathbb{E}^{\frac{\gamma}{1+\gamma}}[|\overline{Q}_\ell^f|^{\frac{2(1+\gamma)}{\gamma}}] + \mathbb{E}^{\frac{\gamma}{1+\gamma}}[|\overline{Q}_\ell^c|^{\frac{2(1+\gamma)}{\gamma}}]\right)\left(\mathbb{P}[A_1]\right)^{\frac{1}{1+\gamma}}.\end{aligned}$$

As the payoff function f is bounded, we deduce by Lemma 2.3.2 that for h_ℓ sufficiently small

$$\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_1}] = o(h_\ell^{\frac{q}{1+\gamma}}) \text{ for all } q \text{ such that } 0 < \frac{q}{\eta} \leq p. \tag{2.27}$$

Second event A_2 This event corresponds to the non-extreme paths satisfying

$$|\inf_{t\in[0,T]} Y_t - \mathcal{D}| > h_\ell^{\frac{1}{2}-\eta(1+\varepsilon)} \text{ for } \eta \in (0, 1/2(1+\varepsilon)) \text{ with } \varepsilon > 0.$$

Let us assume that $\inf_{t\in[0,T]} Y_t = Y_\tau$ with $\tau \in [t_i^\ell, t_{i+1}^\ell]$ for a given $i \in \{0, \ldots, 2^\ell\}$. Now, we write

$$\begin{aligned}|\overline{Y}_{t_i^\ell}^{2^\ell} - Y_\tau| &\leq |\overline{Y}_{t_i^\ell}^{2^\ell} - Y_{t_i^\ell}| + |Y_{t_i^\ell} - Y_\tau| \\ &\leq |\overline{Y}_{t_i^\ell}^{2^\ell} - Y_{t_i^\ell}| + \int_{t_i^\ell}^{t_{i+1}^\ell} |L(Y_s)|ds + \gamma \sup_{t\in[t_i^\ell, t_{i+1}^\ell]} |W_t - W_{t_i^\ell}|.\end{aligned}$$

Then, as we work on the non-extreme paths events then for h_ℓ sufficiently small we have that $\left|\overline{Y}_{t_i^\ell}^{2^\ell} - Y_\tau\right| = O(h_\ell^{\frac{1}{2}-\eta})$ and therefore $\left|\overline{Y}_{t_i^\ell}^{2^\ell} - Y_\tau\right| < h_\ell^{\frac{1}{2}-\eta(1+\varepsilon)}$. Similarly, we deduce that $\left|\overline{Y}_{t_i^\ell}^{2^{\ell-1}} - Y_\tau\right| < h_\ell^{\frac{1}{2}-\eta(1+\varepsilon)}$ since on the non-extreme paths events we have that $|\overline{Y}_{t_i^\ell}^{2^\ell} - \overline{Y}_{t_i^\ell}^{2^{\ell-1}}| < h_\ell{}^{1-\eta}$.

- First case: for $Y_\tau < \mathcal{D} - h_\ell^{\frac{1}{2}-\eta(1+\varepsilon)}$, the above estimate yields that $\overline{Y}_{t_i^\ell}^{2^\ell} < \mathcal{D}$. and so $\left|\overline{Y}_{t_i^\ell}^{2^{\ell-1}} - Y_\tau\right| = O(h_\ell^{\frac{1}{2}-\eta})$. Hence, we also have $\overline{Y}_{t_i^\ell}^{2^{\ell-1}} < \mathcal{D}$, for sufficiently small h_ℓ which yields that $\overline{Q}_\ell^f - \overline{Q}_\ell^c = 0$.
- Second case: for $Y_\tau > \mathcal{D} + h_\ell^{\frac{1}{2}-\eta(1+\varepsilon)}$, we have

$$Y_{t_i^\ell} > \mathcal{D} + h_\ell^{\frac{1}{2}-\eta(1+\varepsilon)}.$$

So as $|\overline{Y}_{t_i^\ell}^{2^\ell} - Y_{t_i^\ell}| < h_\ell{}^{1-\eta}$ then $\overline{Y}_{t_i^\ell}^{2^\ell} > \mathcal{D} + h_\ell^{\frac{1}{2}-\eta(1+\varepsilon)}$. Using similar arguments we also get $\overline{Y}_{t_i^\ell}^{2^{\ell-1}} > \mathcal{D} + h_\ell^{\frac{1}{2}-\eta(1+\varepsilon)}$. Then, we easily check that for h_ℓ sufficiently small $\prod_{i=0}^{2^\ell-1}(1-\overline{q}_i^{2^\ell})$ and $\prod_{i=0}^{2^\ell-1}(1-\overline{q}_i^{2^{\ell-1}})$ are both equal to $1 + o(h_\ell{}^a)$ for all $a > 0$. Besides, by Taylor formula we have $|g(x) - g(y)| \leq \frac{[f]_{\text{Lip}}}{\nu}|x-y|(|x|^{-\nu-1} + |y|^{-\nu-1})$ for all $x, y \in I = (0, +\infty)$. Therefore, using that $\overline{Y}_{t_i^\ell}^{2^\ell} > \mathcal{D}$ and $\overline{Y}_{t_i^\ell}^{2^{\ell-1}} > \mathcal{D}$ we deduce that $\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_2}] = O(h_\ell{}^{2(1-\eta)})$.

Third event A_3 For this last event, we proceed exactly as in Theorem 2.3.3 to get

$$\left|\prod_{i=0}^{2^\ell-1}(1-\overline{q}_i^{2^\ell}) - \prod_{i=0}^{2^\ell-1}(1-\overline{q}_i^{2^{\ell-1}})\right| = O(h_\ell{}^{\frac{1}{2}-2\eta(1+\varepsilon)}).$$

Therefore, as we work on the non-extreme paths events, we deduce using condition (2.23) on the payoff function g that

$$|\overline{Q}_\ell^f - \overline{Q}_\ell^c|^2 = \left| g(\overline{Y}_T^{2^\ell}) \prod_{i=0}^{2^\ell-1} (1-\overline{q}_i^{2^\ell}) - g(\overline{Y}_T^{2^{\ell-1}}) \prod_{i=0}^{2^\ell-1} (1-\overline{q}_i^{2^{\ell-1}}) \right|^2$$
$$\leq C|\overline{Y}_T^{2^\ell} - \overline{Y}_T^{2^{\ell-1}}|^2 \left(1 + |\overline{Y}_T^{2^\ell}|^{-2\nu} + |\overline{Y}_T^{2^{\ell-1}}|^{-2\nu}\right) + {h_\ell}^{1-4\eta(1+\varepsilon)} |g(\overline{Y}_T^{2^\ell})|^2.$$

where C is a positive constant that may varies from line to line. The second term of the above upper bound is clearly $O({h_\ell}^{1-4\eta(1+\varepsilon)})$ since g is bounded. For the first term, let us recall that we are in the case where $Y_T \geq Y_\tau \geq D - {h_\ell}^{\frac{1}{2}-\eta(1+\varepsilon)}$ and since on the non-extreme paths events we have that

$$\overline{Y}_T^{2^\ell} - Y_T > -{h_\ell}^{1-\eta(1+\varepsilon)} \text{ and } \overline{Y}_T^{2^\ell-1} - Y_T > -{h_\ell}^{1-\eta(1+\varepsilon)}$$

then

$$\overline{Y}_T^{2^\ell} > D - {h_\ell}^{\frac{1}{2}-\eta(1+\varepsilon)} - {h_\ell}^{1-\eta(1+\varepsilon)} > \frac{D}{2} \text{ for } h_\ell \text{ small enough.}$$

By similar arguments we get that $\overline{Y}_T^{2^{\ell-1}} > \frac{D}{2}$ for h_ℓ small enough. Then, since $\nu > 0$, we get

$$|\overline{Y}_T^{2^\ell} - \overline{Y}_T^{2^{\ell-1}}|^2 \left(1 + |\overline{Y}_T^{2^\ell}|^{-2(\nu+1)} + |\overline{Y}_T^{2^{\ell-1}}|^{-2(\nu+1)}\right) = O({h_\ell}^{2-2\eta})$$

Therefore, we have

$$\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_3}] = O({h_\ell}^{1-4\eta(1+\varepsilon)} \times \mathbb{P}(|\inf_{t\in[0,T]} Y_t - \mathcal{D}| \leq {h_\ell}^{\frac{1}{2}-\eta(1+\varepsilon)}))$$
$$= O({h_\ell}^{\frac{3}{2}-5\eta(1+\varepsilon)})$$

since the random variable $\inf_{t\in[0,T]} Y_t$ has a bounded density on the neighborhood of D. To complete the proof, we choose

$$\eta = \frac{\frac{1}{2}-\delta}{5(1+\varepsilon)},$$

which yields $\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_3}] = O(h_\ell^{1+\delta})$. Now concerning the second event noticing that $\eta \leq \frac{\frac{1}{2}-\delta}{5}$ we easily see that $2(1-\eta) > 1+\delta$, for $\delta \in (0, 1/2)$, which yields $\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_2}] = o(h_\ell^{1+\delta})$. Finally, for the first event, we choose $q = (1+\gamma)(1+\delta)$ to guarantee that $\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_1}] = O(h_\ell^{1+\delta})$ which is satisfied as soon as $p > \frac{(1+\delta)(1+\gamma)5(1+\varepsilon)}{\frac{1}{2}-\delta}$. □

2.4 Application to the CIR process

In this section, we consider the problem of pricing D-O and U-O barrier options

$$\pi_{\mathcal{D}} = \mathbb{E}\Big[f(X_T)\mathbb{1}_{\{\inf_{t\in[0,T]} X_t > \mathcal{D}\}}\Big] \text{ and } \pi_{\mathcal{U}} = \mathbb{E}\Big[f(X_T)\mathbb{1}_{\{\sup_{t\in[0,T]} X_t < \mathcal{U}\}}\Big],$$

where f is a Lipschitz payoff function with Lipschitz constant $[f]_{\text{Lip}}$ and $(X_t)_{0\leq t\leq T}$ denotes the Cox-Ingersoll-Ross (CIR) process solution to

$$\left\{ \begin{array}{rcl} dX_t &=& (a-\kappa X_t)dt + \sigma\sqrt{X_t}dW_t \\ X_0 &=& x > 0, \end{array} \right. \tag{2.28}$$

with $a \geq \sigma^2/2$, $\kappa \in \mathbb{R}$, $\sigma > 0$, $X_0 = x > 0$. It is well known that this SDE admits a unique strong positive solution. Applying the Lamperti transformation, the process $(Y_t)_{0\leq t\leq T}$ given by $Y_t = \sqrt{X_t}$ satisfies

$$\left\{ \begin{array}{rcl} dY_t & = & L(Y_t)dt + \gamma dW_t \\ Y_0 & = & \sqrt{x}, \end{array} \right. \tag{2.29}$$

where $L(y) = \dfrac{a-\sigma^2/4}{2y} - \dfrac{\kappa}{2}y$ and $\gamma = \dfrac{\sigma}{2}$. Thus, for $g : x \in \mathbb{R} \mapsto g(x) = f(x^2)$ we get

$$\pi_{\mathcal{D}} = \mathbb{E}\Big[g(Y_T)\mathbb{1}_{\{\inf_{t\in[0,T]} Y_t > \sqrt{\mathcal{D}}\}}\Big] \text{ and } \pi_{\mathcal{U}} = \mathbb{E}\Big[g(Y_T)\mathbb{1}_{\{\sup_{t\in[0,T]} Y_t < \sqrt{\mathcal{U}}\}}\Big].$$

As $a-\sigma^2/4 > 0$, we easily check assumptions (H1) and (H4). Besides, noticing that $\lim\limits_{y\to 0^+} L'(y) = \lim\limits_{y\to 0^+} -\frac{(a-\sigma^2/4)}{2y^2} - \frac{\kappa}{2} = -\infty$, we deduce that L is decreasing on $(0,\tilde{\varepsilon})$ for $\tilde{\varepsilon} > 0$ small enough. It is also globally Lipschitz on $[\tilde{\varepsilon}, +\infty)$ so that assumption (H3) is satisfied with $A = \tilde{\varepsilon}$ and $L'_A = \frac{|a-\sigma^2/4|}{2\varepsilon^2} + \frac{\kappa}{2}$. Now, to check ($\tilde{\text{H}}$2) it is enough to show that

$$\sup_{t\in[0,T]} \mathbb{E}\big[|L'(Y_t)L(Y_t)|^p + |L''(Y_t)|^p + |L'(Y_t)|^{(2\vee p)} + |L(Y_t)|^p\big] < \infty$$

which is clearly satisfied as soon as

$$\sup_{t\in[0,T]} \mathbb{E}\big[Y_t^{-(4\vee 3p)}\big] = \sup_{t\in[0,T]} \mathbb{E}\big[X_t^{-(2\vee\frac{3}{2}p)}\big] < \infty \text{ and } \sup_{t\in[0,T]} \mathbb{E}\big[Y_t\big] = \sup_{t\in[0,T]} \mathbb{E}\big[|X_t|^{\frac{1}{2}}\big] < \infty. \tag{2.30}$$

Recalling that $\sup_{t\in[0,T]} \mathbb{E}\big[X_t^q\big] < \infty$ for all $q > -\frac{2a}{\sigma^2}$ (see e.g. [28, 12]), we easily conclude that condition (2.30) is satisfied when $\sigma^2 < a$ and $p < \frac{4}{3}\frac{a}{\sigma^2}$. Now, as $|g(x)-g(y)| \leq [f]_{\text{Lip}}|x-y|(|x|+|y|)$ for all $x, y \in I = (0,+\infty)$, then g satisfies condition (2.23) with $\nu = 1$. Consequently, for $\delta \in (0,1/2)$, if we choose $\frac{4}{3}\frac{a}{\sigma^2} > p > \frac{(1+\delta)(1+\gamma)[7(1+\varepsilon)+2\nu]}{\frac{1}{2}-\delta} > 18$ for ε, δ close to zero and h_ℓ sufficiently small s.t. $L'_A = \frac{|a-\sigma^2/4|}{2\varepsilon^2} + \frac{\kappa}{2} < \frac{1}{2h_\ell}$ then Theorem 2.3.3 is valid provided that $\inf_{t\in[0,T]} Y_t$ (resp. $\sup_{t\in[0,T]} Y_t$) has a bounded density in the neighborhood of the barrier $\sqrt{\mathcal{D}}$ (resp. $\sqrt{\mathcal{U}}$). More precisely, by the monotone property of $x \in \mathbb{R}_+^* \mapsto \sqrt{x}$ we have the relationship $\inf_{t\in[0,T]} Y_t = \sqrt{\inf_{t\in[0,T]} X_t}$ (resp. $\sup_{t\in[0,T]} Y_t = \sqrt{\sup_{t\in[0,T]} X_t}$), then it is sufficient to prove that $\inf_{t\in[0,T]} X_t$ (resp. $\sup_{t\in[0,T]} X_t$) has a continuous density in the neighborhood of the barrier which is the aim of the following subsection.

2.4.1 Running maximum of the CIR process

The aim of this subsection is to prove that the maximum of the CIR process (2.28) admits a continuous density. To do so, let us introduce firstly the confluent hypergeometric function ${}_1F_1(x,b,y)$ defined for all y, $x \in \mathbb{C}$ and $b \in \mathbb{C} \setminus \{0,-1,-2,\cdots\}$ by

$${}_1F_1(x,b,y) = \sum_{n=0}^{\infty} \frac{(x)_n}{(b)_n n!} y^n, \tag{2.31}$$

where $(x)_n = x(x+1)...(x+n-1)$ stands for the Pochhammer symbol.

Theorem 2.4.1. *Let $(X_t)_{0\leq t\leq T}$ denote the CIR process solution to* (2.28)*. Then $\sup_{t\in[0,T]} X_t$ has a continuous density on any compact set $K \subset (X_0, +\infty)$, given by*

$$z \in K \mapsto P_{\text{CIR,Max}}(z) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{e}^{(1+iu)T}\,\hat{\phi}(u,z)du \tag{2.32}$$

with

$$\hat{\phi}(u,z) = -\frac{{}_1F_1((1+iu)/\kappa, 2a/\sigma^2, 2\kappa X_0/\sigma^2)\,{}_1F_1((1+iu)/\kappa+1, 2a/\sigma^2+1, 2\kappa z/\sigma^2)}{a\,{}_1F_1((1+iu)/\kappa, 2a/\sigma^2, 2\kappa z/\sigma^2)^2}.$$

Proof. At first, let us recall that for the CIR process, we have

$$\begin{aligned}\mathbb{P}\big[\sup_{0\le s\le t} X_s > z\big] &= \frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty} \frac{e^{ts}}{s}\frac{{}_1F_1(s/\kappa, 2a/\sigma^2, 2\kappa X_0/\sigma^2)}{{}_1F_1(s/\kappa, 2a/\sigma^2, 2\kappa z/\sigma^2)}ds,\\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{e}^{(1+iu)t}\,\phi(u,z)du,\\ \text{with } \phi(u,z) &:= \frac{{}_1F_1((1+iu)/\kappa, 2a/\sigma^2, 2\kappa X_0/\sigma^2)}{(1+iu)\,{}_1F_1((1+iu)/\kappa, 2a/\sigma^2, 2\kappa z/\sigma^2)},\end{aligned}$$

where we recall that for $b, y > 0$, the x-zeros of ${}_1F_1(x,b,y)$ are negative real and simple (see e.g. [61] and [33]). Thereafter, by formula (13.4.8) in [32] the derivative of the Kummer confluent hypergeometric function is given by

$$\frac{\partial_1 F_1(a,b,z)}{\partial z} = \frac{a}{b}\,{}_1F_1(a+1, b+1, z), \tag{2.33}$$

which gives that

$$\frac{\partial\phi}{\partial z}(u,z) = -\frac{{}_1F_1((1+iu)/\kappa, 2a/\sigma^2, 2\kappa X_0/\sigma^2)\,{}_1F_1((1+iu)/\kappa+1, 2a/\sigma^2+1, 2\kappa z/\sigma^2)}{a\,{}_1F_1((1+iu)/\kappa, 2a/\sigma^2, 2\kappa z/\sigma^2)^2}.$$

On the one hand, by formula (10.3.53) of [78] we have for $x, b, y \in \mathbb{C}$

$${}_1F_1(x,b,y) \sim \left(\frac{y}{x}\right)^{\frac{1-b}{2}} \frac{\Gamma(b)\Gamma(x-b+1)}{\Gamma(x)} e^{\frac{y}{2}} I_{b-1}(2\sqrt{xy}), \text{ as } x \to +\infty$$

which is valid inside the sector $|\arg(x)| < \pi$ and uniformly in bounded b and y-domains, where I_ν stands for the modified Bessel functions of the first kind. By formula (9.3.14), we have $I_\nu(x) \sim \frac{\mathrm{e}^x}{\sqrt{2\pi x}}$, as $x \to \infty$ uniformly in the sector $|\arg(x)| < \frac{\pi}{2}$, we then deduce that

$$\begin{aligned}{}_1F_1((1+iu)/\kappa + j, 2a/\sigma^2 + j, 2\kappa v/\sigma^2) \sim{}& \frac{\mathrm{e}^{\frac{\kappa v}{\sigma^2}}}{\sqrt{4\pi}}\left(\frac{\frac{2\kappa v}{\sigma^2}}{\frac{iu+1}{\kappa}+j}\right)^{\frac{1}{2}-\frac{j}{2}-\frac{a}{\sigma^2}} \frac{\Gamma(\frac{2a}{\sigma^2}+j)\Gamma(\frac{iu+1}{\kappa}-\frac{2a}{\sigma^2}+1)}{\Gamma(\frac{iu+1}{\kappa}+j)}\\ &\times\left(\frac{2\kappa v}{\sigma^2}\left(\frac{iu+1}{\kappa}+j\right)\right)^{-\frac{1}{4}} \exp\left(2\sqrt{\frac{2\kappa v}{\sigma^2}(\frac{1+iu}{\kappa}+j)}\right)\end{aligned}$$

as $u \to \infty$ uniformly on v-bounded domain. Using that $\Gamma(x+a)/\Gamma(x+b) \sim x^{a-b}$ as $x \to \infty$ uniformly inside the sector $|\arg(x)| < \pi$, (see e.g. (6.5.72) of [78]), and that

$$\exp\left(2\sqrt{\frac{2\kappa v}{\sigma^2}(\frac{1+iu}{\kappa}+j)}\right) \sim \exp\left(\frac{2\sqrt{v}}{\sigma}(1+i)\sqrt{u}\right)$$

as $u \to +\infty$ uniformly in bounded v-domain, we get

$$\begin{aligned}{}_1F_1((1+iu)/\kappa + j, 2a/\sigma^2 + j, 2\kappa v/\sigma^2) \sim{}& \frac{\mathrm{e}^{\frac{\kappa v}{\sigma^2}}}{2\sqrt{\pi}}\left(\frac{2\kappa v}{\sigma^2}\right)^{\frac{1}{4}-\frac{j}{2}-\frac{a}{\sigma^2}} \Gamma(\frac{2a}{\sigma^2}+j)\left(\frac{iu+1}{\kappa}+j\right)^{\frac{1}{4}-\frac{j}{2}-\frac{a}{\sigma^2}}\\ &\times e^{2(1+i)\sqrt{\frac{v}{\sigma^2}u}}\end{aligned} \tag{2.34}$$

as $u \to \infty$ uniformly in bounded v-domain. Thus, we deduce that

$$\frac{\partial \phi}{\partial z}(u,z) \sim -\frac{\sqrt{2}}{\sigma}\Big(X_0\Big)^{-\frac{a}{\sigma^2}+\frac{1}{4}}\Big(z\Big)^{\frac{a}{\sigma^2}-\frac{3}{4}}(1+iu)^{\frac{a}{\sigma^2}-\frac{1}{4}}\big((1+\kappa)+iu\big)^{-\frac{a}{\sigma^2}-\frac{1}{4}}$$
$$\exp\Big((\kappa/\sigma^2)(X_0-z)\Big)\exp\left[\frac{2}{\sigma}\sqrt{u}(1+i)(\sqrt{X_0}-\sqrt{z})\right],$$

and therefore

$$\left|\frac{\partial \phi}{\partial z}(u,z)\right| \sim \frac{\sqrt{2}}{\sigma}\Big(X_0\Big)^{-\frac{a}{\sigma^2}+\frac{1}{4}}\Big(z\Big)^{\frac{a}{\sigma^2}-\frac{3}{4}}\exp\Big[(\kappa/\sigma^2)(X_0-z)\Big]u^{-\frac{1}{2}}\exp\left[\frac{2}{\sigma}\sqrt{u}(\sqrt{X_0}-\sqrt{z})\right]$$

as $u \to \infty$, uniformly in bounded z-domain. Hence, $\int_1^{+\infty}\left|\mathrm{e}^{(1+iu)t}\,\frac{\partial\phi}{\partial z}(u,z)\right|du < \infty$ is uniformly convergent in bounded z-domain too. On the other hand, for the integral from $-\infty$ to -1, by (10.3.58) of [78] we have for $x, b, y \in \mathbb{C}$

$${}_1F_1(-x,b,y) \sim \left(\frac{y}{x}\right)^{\frac{1-b}{2}}\frac{\Gamma(b)\Gamma(x+1)}{\Gamma(x+b)}e^{\frac{y}{2}}J_{b-1}(2\sqrt{xy}),\ \text{as } x \to +\infty$$

which is valid inside the sector $|\arg(x)| < \pi$ and uniformly in bounded b and y-domains, where J_ν stands for the Bessel functions of the first kind. By (9.2.1) of [1], as $|x| \to \infty$ we have that

$$J_\nu(x) = \sqrt{\frac{2}{\pi x}}\left(\cos\left(x-\frac{\nu\pi}{2}-\frac{\pi}{4}\right)+\mathrm{e}^{|\operatorname{Im}(x)|}\,\mathrm{O}\left(|x|^{-1}\right)\right),\ |\arg(x)| < \pi.$$

Combining this with the following standard asymptotic expansions valid for any $\alpha \in \mathbb{R}$, $\beta > 0$ and $u \to +\infty$ $\cos(\alpha+i\beta u) = \frac{1}{2}\,\mathrm{e}^{-i\alpha+\beta u}+o(e^{\beta|u|})$, we get

$${}_1F_1((1-iu)/\kappa+j, 2a/\sigma^2+j, 2\kappa v/\sigma^2) \sim \frac{\mathrm{e}^{\frac{\kappa v}{\sigma^2}}}{\sqrt{\pi}}\left(\frac{\frac{2\kappa v}{\sigma^2}}{\frac{iu-1}{\kappa}-j}\right)^{\frac{1}{2}-\frac{j}{2}-\frac{a}{\sigma^2}}\frac{\Gamma(\frac{2a}{\sigma^2}+j)\Gamma(\frac{iu-1}{\kappa}-j+1)}{\Gamma(\frac{iu-1}{\kappa}+\frac{2a}{\sigma^2})}$$
$$\times\left(\frac{2\kappa v}{\sigma^2}\left(\frac{iu-1}{\kappa}-j\right)\right)^{-\frac{1}{4}}\cos\left(2\sqrt{\frac{2\kappa v}{\sigma^2}\left(\frac{iu-1}{\kappa}-j\right)}+\frac{\pi}{2}(\frac{2a}{\sigma^2}+j-1)+\frac{\pi}{4}\right).$$

Using that

$$\cos\left(2\sqrt{\frac{2\kappa v}{\sigma^2}\left(\frac{iu-1}{\kappa}-j\right)}+\frac{\pi}{2}(\frac{2a}{\sigma^2}+j-1)+\frac{\pi}{4}\right) \sim \frac{1}{2}e^{-i\pi\left(\frac{a}{\sigma^2}+\frac{j}{2}+\frac{1}{4}\right)}e^{2(1-i)\sqrt{\frac{v}{\sigma^2}u}},$$

as $u \to \infty$ uniformly on v-bounded domain and that $\Gamma(z+a)/\Gamma(z+b) \sim z^{a-b}$ as $z \to \infty$ uniformly inside the sector $|\arg(z)| < \pi$, (see e.g. (6.5.72) of [78]), we get

$$\begin{aligned}{}_1F_1((1-iu)/\kappa+j, 2a/\sigma^2+j, 2\kappa v/\sigma^2) &\sim \frac{\mathrm{e}^{\frac{\kappa v}{\sigma^2}}}{2\sqrt{\pi}}\left(\frac{\frac{2\kappa v}{\sigma^2}}{\frac{iu-1}{\kappa}-j}\right)^{\frac{1}{2}-\frac{j}{2}-\frac{a}{\sigma^2}}\Gamma(\frac{2a}{\sigma^2}+j)\left(\frac{iu-1}{\kappa}-j\right)^{1-j-\frac{2a}{\sigma^2}}\\
&\times\left(\frac{2\kappa v}{\sigma^2}\left(\frac{iu-1}{\kappa}-j\right)\right)^{-\frac{1}{4}}e^{-i\pi\left(\frac{a}{\sigma^2}+\frac{j}{2}+\frac{1}{4}\right)}e^{2(1-i)\sqrt{\frac{v}{\sigma^2}u}}\\
&\sim \frac{\mathrm{e}^{\frac{\kappa v}{\sigma^2}}}{2\sqrt{\pi}}\left(\frac{2\kappa v}{\sigma^2}\right)^{\frac{1}{4}-\frac{j}{2}-\frac{a}{\sigma^2}}\Gamma(\frac{2a}{\sigma^2}+j)\left(\frac{iu-1}{\kappa}-j\right)^{\frac{1}{4}-\frac{j}{2}-\frac{a}{\sigma^2}}\\
&\times e^{-i\pi\left(\frac{a}{\sigma^2}+\frac{j}{2}+\frac{1}{4}\right)}e^{2(1-i)\sqrt{\frac{v}{\sigma^2}u}}\end{aligned} \tag{2.35}$$

and therefore

$$\left|\frac{\partial\phi}{\partial z}(-u,z)\right| \sim \frac{\sqrt{2}}{\sigma}(X_0)^{\frac{1}{4}-\frac{a}{\sigma^2}}(z)^{\frac{a}{\sigma^2}-\frac{3}{4}}\exp\Big(\frac{\kappa}{\sigma^2}(X_0-z)\Big)u^{-\frac{1}{2}}\exp\Big(\frac{2}{\sigma}\sqrt{u}(\sqrt{X_0}-\sqrt{z})\Big)$$

as $u\to\infty$, uniformly in bounded z-domain. Hence, we deduce that $\int_{-\infty}^{-1}\left|\mathrm{e}^{(1+iu)t}\,\frac{\partial\phi}{\partial z}(u,z)\right|du<\infty$ uniformly in bounded z-domain. Finally, we complete the proof by noticing that $(u,z)\in\mathbb{R}\times K\mapsto \mathrm{e}^{(1+iu)t}\,\frac{\partial\phi}{\partial z}(u,z)$ is a continuous function for any compact set $K\subset(X_0,+\infty)$. (see e.g. [70, Theorem B.3]) □

2.4.2 Running minimum of the CIR process

In the current subsection, we focus on studying the density of the running minimum of the CIR process (2.28). For this aim, we introduce the Tricomi confluent hypergeometric function $U(a,b,z)$ defined for all $a,\ z\in\mathbb{C}$ and $b\in\mathbb{C}\setminus\{\pm0,\pm1,\pm2,...\}$ by

$$U(a,b,z)=\frac{\Gamma(1-b)}{\Gamma(1+a-b)}{}_1F_1(a,b,z)+\frac{\Gamma(b-1)}{\Gamma(a)}z^{1-b}{}_1F_1(1+a-b,2-b,z). \tag{2.36}$$

Let us denote by $\tau_{X_0\downarrow z}:=\inf\{t\geq 0: X_t=z\}$ the first time that the CIR process $(X_t)_{t\geq0}$ starting at X_0 hits the level z satisfying $0<z<X_0$. By [18, Theorem 3], the Laplace Transform of $\tau_{X_0\downarrow z}$ is explicit and given by

$$\mathbf{E}[\mathrm{e}^{-s\tau_{X_0\downarrow z}}]=\frac{U(s/\kappa,2a/\sigma^2,2\kappa X_0/\sigma^2)}{U(s/\kappa,2a/\sigma^2,2\kappa z/\sigma^2)},\ \text{for } s>0. \tag{2.37}$$

Theorem 2.4.2. *Let $(X_t)_{0\leq t\leq T}$ denote the CIR process solution to* (2.28)*. Then $\inf_{t\in[0,T]}X_t$ has a continuous density on any compact set $K\subset(0,X_0)$, given by*

$$z\in K\mapsto P_{\mathrm{CIR,Min}}(z)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\mathrm{e}^{(1+iu)T}\,\hat{\psi}(u,z)du \tag{2.38}$$

with

$$\hat{\psi}(u,z)=\frac{2U((1+iu)/\kappa,2a/\sigma^2,2\kappa X_0/\sigma^2)U((1+iu)/\kappa+1,2a/\sigma^2+1,2\kappa z/\sigma^2)}{\sigma^2U((1+iu)/\kappa,2a/\sigma^2,2\kappa z/\sigma^2)^2}.$$

Proof. Making use of an inverse Laplace transform, the cumulative distribution function of the running minimum CIR process can be expressed as

$$\begin{aligned}\mathbb{P}\left[\inf_{0\leq s\leq t}X_s\leq z\right]&=\mathbb{P}\left[\tau_{X_0\downarrow z}\leq t\right]\\&=\frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty}\frac{e^{ts}}{s}\frac{U(s/\kappa,2a/\sigma^2,2\kappa X_0/\sigma^2)}{U(s/\kappa,2a/\sigma^2,2\kappa z/\sigma^2)}ds\\&=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\mathrm{e}^{(1+iu)t}\,\psi(u,z)du,\\\text{with }\psi(u,z)&:=\frac{U((1+iu)/\kappa,2a/\sigma^2,2\kappa X_0/\sigma^2)}{(1+iu)U((1+iu)/\kappa,2a/\sigma^2,2\kappa z/\sigma^2)}\end{aligned}$$

and where we recall that for $b,y>0$, the x-zeros of $U(x,b,y)$ are negative real and simple (see e.g. [61]). By formula (13.4.21) in [32] the derivative of the Tricomi confluent hypergeometric function is given by

$$\frac{\partial U(a,b,z)}{\partial z}=-aU(a+1,b+1,z) \tag{2.39}$$

which gives that

$$\frac{\partial\psi(u,z)}{\partial z} = \frac{2U((1+iu)/\kappa, 2a/\sigma^2, 2\kappa X_0/\sigma^2)U((1+iu)/\kappa+1, 2a/\sigma^2+1, 2\kappa z/\sigma^2)}{\sigma^2 U((1+iu)/\kappa, 2a/\sigma^2, 2\kappa z/\sigma^2)^2}.$$

Now using formulas (10.3.37), (10.3.31) and (9.1.3) in [78], we get for x, b and $y \in \mathbb{C}$

$$U(x,b,y) \sim \frac{\sqrt{\pi}}{\Gamma(x)} x^{-\frac{3}{4}+\frac{1}{2}b} y^{\frac{1}{4}-\frac{1}{2}b}\, \mathrm{e}^{\frac{y}{2}-2\sqrt{xy}}, \quad \text{as } x \to +\infty, \tag{2.40}$$

inside the sector $|\arg(x)| < \pi$ and uniformly in bounded b and y-domains. For $j \in \{0,1\}$ and $v \in \{X_0, z\}$, we get

$$U((1+iu)/\kappa + j, 2a/\sigma^2 + j, 2\kappa v/\sigma^2) \sim \frac{\sqrt{\pi}}{\Gamma\left(\frac{1+iu}{\kappa}+j\right)} \exp\left(\frac{\kappa v}{\sigma^2} - 2\sqrt{\frac{2\kappa v}{\sigma^2}(\frac{1+iu}{\kappa}+j)}\right) \left(\frac{2\kappa v}{\sigma^2}\right)^{\frac{1}{4}-\frac{a}{\sigma^2}-\frac{j}{2}} \left(\frac{1+iu}{\kappa}+j\right)^{-\frac{3}{4}+\frac{a}{\sigma^2}+\frac{j}{2}} \tag{2.41}$$

as $u \to \infty$ uniformly in bounded v-domain. Besides, we easily check that

$$\exp\left(\frac{\kappa v}{\sigma^2} - 2\sqrt{\frac{2\kappa v}{\sigma^2}(\frac{1+iu}{\kappa}+j)}\right) \sim \exp\left(\frac{\kappa v}{\sigma^2} - \frac{2\sqrt{v}}{\sigma}(1+i)\sqrt{u}\right), \text{ as } u \to +\infty, \tag{2.42}$$

also uniformly in bounded v-domain. We deduce that

$$\frac{\partial\psi(u,z)}{\partial z} \sim \frac{2}{\sigma^2}\left(\frac{1+iu}{\kappa}\right)^{-\frac{1}{4}-\frac{a}{\sigma^2}} \left(\frac{2\kappa X_0}{\sigma^2}\right)^{\frac{1}{4}-\frac{a}{\sigma^2}} \left(\frac{(1+\kappa)+iu}{\kappa}\right)^{-\frac{1}{4}+\frac{a}{\sigma^2}} \left(\frac{2\kappa z}{\sigma^2}\right)^{-\frac{3}{4}+\frac{a}{\sigma^2}} \exp\Big((\kappa/\sigma^2)(X_0-z)\Big) \exp\left(\frac{2}{\sigma}(1+i)(\sqrt{z}-\sqrt{X_0})\sqrt{u}\right)$$

and then

$$\left|\frac{\partial\psi(u,z)}{\partial z}\right| \sim \frac{\sqrt{2}}{\sigma}(X_0)^{\frac{1}{4}-\frac{a}{\sigma^2}}(z)^{-\frac{3}{4}+\frac{a}{\sigma^2}} \exp\Big((\kappa/\sigma^2)(X_0-z)\Big) u^{-\frac{1}{2}} \exp\left(\frac{2}{\sigma}(\sqrt{z}-\sqrt{X_0})\sqrt{u}\right)$$

as $u \to \infty$ uniformly in bounded z-domain in $(0, X_0)$. Hence, $\int_1^{+\infty} \left|\mathrm{e}^{(1+iu)t}\frac{\partial\psi}{\partial z}(u,z)\right| du < \infty$ is uniformly convergent in any bounded z-domain in $(0, X_0)$.

On the other hand, for the integral from $-\infty$ to -1, we have by formula (10.3.68) of [78] that for x, b and $y \in \mathbb{C}$

$$U(-x,b,y) \sim \left(\frac{y}{x}\right)^{\frac{1-b}{2}} \Gamma(x+1)\, \mathrm{e}^{\frac{y}{2}} \left(\cos(\pi x) J_{b-1}(2\sqrt{xy}) + \sin(\pi x) Y_{b-1}(2\sqrt{xy})\right), \quad \text{as } x \to +\infty, \tag{2.43}$$

which is valid inside the sector $|\arg(x)| < \pi$ and uniformly in bounded b and y-domains, where J_ν and Y_ν stand for the Bessel functions of respectively the first and the second kind. By (9.2.1) and (9.2.2) of [1] we have as $|x| \to \infty$

$$J_\nu(x) = \sqrt{\frac{2}{\pi x}}\left(\cos\left(x - \frac{\nu\pi}{2} - \frac{\pi}{4}\right) + \mathrm{e}^{|\operatorname{Im}(x)|}\,\mathrm{O}\left(|x|^{-1}\right)\right), \ |\arg(x)| < \pi \tag{2.44}$$

$$Y_\nu(x) = \sqrt{\frac{2}{\pi x}}\left(\sin\left(x - \frac{\nu\pi}{2} - \frac{\pi}{4}\right) + \mathrm{e}^{|\operatorname{Im}(x)|}\,\mathrm{O}\left(|x|^{-1}\right)\right), \ |\arg(x)| < \pi. \tag{2.45}$$

Combining all this with the following standard asymptotic expansions valid for any $\alpha \in \mathbb{R}$, $\beta > 0$ and $u \to +\infty$ $\cos(\alpha + i\beta u) = O(e^{\beta u})$, $\sin(\alpha + i\beta u) = O(e^{\beta u})$ and with the relation $\cos(z_1)\cos(z_2) + \sin(z_1)\sin(z_2) = \cos(z_1 - z_2)$, $z_1, z_2 \in \mathbb{C}$, we get

$$U\left(-(\frac{iu-1}{\kappa} - j), \frac{2a}{\sigma^2 + j}, 2\frac{\kappa v}{\sigma^2}\right) \sim \frac{\mathrm{e}^{\frac{\kappa v}{\sigma^2}}}{\sqrt{\pi}} \left(\frac{\frac{2\kappa v}{\sigma^2}}{\frac{iu-1}{\kappa} - j}\right)^{\frac{1-j}{2} - \frac{a}{\sigma^2}} \left(\frac{2\kappa v}{\sigma^2}\left(\frac{iu-1}{\kappa} - j\right)\right)^{-\frac{1}{4}}$$
$$\times \Gamma\left(\frac{iu-1}{\kappa} - j + 1\right) \cos\left(\pi\left(\frac{iu-1}{\kappa} - j\right) - 2\sqrt{\frac{2\kappa v}{\sigma^2}\left(\frac{iu-1}{\kappa} - j\right)} + \frac{\pi}{2}(\frac{2a}{\sigma^2} + j - 1) + \frac{\pi}{4}\right)$$

as $u \to +\infty$ uniformly in bounded v-domain. Using that

$$\cos\left(\pi\left(\frac{iu-1}{\kappa} - j\right) - 2\sqrt{\frac{2\kappa v}{\sigma^2}\left(\frac{iu-1}{\kappa} - j\right)} + \frac{\pi}{2}(\frac{2a}{\sigma^2} + j - 1) + \frac{\pi}{4}\right)$$
$$\sim \frac{1}{2} e^{-i\pi\left(\frac{a}{\sigma^2} - \frac{1}{\kappa} - \frac{j}{2} + \frac{1}{4}\right)} e^{\frac{\pi}{\kappa}u - 2(1-i)\sqrt{\frac{v}{\sigma^2}u}},$$

we get

$$U\left(-(\frac{iu-1}{\kappa} - j), \frac{2a}{\sigma^2 + j}, 2\frac{\kappa v}{\sigma^2}\right) \sim \frac{e^{\frac{\kappa v}{\sigma^2}}}{2\sqrt{\pi}} e^{-i\pi\left(\frac{a}{\sigma^2} - \frac{1}{\kappa} - \frac{j}{2} + \frac{1}{4}\right)} e^{\frac{\pi}{\kappa}u - 2(1-i)\sqrt{\frac{v}{\sigma^2}u}} \left(\frac{2\kappa v}{\sigma^2}\right)^{\frac{1}{4} - \frac{j}{2} - \frac{a}{\sigma^2}}$$
$$\times \Gamma\left(\frac{iu-1}{\kappa} - j + 1\right)\left(\frac{iu-1}{\kappa} - j\right)^{-\frac{3}{4} + \frac{j}{2} + \frac{a}{\sigma^2}},$$

as $u \to +\infty$ uniformly in bounded v-domain. Thus,

$$\left|\frac{\partial \psi(u,z)}{\partial z}\right| \sim \frac{\sqrt{2}}{\sigma} (X_0)^{\frac{1}{4} - \frac{a}{\sigma^2}} (z)^{-\frac{3}{4} + \frac{a}{\sigma^2}} \exp\Big((\kappa/\sigma^2)(X_0 - z)\Big) u^{-\frac{1}{2}} \exp\left(\frac{2}{\sigma}(\sqrt{z} - \sqrt{X_0})\sqrt{u}\right)$$

as $u \to \infty$ uniformly in bounded z-domain in $(0, X_0)$. Hence, we deduce that

$$\int_{-\infty}^{-1} \left|\mathrm{e}^{(1+iu)t} \frac{\partial \psi}{\partial z}(u,z)\right| du < \infty$$

uniformly in any bounded z-domain subset of $(0, X_0)$. Finally, we complete the proof by noticing that $(u,z) \in \mathbb{R} \times K \mapsto \mathrm{e}^{(1+iu)t} \frac{\partial \psi}{\partial z}(u,z)$ is a continuous function for any compact set $K \subset (0, X_0)$ (see e.g. [70, Theorem B.3]). □

2.4.3 Numerical tests

For these numerical tests, we consider the problem of pricing D-O and U-O barrier options $\pi_{\mathcal{D}} = \mathbb{E}\Big[f(X_T)\mathbb{1}_{\{\inf_{t\in[0,T]} X_t > \mathcal{D}\}}\Big]$ and $\pi_{\mathcal{U}} = \mathbb{E}\Big[f(X_T)\mathbb{1}_{\{\sup_{t\in[0,T]} X_t < \mathcal{U}\}}\Big]$, where the payoff function $f(x) = \mathrm{e}^{-rT}(x - K)_+$ and $(X_t)_{0\le t\le T}$ is the CIR process solution to (2.28). By the Lamperti transform we get

$$\pi_{\mathcal{D}} = \mathbb{E}\Big[g(Y_T)\mathbb{1}_{\{\inf_{t\in[0,T]} Y_t > \sqrt{\mathcal{D}}\}}\Big] \text{ and } \pi_{\mathcal{U}} = \mathbb{E}\Big[g(Y_T)\mathbb{1}_{\{\sup_{t\in[0,T]} Y_t < \sqrt{\mathcal{U}}\}}\Big],$$

where $g(x) = \mathrm{e}^{-rT}(x^2 - K)_+$ and $(Y_t)_{t\in[0,T]}$ is solution to (2.29). We approximate $\pi_{\mathcal{D}}$ (resp. $\pi_{\mathcal{U}}$) by the improved MLMC algorithm $\bar{Q}_{\mathcal{D}}$ given in (3.13) (resp. $\bar{P}_{\mathcal{U}}$ given in (3.12)), where we used our interpolated drift implicit scheme

$$\overline{Y}_t^n = \overline{Y}_{t_i}^n + \left(\frac{a - \gamma^2}{2\overline{Y}_{t_{i+1}}^n} - \frac{\kappa}{2}\overline{Y}_{t_{i+1}}^n\right)(t - t_i) + \gamma(W_t - W_{t_i}), \qquad \text{for } t \in [t_i, t_{i+1}]$$
$$Y_0 = \sqrt{X_0},$$

with $\gamma = \frac{\sigma}{2}$. For n large enough, the positive solution to the above implicit scheme is explicit and given by

$$\overline{Y}^n_{t_{i+1}} = \frac{\sqrt{(2+\kappa\frac{T}{n})(a-\gamma^2)\frac{T}{n} + (\gamma(W_{t_{i+1}} - W_{t_i}) + \overline{Y}^n_{t_i})^2} + \gamma(W_{t_{i+1}} - W_{t_i}) + \overline{Y}^n_{t_i}}{2+\kappa\frac{T}{n}}.$$

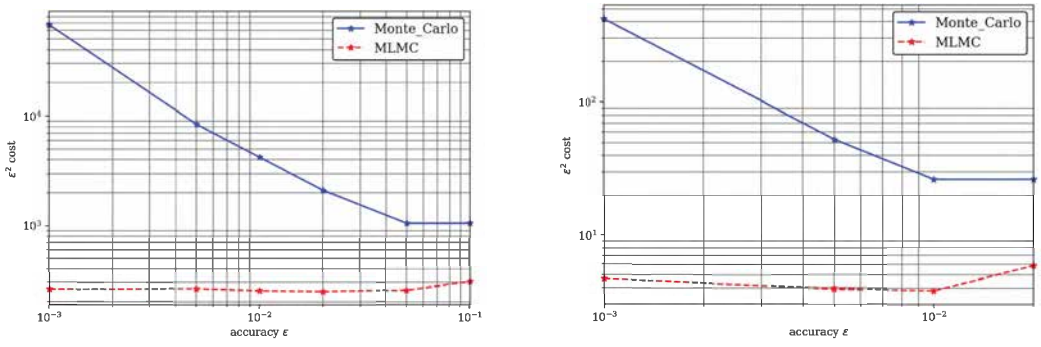
To illustrate the performance of our MLMC algorithms we consider the same comparison procedure as in [35]. We considered the same model and option parameters proposed by [25]. We take $r = 0.1$, $X_0 = 100$, $a = 0$, $\kappa = -0.1$, $\sigma = 2.5$ and $T = 0.5$. For the D-O option the strike is $K = 95$, and the barrier $\mathcal{D} = 90$ and for the U-O option the strike is $K = 105$ and the barrier $\mathcal{U} = 120$. The benchmark prices given in [25] for the D-O (resp. U-O) option is 10.6013 (resp. 0.7734). The performance of the improved MLMC is given in the tables and figure below.

Accuracy	Price	MLMC cost	MC cost	Saving
10^{-3}	10.669	2.588×10^8	6.752×10^{10}	260.91
5×10^{-3}	10.668	1.051×10^7	3.376×10^8	32.13
10^{-2}	10.668	2.510×10^6	4.220×10^7	16.81
2×10^{-2}	10.677	6.187×10^5	5.275×10^6	8.52

Table 2.1: MLMC complexity tests for D-O barrier option pricing of $\pi_{\mathcal{D}}$

Accuracy	Price	MLMC cost	MC cost	Saving
10^{-3}	0.77200	4.674×10^6	4.221×10^8	90.32
5×10^{-3}	0.76926	1.571×10^5	2.11×10^6	13.44
10^{-2}	0.77015	3.809×10^4	2.638×10^5	6.93
2×10^{-2}	0.78168	1.463×10^4	6.596×10^4	4.51

Table 2.2: MLMC complexity tests for U-O barrier option pricing $\pi_{\mathcal{U}}$



(a) Approximation of $\pi_{\mathcal{D}}$ **(b)** Approximation of $\pi_{\mathcal{U}}$

Fig 2.1. Comparison for the performances of MLMC vs classical MC algorithm under the CIR model.

The numerical results illustrates Theorem 2.3.3, that is the improved MLMC algorithm reaches for a given precision ε the optimal time complexity $O(\varepsilon^{-2})$ for D-O and U-O barrier options under the CIR model.

2.5 Application to CEV process

In this section, we consider the CEV process solution to

$$dX_t = \mu X_t dt + \sigma X_t^{\alpha} dW_t, \;\; t \geq 0, \;\; , X_0 > 0, \;\; \mu \in \mathbb{R} \;\; \text{and } \alpha > 1. \tag{2.46}$$

We consider the problem of pricing an U-O barrier option $\Pi_{\mathcal{D}}^{\text{U-O},X} := \mathbb{E}\Big[f(X_T)\mathbb{1}_{\{\sup_{t\in[0,T]} X_t < \mathcal{D}\}}\Big]$ with barrier $\mathcal{D}$ where f is a bounded Lipschitz function with Lipschitz constant $[f]_{\text{Lip}}$. For $\alpha > 1$, by Feller's test the solution of (2.46) is known to be positive (see e.g. [53, Lemma 6.4.4.1]). Applying the Lamperti transformation, the process $(Y_t)_{0\leq t\leq T}$ given by $Y_t = X_t^{1-\alpha}$ is well defined on $I = (0, +\infty)$ and satisfies

$$\left\{ \begin{array}{rcl} dY_t &=& L(Y_t)dt + \gamma dW_t \\ Y_0 &=& X_0^{1-\alpha}, \end{array} \right. \tag{2.47}$$

where $L(y) = (1-\alpha)\left(\mu y - \alpha\frac{\sigma^2}{2}y^{-1}\right)$ and $\gamma = \sigma(1-\alpha)$. Thus, as the map $x \mapsto x^{1-\alpha}$ is decreasing, our initial pricing problem is transformed as follows on the Lamperti transform space

$$\Pi_{\mathcal{D}}^{\text{U-O},X} = \mathbb{E}\Big[g(Y_T)\mathbb{1}_{\{\inf_{t\in[0,T]} Y_t > \mathcal{D}^{1-\alpha}\}}\Big], \tag{2.48}$$

with $g : x \in \mathbb{R} \mapsto f(x^{\frac{1}{1-\alpha}})$. As $\lim\limits_{y\to 0^+} L'(y) = \lim\limits_{y\to 0^+}(1-\alpha)(\mu + \alpha\frac{\sigma^2}{2}y^{-2}) = -\infty$, we deduce that L is decreasing on $(0, \tilde{\varepsilon})$ for $\tilde{\varepsilon} > 0$ small enough and it is clearly globally Lipschitz on $[\tilde{\varepsilon}, +\infty)$ so that assumption (H3) is satisfied with $A = \tilde{\varepsilon}$ and $L'_A = (\alpha-1)\left(|\mu| + \alpha\frac{\sigma^2}{2}\varepsilon^{-2}\right)$. Also, we easily check assumptions (H1) and (H4). On the one hand, by Itô's formula the process $(Z_t)_{0\leq t\leq T}$ given by $Z_t = \frac{X_t^{-2(\alpha-1)}}{4(\alpha-1)^2}$ is a CIR process solution to

$$\left\{ \begin{array}{rcl} dZ_t &=& (a - \kappa Z_t)dt - \sigma\sqrt{Z_t}dW_t, \\ Z_0 &=& \frac{X_0^{-2(\alpha-1)}}{4(\alpha-1)^2}, \end{array} \right.$$

with $a = \frac{\sigma^2(2\alpha-1)}{4(\alpha-1)}$ and $\kappa = 2\mu(\alpha-1)$. Thanks to this second transformation we deduce that $\sup_{t\in[0,T]}\mathbb{E}[Y_t^q] < \infty$ for $q > -\frac{2\alpha-1}{2(\alpha-1)}$. On the other hand to check assumption ($\tilde{\text{H}}2$) it is enough to show that

$$\sup_{t\in[0,T]} \mathbb{E}\big[|L'(Y_t)L(Y_t)|^p + |L''(Y_t)|^p + |L'(Y_t)|^{(2\vee p)} + |L(Y_t)|^p\big] < \infty \tag{2.49}$$

which is satisfied if $\sup_{t\in[0,T]}\mathbb{E}\big[Y_t^{-(4\vee 3p)}\big] < \infty$. This condition is satisfied when $4 < \frac{2\alpha-1}{2(\alpha-1)}$ (which corresponds to have $\alpha \in (1, \frac{7}{6})$) and $p < \frac{2\alpha-1}{6(\alpha-1)}$.

Besides, since by Taylor formula we have $|g(x)-g(y)| \leq \frac{[f]_{\text{Lip}}}{\alpha-1}|x-y|(|x|^{-\frac{\alpha}{\alpha-1}} + |y|^{-\frac{\alpha}{\alpha-1}})$ for all $x, y \in I = (0, +\infty)$, then g satisfies condition (2.26) with $\nu = -\frac{\alpha}{\alpha-1}$. Hence, for $\delta \in (0, 1/2)$, if we choose α such that $1 < \alpha < \frac{59}{58} < \frac{7}{6}$ then we can find p such that $\frac{2\alpha-1}{6(\alpha-1)} > p > \frac{(1+\delta)(1+\gamma)5(1+\varepsilon)}{\frac{1}{2}-\delta} > 10$. Finally, if we choose h_ℓ sufficiently small such that $L'_A = (\alpha-1)\left(|\mu| + \alpha\frac{\sigma^2}{2}\varepsilon^{-2}\right) < \frac{1}{2h^\ell}$, then Theorem 2.3.5 is valid provided that $\inf_{t\in[0,T]} Y_t$ has a bounded density in the neighbourhood of the barrier $\mathcal{D}^{1-\alpha}$. By the monotone property of $x \in \mathbb{R}_+^* \mapsto x^{1-\alpha}$ we have the relationship $\inf_{t\in[0,T]} Y_t = (\sup_{t\in[0,T]} X_t)^{1-\alpha}$, then it is sufficient to prove that $\sup_{t\in[0,T]} X_t$ has a continuous density in the neighborhood of the barrier which is the aim of the following subsection.

Remark 2.5.1. *One can also consider the CEV process for $\alpha \in (\frac{1}{2}, 1)$ solution to*

$$dX_t = (a - \kappa X_t)dt + \sigma Y_t^{\alpha} dW_t, \text{ with } X_0 > 0, a > 0. \tag{2.50}$$

It can be easily checked that for $a > 0$ this SDE is well defined on $I = (0, +\infty)$ (see [5, Section 3]). The associated drift implicit Euler scheme is well defined on I too and satisfy the conditions of our theoretical setting. However, the condition that $\inf_{t\in[0,T]} X_t$ or $\sup_{t\in[0,T]} X_t$ admits a continuous density in the neighborhood of the barrier seems to be a challenging problem, since we dont have explicit Laplace transform of the corresponding hitting times as it is the case for the previous CEV process solution to (2.46)*. In counterpart, the efficiency of the MLMC method is still confirmed by our numerical tests for the model* (2.50)*.*

2.5.1 Running maximum of the CEV process

Let us denote by $\tau_{X_0\uparrow z} := \inf\{t \geq 0 : X_t = z\}$ the first time that the CEV process $(X_t)_{t\geq 0}$ starting at X_0 hits the level $z > X_0$. From [53, subsections 5.3.6 and 6.4.5], the Laplace transform of the hitting time $\tau_{X_0\uparrow z}$ is given by

$$\mathbf{E}[\mathrm{e}^{-s\tau_{X_0\uparrow z}}] = \Big(\frac{X_0}{z}\Big)^{\beta+\frac{1}{2}} \exp\Big(\frac{\epsilon}{2}c(X_0^{-2\beta} - z^{-2\beta})\Big)\frac{W_{k,n}(cX_0^{-2\beta})}{W_{k,n}(cz^{-2\beta})}, \tag{2.51}$$

with $\epsilon = \mathrm{sign}(\mu\beta)$, $n = \frac{1}{4\beta}$, $k = \epsilon\Big(\frac{1}{2} + \frac{1}{4\beta}\Big) - \frac{s}{2|\mu\beta|}$ and $W_{k,n}$ the Whittaker's function $W_{k,n}(y) = y^{n+\frac{1}{2}}\,\mathrm{e}^{-y/2}\,U(n - k + \frac{1}{2}, 2n + 1, y)$, where U denotes the confluent hypergeometric function of second kind defined in (2.36) and with $\beta = \alpha - 1$ and $c = \dfrac{|\mu|}{\beta\sigma^2}$.

Theorem 2.5.2. *Let $(X_t)_{0\leq t\leq T}$ denotes the CEV process solution to* (2.46)*. Then $\sup_{t\in[0,T]} X_t$ has a continuous density on any compact set $K \subset (X_0, +\infty)$, given by*

$$z \in K \mapsto P_{CEV,\ Max}(z) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{e}^{(1+iu)T}\,\hat{\Phi}(z, u)du, \tag{2.52}$$

with

$$\hat{\Phi}(z, u) = -\frac{c}{\mu}z^{-2\beta-1}\frac{U(\frac{1+iu}{2\mu\beta}, 1 + \frac{1}{2\beta}, cX_0^{-2\beta})U(\frac{1+iu}{2\mu\beta} + 1, 2 + \frac{1}{2\beta}, cz^{-2\beta})}{U(\frac{1+iu}{2\mu\beta}, 1 + \frac{1}{2\beta}, cz^{-2\beta})^2}, \text{ for } \mu > 0$$

and

$$\hat{\Phi}(z, u) = -cz^{-2\beta-1}\left(\frac{2\beta+1}{1+iu} - \frac{1}{\mu}\right)\frac{U(1 + \frac{1}{2\beta} - \frac{1+iu}{2\mu\beta}, 1 + \frac{1}{2\beta}, cX_0^{-2\beta})U(2 + \frac{1}{2\beta} - \frac{1+iu}{2\mu\beta}, 2 + \frac{1}{2\beta}, cz^{-2\beta})}{U(1 + \frac{1}{2\beta} - \frac{1+iu}{2\mu\beta}, 1 + \frac{1}{2\beta}, cz^{-2\beta})^2},$$

for $\mu < 0$.

Proof. • **Case $\mu > 0$.** By (2.51), we have

$$\mathbf{E}[\mathrm{e}^{-s\tau_{X_0\uparrow z}}] = \frac{U\Big(\frac{s}{2\mu\beta}, \frac{1}{2\beta} + 1, cX_0^{-2\beta}\Big)}{U\Big(\frac{s}{2\mu\beta}, \frac{1}{2\beta} + 1, cz^{-2\beta}\Big)}.$$

So, we deduce the law of the running maximum of the CEV process

$$\begin{aligned}\mathbb{P}\left[\sup_{0\le s\le t} X_s \ge z\right] &= \mathbb{P}\left[\tau_{X_0\uparrow z} \le t\right] \\ &= \frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty} \frac{e^{ts}}{s}\frac{U(s/2\mu\beta, 1/2\beta+1, cX_0^{-2\beta})}{U(s/2\mu\beta, 1/2\beta+1, cz^{-2\beta})}ds \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{e}^{(1+iu)t}\,\Phi(z,u)du, \end{aligned} \tag{2.53}$$

with $\Phi(z,u) = \dfrac{1}{1+iu}\dfrac{U(\frac{1+iu}{2\mu\beta}, 1+\frac{1}{2\beta}, cX_0^{-2\beta})}{U(\frac{1+iu}{2\mu\beta}, 1+\frac{1}{2\beta}, cz^{-2\beta})}$. By (2.39), we have

$$\frac{\partial\Phi}{\partial z}(z,u) = -\frac{c}{\mu}z^{-2\beta-1}\frac{U(\frac{1+iu}{2\mu\beta}, 1+\frac{1}{2\beta}, cX_0^{-2\beta})U(\frac{1+iu}{2\mu\beta}+1, 2+\frac{1}{2\beta}, cz^{-2\beta})}{U(\frac{1+iu}{2\mu\beta}, 1+\frac{1}{2\beta}, cz^{-2\beta})^2}.$$

We proceed similarly as in the proof of Theorem 2.4.2 and apply (2.40) to get for $j\in\{0,1\}$ and $v\in\{X_0, z\}$

$$\begin{aligned}&U((1+iu)/2\mu\beta + j, 1+1/2\beta+j, cv^{-2\beta}) \sim \\ &\frac{\sqrt{\pi}\,\mathrm{e}^{\frac{cv^{-2\beta}}{2}}}{\Gamma\left(\frac{1+iu}{2\mu\beta}+j\right)}\exp\left(-2\sqrt{cv^{-2\beta}(\frac{1+iu}{2\mu\beta}+j)}\right)\left(cv^{-2\beta}\right)^{-\frac{1}{4}-\frac{1}{4\beta}-\frac{j}{2}}\left(\frac{1+iu}{2\mu\beta}+j\right)^{-\frac{1}{4}+\frac{1}{4\beta}+\frac{j}{2}}\end{aligned} \tag{2.54}$$

and then use similar relation as (2.42) to get $\exp\left(-2\sqrt{cv^{-2\beta}(\frac{1+iu}{2\mu\beta}+j)}\right) \sim \exp\left(-\sqrt{\frac{c}{\mu\beta}}v^{-\beta}(1+i)\sqrt{u}\right)$, as $u\to+\infty$, uniformly in bounded v-domain. We then deduce that

$$\begin{aligned}\frac{\partial\Phi}{\partial z}(z,u) \sim -\frac{\sqrt{c}}{\mu}(X_0)^{\frac{1}{2}(\beta+1)}z^{-\frac{3}{2}(\beta+1)}\left(\frac{1+iu}{2\mu\beta}\right)^{-\frac{3}{4}-\frac{1}{4\beta}}\left(\frac{1+iu}{2\mu\beta}+1\right)^{\frac{1}{4}+\frac{1}{4\beta}} \\ \times\exp\left(\sqrt{\frac{c}{\mu\beta}}(1+i)\sqrt{u}(z^{-\beta}-X_0^{-\beta})\right)\exp\left(\frac{c}{2}(X_0^{-2\beta}-z^{-2\beta})\right)\end{aligned}$$

and then

$$\left|\frac{\partial\Phi}{\partial z}(z,u)\right| \sim \sqrt{\frac{2\beta c}{\mu}}(X_0)^{\frac{1}{2}(\beta+1)}z^{-\frac{3}{2}(\beta+1)}u^{-\frac{1}{2}}\exp\left(\sqrt{\frac{c}{\mu\beta}}\sqrt{u}(z^{-\beta}-X_0^{-\beta})\right)\exp\left(\frac{c}{2}(X_0^{-2\beta}-z^{-2\beta})\right),$$

as $u\to+\infty$ uniformly in any bounded z-domain subset of $(X_0,+\infty)$. As c,β and μ are positive constants it follows that $\int_1^{+\infty}\left|\mathrm{e}^{(1+iu)t}\frac{\partial\Phi}{\partial z}(z,u)\right|du<\infty$ is uniformly convergent in any bounded z-domain in $(X_0,+\infty)$. Besides, for the integral from $-\infty$ to -1, by (2.43), (2.44) and (2.45) we have that

$$\begin{aligned}U\Big(-(\frac{iu-1}{2\mu\beta}-j), \frac{1}{2\beta}+1+j, cv^{-2\beta}\Big) &\sim \frac{\mathrm{e}^{\frac{cv^{-2\beta}}{2}}}{\sqrt{\pi}}\left(cv^{-2\beta}(\frac{iu-1}{2\mu\beta}-j)\right)^{-\frac{1}{4}}\left(\frac{cv^{-2\beta}}{\frac{iu-1}{2\mu\beta}-j}\right)^{\frac{-j-\frac{1}{2\beta}}{2}} \\ &\times\Gamma\left(\frac{iu-1}{2\mu\beta}-j+1\right)\cos\left(\pi(\frac{iu-1}{2\mu\beta}-j)-2\sqrt{cv^{-2\beta}(\frac{iu-1}{2\mu\beta}-j)}+\frac{\pi}{2}(\frac{1}{2\beta}+j)+\frac{\pi}{4}\right) \\ &\sim \frac{\mathrm{e}^{\frac{cv^{-2\beta}}{2}}}{2\sqrt{\pi}}e^{-i\pi\left(\frac{1}{4\beta}-\frac{1}{2\mu\beta}-\frac{j}{2}+\frac{1}{4}\right)}e^{\frac{\pi}{2\mu\beta}u-(1-i)\sqrt{\frac{cv^{-2\beta}}{\mu\beta}u}} \\ &\times\left(cv^{-2\beta}(\frac{iu-1}{2\mu\beta}-j)\right)^{-\frac{1}{4}}\Gamma\left(\frac{iu-1}{2\mu\beta}-j+1\right)\left(\frac{cv^{-2\beta}}{\frac{iu-1}{2\mu\beta}-j}\right)^{\frac{-j-\frac{1}{2\beta}}{2}}\end{aligned} \tag{2.55}$$

as $u \to +\infty$ uniformly in bounded v-domain. Thus,

$$|\frac{\partial \Phi}{\partial z}(z,-u)| \sim \sqrt{\frac{2\beta c}{\mu}}(X_0)^{\frac{1}{2}(\beta+1)} z^{-\frac{3}{2}(\beta+1)} u^{-\frac{1}{2}} \exp\left(\sqrt{\frac{c}{\mu\beta}}\sqrt{u}(z^{-\beta}-X_0^{-\beta})\right) \exp\left(\frac{c}{2}(X_0^{-2\beta}-z^{-2\beta})\right)$$

as $u \to +\infty$ uniformly in any bounded z-domain subset of $(X_0,+\infty)$ that $\int_1^{+\infty} \left|\mathrm{e}^{(1+iu)t}\,\frac{\partial \Phi}{\partial z}(z,u)\right| du < \infty$ is uniformly convergent in any bounded z-domain in $(X_0,+\infty)$. We deduce in the same way as above that $\int_{-\infty}^{-1} \left|\mathrm{e}^{(1+iu)t}\,\frac{\partial \Phi}{\partial z}(z,u)\right| du < \infty$ is uniformly convergent in any bounded z-domain in $(X_0,+\infty)$. We complete the proof by noticing that $(u,z) \in \mathbb{R}\times K \mapsto \mathrm{e}^{(1+iu)t}\,\frac{\partial \Phi}{\partial z}(u,z)$ is a continuous function for any compact set $K \subset (X_0,+\infty)$.

• **Case $\mu < 0$.** For this case, by (2.51) we have

$$\mathbf{E}[\mathrm{e}^{-s\tau_{X_0\uparrow z}}] = \exp\Big(-c(X_0^{-2\beta}-y^{-2\beta})\Big)\frac{U\Big(1+\frac{1}{2\beta}-\frac{s}{2\mu\beta},\frac{1}{2\beta}+1,cX_0^{-2\beta}\Big)}{U\Big(1+\frac{1}{2\beta}-\frac{s}{2\mu\beta},\frac{1}{2\beta}+1,cz^{-2\beta}\Big)}$$

and

$$\mathbb{P}\big[\sup_{0\le s\le t} X_s > z\big] = \frac{\exp\Big(-c(X_0^{-2\beta}-y^{-2\beta})\Big)}{2\pi}\int_{-\infty}^{+\infty} \mathrm{e}^{(1+iu)t}\Phi(z,u)du,$$

with

$$\Phi(z,u) = \frac{1}{1+iu}\frac{U\Big(1+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta},\frac{1}{2\beta}+1,cX_0^{-2\beta}\Big)}{U\Big(1+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta},\frac{1}{2\beta}+1,cz^{-2\beta}\Big)}.$$

By (2.39), we obtain

$$\frac{\partial \Phi}{\partial z}(z,u) = -cz^{-2\beta-1}\left(\frac{2\beta+1}{1+iu}-\frac{1}{\mu}\right)\frac{U(1+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta},1+\frac{1}{2\beta},cX_0^{-2\beta})U(2+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta},2+\frac{1}{2\beta},cz^{-2\beta})}{U(1+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta},1+\frac{1}{2\beta},cz^{-2\beta})^2}.$$

Proceeding in the same way as for the previous case, we first apply (2.40) to get for $j \in \{0,1\}$ and $v \in \{X_0, z\}$

$$U(1+j+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta},1+j+\frac{1}{2\beta},cv^{-2\beta}) \sim \sqrt{\pi}\frac{\Big(1+j+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta}\Big)^{-\frac{1}{4}+\frac{j}{2}+\frac{1}{4\beta}}}{\Gamma\Big(1+j+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta}\Big)}\Big(cv^{-2\beta}\Big)^{-\frac{1}{4}-\frac{j}{2}-\frac{1}{4\beta}}$$
$$\times \exp\left(\frac{cv^{-2\beta}}{2}-2\sqrt{cv^{-2\beta}\left(1+j+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta}\right)}\right)$$

and then we use that $\exp\left(-2\sqrt{cv^{-2\beta}\left(1+j+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta}\right)}\right) \sim \exp\left(-\sqrt{-\frac{cv^{-2\beta}}{\mu\beta}}(1+i)\sqrt{u}\right)$ as $u \to +\infty$, uniformly in bounded v-domain. We then deduce that

$$\frac{\partial \Phi}{\partial z}(z,u) \sim -cz^{-2\beta-1}\left(\frac{2\beta+1}{1+iu}-\frac{1}{\mu}\right)\Big(1+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta}\Big)^{-\frac{3}{4}-\frac{1}{4\beta}}\Big(2+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta}\Big)^{\frac{1}{4}+\frac{1}{4\beta}}$$
$$\times\Big(cX_0^{-2\beta}\Big)^{-\frac{1}{4}-\frac{1}{4\beta}}\Big(cz^{-2\beta}\Big)^{-\frac{1}{4}+\frac{1}{4\beta}}\exp\left(\frac{c}{2}\Big(X_0^{-2\beta}-z^{-2\beta}\Big)\right)\exp\left(\sqrt{-\frac{c}{\mu\beta}}(1+i)\sqrt{u}(z^{-\beta}-X_0^{-\beta})\right)$$

and

$$\left|\frac{\partial\Phi}{\partial z}(z,u)\right| \sim \sqrt{-\frac{2\beta c}{\mu}}(X_0)^{\frac{1}{2}(\beta+1)} z^{-\frac{3}{2}(\beta+1)} u^{-\frac{1}{2}} \exp\left(\sqrt{-\frac{c}{\mu\beta}}\sqrt{u}(z^{-\beta}-X_0^{-\beta})\right)\exp\left(\frac{c}{2}(X_0^{-2\beta}-z^{-2\beta})\right)$$

as $u \to +\infty$ uniformly in any bounded z-domain subset of $(X_0, +\infty)$. As c, β and $-\mu$ are positive constants it follows that $\int_1^{+\infty}\left|\mathrm{e}^{(1+iu)t}\,\frac{\partial\Phi}{\partial z}(z,u)\right| du < \infty$ is uniformly convergent in any bounded z-domain in $(X_0, +\infty)$. Concerning the integral from $-\infty$ to -1, we proceed as above and use (2.43), (2.44) and (2.45) to get for $j \in \{0,1\}$ and $v \in \{X_0, z\}$

$$\begin{aligned}
&U(-(-1-j-\frac{1}{2\beta}+\frac{1-iu}{2\mu\beta}), 1+j+\frac{1}{2\beta}, cv^{-2\beta}) \sim \frac{\mathrm{e}^{\frac{cv^{-2\beta}}{2}}}{\sqrt{\pi}}\left(\frac{cv^{-2\beta}}{-1-j-\frac{1}{2\beta}+\frac{1-iu}{2\mu\beta}}\right)^{-\frac{j}{2}-\frac{1}{4\beta}}\\
&\times\left(cv^{-2\beta}\left(-1-j-\frac{1}{2\beta}+\frac{1-iu}{2\mu\beta}\right)\right)^{-\frac{1}{4}}\\
&\times\cos\left(\pi\left(-1-j-\frac{1}{2\beta}+\frac{1-iu}{2\mu\beta}\right)-2\sqrt{cv^{-2\beta}\left(-1-j-\frac{1}{2\beta}+\frac{1-iu}{2\mu\beta}\right)}+\frac{\pi}{2}(j+\frac{1}{2\beta})+\frac{\pi}{4}\right)
\end{aligned}$$

as $u \to +\infty$ uniformly in bounded v-domain. Using that

$$\begin{aligned}
&\cos\left(\pi\left(-1-j-\frac{1}{2\beta}+\frac{1-iu}{2\mu\beta}\right)-2\sqrt{cv^{-2\beta}\left(-1-j-\frac{1}{2\beta}+\frac{1-iu}{2\mu\beta}\right)}+\frac{\pi}{2}(j+\frac{1}{2\beta})+\frac{\pi}{4}\right)\\
&\qquad\sim \frac{1}{2}e^{-i\pi\left(-\frac{3}{4}-\frac{j}{2}-\frac{1}{4\beta}+\frac{1}{2\mu\beta}\right)}e^{-\frac{\pi}{2\mu\beta}u-(1-i)\sqrt{-\frac{cv^{-2\beta}}{\mu\beta}u}},
\end{aligned}$$

as $u \to +\infty$ uniformly in bounded v-domain. Thus, we get

$$\left|\frac{\partial\Phi}{\partial z}(z,u)\right| \sim \sqrt{-\frac{2\beta c}{\mu}}(X_0)^{\frac{1}{2}(\beta+1)} z^{-\frac{3}{2}(\beta+1)} u^{-\frac{1}{2}} \exp\left(\sqrt{-\frac{c}{\mu\beta}}\sqrt{u}(z^{-\beta}-X_0^{-\beta})\right)\exp\left(\frac{c}{2}(X_0^{-2\beta}-z^{-2\beta})\right)$$

as $u \to +\infty$ uniformly in any bounded z-domain subset of $(X_0, +\infty)$. As c, β and $-\mu$ are positive constants it follows that $\int_{-\infty}^{-1}\left|\mathrm{e}^{(1+iu)t}\,\frac{\partial\Phi}{\partial z}(z,u)\right| du < \infty$ is uniformly convergent in any bounded z-domain in $(X_0, +\infty)$. We complete the proof by noticing that $(u,z) \in \mathbb{R}\times K \mapsto \mathrm{e}^{(1+iu)t}\,\frac{\partial\Phi}{\partial z}(u,z)$ is a continuous function for any compact set $K \subset (X_0, +\infty)$ (see e.g. [70, Theorem B.3]). □

2.5.2 Minimum of CEV process

Let us denote by $\tau_{X_0\downarrow z} := \inf\{t \geq 0 : X_t = z\}$ the first time that the CEV process $(X_t)_{t\geq 0}$ starting at X_0 hits the level $0 < z < X_0$. By [53, subsections 5.3.6 and 6.4.5], the Laplace transform of the hitting time $\tau_{X_0\downarrow z} := \inf\{t \geq 0 : X_t = z\}$ is given by

$$\mathbf{E}[\mathrm{e}^{-s\tau_{X_0\downarrow z}}] = \left(\frac{X_0}{z}\right)^{\beta+\frac{1}{2}}\exp\left(\frac{\epsilon}{2}c(X_0^{-2\beta}-z^{-2\beta})\right)\frac{M_{k,n}(cX_0^{-2\beta})}{M_{k,n}(cz^{-2\beta})} \tag{2.56}$$

with $\epsilon = \mathrm{sign}(\mu\beta)$, $n = \frac{1}{4\beta}$, $k = \epsilon\left(\frac{1}{2}+\frac{1}{4\beta}\right) - \frac{s}{2\beta|\mu|}$ and the Whittaker function $M_{k,n}(y) = y^{n+\frac{1}{2}}\,\mathrm{e}^{-\frac{y}{2}}\,{}_1F_1(n-k+\frac{1}{2}, 2n+1, y)$, where ${}_1F_1$ denotes the confluent hypergeometric function of the first kind defined in (2.31) and with $\beta = \alpha - 1$ and $c = \dfrac{|\mu|}{\beta\sigma^2}$.

Theorem 2.5.3. *Let $(X_t)_{0\le t\le T}$ denotes the CEV process solution to* (2.46)*. Then* $\inf_{t\in[0,T]} X_t$ *has a continuous density on any compact set $K\subset(0,X_0)$, given by*

$$z\in K\mapsto P_{CEV,\ Min}(z)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\mathrm{e}^{(1+iu)T}\,\hat{\Psi}(z,u)du, \tag{2.57}$$

with

$$\hat{\Psi}(z,u)=-\frac{cz^{-2\beta-1}}{\mu(1+\frac{1}{2\beta})}\frac{{}_1F_1(\frac{1+iu}{2\mu\beta},1+\frac{1}{2\beta},cX_0^{-2\beta})\,{}_1F_1(\frac{1+iu}{2\mu\beta}+1,2+\frac{1}{2\beta},cz^{-2\beta})}{{}_1F_1(\frac{1+iu}{2\mu\beta},1+\frac{1}{2\beta},cz^{-2\beta})^2},\ \textit{for } \mu>0$$

and

$$\begin{aligned}\hat{\Psi}(z,u)=&-cz^{-2\beta-1}\left(\frac{2\beta}{1+iu}-\frac{1}{\mu(1+\frac{1}{2\beta})}\right)\\&\times\frac{{}_1F_1(1+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta},1+\frac{1}{2\beta},cX_0^{-2\beta})\,{}_1F_1(2+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta},2+\frac{1}{2\beta},cz^{-2\beta})}{{}_1F_1(1+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta},1+\frac{1}{2\beta},cz^{-2\beta})^2},\ \textit{for } \mu<0.\end{aligned}$$

Proof. • **Case $\mu>0$.** In this case, by (2.56) we have $\mathbf{E}[\mathrm{e}^{-s\tau_{X_0\downarrow z}}]=\dfrac{{}_1F_1\Big(\frac{s}{2\mu\beta},\frac{1}{2\beta}+1,cX_0^{-2\beta}\Big)}{{}_1F_1\Big(\frac{s}{2\mu\beta},\frac{1}{2\beta}+1,cz^{-2\beta}\Big)}$ and the law of the running minimum of the CEV process is given by

$$\mathbb{P}\left[\inf_{0\le s\le t}X_s\ge z\right]=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\mathrm{e}^{(1+iu)t}\Psi(z,u)du \tag{2.58}$$

with $\Psi(z,u)=\dfrac{1}{1+iu}\dfrac{{}_1F_1(\frac{1+iu}{2\mu\beta},1+\frac{1}{2\beta},cX_0^{-2\beta})}{{}_1F_1(\frac{1+iu}{2\mu\beta},1+\frac{1}{2\beta},cz^{-2\beta})}$. By (2.33), we have

$$\frac{\partial\Psi}{\partial z}(z,u)=-\frac{cz^{-2\beta-1}}{\mu(1+\frac{1}{2\beta})}\frac{{}_1F_1(\frac{1+iu}{2\mu\beta},1+\frac{1}{2\beta},cX_0^{-2\beta})\,{}_1F_1(\frac{1+iu}{2\mu\beta}+1,2+\frac{1}{2\beta},cz^{-2\beta})}{{}_1F_1(\frac{1+iu}{2\mu\beta},1+\frac{1}{2\beta},cz^{-2\beta})^2}.$$

We proceed in the same way as we used to obtain (2.34) to get for $j\in\{0,1\}$ and $v\in\{X_0,z\}$

$$\begin{aligned}{}_1F_1(\frac{1+iu}{2\mu\beta}+j,1+j+\frac{1}{2\beta},cv^{-2\beta})\sim&\frac{\mathrm{e}^{\frac{1}{2}cv^{-2\beta}}}{\sqrt{4\pi}}\left(\frac{cv^{-2\beta}}{\frac{1+iu}{2\mu\beta}+j}\right)^{-\frac{j}{2}-\frac{1}{4\beta}}\frac{\Gamma(1+j+\frac{1}{2\beta})\Gamma(\frac{1+iu}{2\mu\beta}-\frac{1}{2\beta})}{\Gamma(\frac{1+iu}{2\mu\beta}+j)}\\&\left(cv^{-2\beta}\left(\frac{1+iu}{2\mu\beta}+j\right)\right)^{-\frac{1}{4}}\exp\left(2\sqrt{cv^{-2\beta}\left(\frac{1+iu}{2\mu\beta}+j\right)}\right).\end{aligned}$$

Using that

$$\Gamma(x+a)/\Gamma(x+b)\sim x^{a-b},\ \text{ as } x\to\infty \tag{2.59}$$

uniformly inside the sector $|\arg(x)|<\pi$, (see e.g. (6.5.72) of [78]), and that $\exp\left(2\sqrt{cv^{-2\beta}(\frac{1+iu}{2\mu\beta}+j)}\right)\sim\exp\left(\sqrt{\frac{cv^{-2\beta}}{\mu\beta}}(1+i)\sqrt{u}\right)$, as $u\to+\infty$ uniformly in bounded v-domain, we get

$$\begin{aligned}{}_1F_1(\frac{1+iu}{2\mu\beta}+j,1+j+\frac{1}{2\beta},cv^{-2\beta})\sim&\frac{\mathrm{e}^{\frac{1}{2}cv^{-2\beta}}}{\sqrt{4\pi}}\left(cv^{-2\beta}\left(\frac{1+iu}{2\mu\beta}+j\right)\right)^{-\frac{j}{2}-\frac{1}{4\beta}-\frac{1}{4}}\Gamma(1+j+\frac{1}{2\beta})\\&\exp\left(\sqrt{\frac{cv^{-2\beta}}{\mu\beta}}(1+i)\sqrt{u}\right)\end{aligned}$$

uniformly in bounded v-domain. Hence, we deduce that

$$|\frac{\partial\Psi}{\partial z}(z,u)| \sim \sqrt{\frac{2\beta c}{\mu}}(X_0)^{\frac{1}{2}(\beta+1)}z^{-\frac{3}{2}(\beta+1)}u^{-\frac{1}{2}}\exp\left(\sqrt{\frac{c}{\mu\beta}}\sqrt{u}(X_0^{-\beta}-z^{-\beta})\right)\exp\left(\frac{c}{2}(X_0^{-2\beta}-z^{-2\beta})\right),$$

as $u\to+\infty$ uniformly in any bounded z-domain subset of $(0,X_0)$. As c,β and μ are positive constants it follows that $\int_1^{+\infty}\left|\mathrm{e}^{(1+iu)t}\,\frac{\partial K}{\partial z}(z,u)\right|du<\infty$ is uniformly convergent in any bounded z-domain in $(0,X_0)$. Besides, for the integral from $-\infty$ to -1, we use similar steps as the ones that led us to get (2.35), and we have

$$\begin{aligned}{}_1F_1(\frac{1-iu}{2\mu\beta}+j,1+j+\frac{1}{2\beta},cv^{-2\beta}) \sim \frac{\mathrm{e}^{\frac{1}{2}cv^{-2\beta}}}{\sqrt{\pi}}&\left(\frac{cv^{-2\beta}}{\frac{iu-1}{2\mu\beta}-j}\right)^{-\frac{j}{2}-\frac{1}{4\beta}}\frac{\Gamma(1+j+\frac{1}{2\beta})\Gamma(\frac{iu-1}{2\mu\beta}-j)}{\Gamma(\frac{iu-1}{2\mu\beta}+1+\frac{1}{2\beta})}\\&\left(cv^{-2\beta}\left(\frac{iu-1}{2\mu\beta}-j\right)\right)^{-\frac{1}{4}}\\&\times\cos\left(2\sqrt{cv^{-2\beta}\left(\frac{iu-1}{2\mu\beta}-j\right)}-\frac{\pi}{2}(j+\frac{1}{2\beta})-\frac{\pi}{4}\right)\end{aligned}$$

as $u\to\infty$ uniformly on v-bounded domain. Now, using (2.59) and that

$$\cos\left(2\sqrt{cv^{-2\beta}\left(\frac{iu-1}{2\mu\beta}-j\right)}-\frac{\pi}{2}(j+\frac{1}{2\beta})-\frac{\pi}{4}\right)\sim\frac{1}{2}e^{i\pi\left(\frac{j}{2}+\frac{1}{4\beta}+\frac{1}{4}\right)}e^{(1-i)\sqrt{\frac{cv^{-2\beta}}{\mu\beta}u}},$$

as $u\to\infty$ uniformly on v-bounded domain, we get

$$\begin{aligned}{}_1F_1(\frac{1-iu}{2\mu\beta}+j,1+j+\frac{1}{2\beta},cv^{-2\beta}) \sim \frac{\mathrm{e}^{\frac{1}{2}cv^{-2\beta}}}{2\sqrt{\pi}}&\left(cv^{-2\beta}\right)^{-\frac{j}{2}-\frac{1}{4\beta}-\frac{1}{4}}\Gamma(1+j+\frac{1}{2\beta})\\&\left(\frac{iu-1}{2\mu\beta}-j\right)^{-\frac{j}{2}-\frac{1}{4\beta}-\frac{5}{4}}\mathrm{e}^{i\pi\left(\frac{j}{2}+\frac{1}{4\beta}+\frac{1}{4}\right)}\exp\left((1-i)\sqrt{\frac{cv^{-2\beta}}{\mu\beta}u}\right)\end{aligned}$$

and that

$$|\frac{\partial\Psi}{\partial z}(z,-u)| \sim \sqrt{\frac{2\beta c}{\mu}}(X_0)^{\frac{1}{2}(\beta+1)}z^{-\frac{3}{2}(\beta+1)}u^{-\frac{1}{2}}\exp\left(\sqrt{\frac{c}{\mu\beta}}\sqrt{u}(X_0^{-\beta}-z^{-\beta})\right)\exp\left(\frac{c}{2}(X_0^{-2\beta}-z^{-2\beta})\right),$$

as $u\to+\infty$ uniformly in any bounded z-domain subset of $(0,X_0)$. Thus, as c,β and μ are positive constants we deduce that $\int_{-\infty}^{-1}\left|\mathrm{e}^{(1+iu)t}\,\frac{\partial\Psi}{\partial z}(z,u)\right|du<\infty$ is uniformly convergent in any bounded z-domain in $(0,X_0)$. We complete the proof of the case $\mu>0$ by noticing that $(u,z)\in\mathbb{R}\times K\mapsto\mathrm{e}^{(1+iu)t}\,\frac{\partial\Psi}{\partial z}(u,z)$ is a continuous function for any compact set $K\subset(0,X_0)$.

• **Case $\mu<0$.** In this case by (2.56) we have,

$$\mathbf{E}[\mathrm{e}^{-s\tau_{X_0\downarrow z}}]=\exp\Big(-c(X_0^{-2\beta}-z^{-2\beta})\Big)\frac{{}_1F_1\Big(\frac{1}{2\beta}+1-\frac{s}{2\mu\beta},\frac{1}{2\beta}+1,cX_0^{-2\beta}\Big)}{{}_1F_1\Big(\frac{1}{2\beta}+1-\frac{s}{2\mu\beta},\frac{1}{2\beta}+1,cz^{-2\beta}\Big)}$$

and

$$\mathbb{P}\left[\inf_{0\le s\le t}X_s\ge z\right]=\frac{\exp\Big(-c(X_0^{-2\beta}-z^{-2\beta})\Big)}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{(1+iu)t}\Psi(z,u)du$$

with $\Psi(z,u)=\dfrac{1}{1+iu}\dfrac{{}_1F_1\Big(\frac{1}{2\beta}+1-\frac{1+iu}{2\mu\beta},\frac{1}{2\beta}+1,cX_0^{-2\beta}\Big)}{{}_1F_1\Big(\frac{1}{2\beta}+1-\frac{1+iu}{2\mu\beta},\frac{1}{2\beta}+1,cz^{-2\beta}\Big)}$. By (2.33), we have

$$\frac{\partial\Psi}{\partial z}(z,u)=-cz^{-2\beta-1}\left(\frac{2\beta}{1+iu}-\frac{1}{\mu(1+\frac{1}{2\beta})}\right) \tag{2.60}$$

$$\times\frac{{}_1F_1(1+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta},1+\frac{1}{2\beta},cX_0^{-2\beta}){}_1F_1(2+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta},2+\frac{1}{2\beta},cz^{-2\beta})}{{}_1F_1(1+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta},1+\frac{1}{2\beta},cz^{-2\beta})^2}. \tag{2.61}$$

For $j\in\{0,1\}$ and $v\in\{X_0,z\}$, we follow the similar steps that led us to get (2.34) and we obtain

$$\begin{aligned}{}_1F_1(1+j+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta},1+j+\frac{1}{2\beta},cv^{-2\beta})&\sim\frac{\mathrm{e}^{\frac{cv^{-2\beta}}{2}}}{\sqrt{4\pi}}\left(\frac{cv^{-2\beta}}{1+j+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta}}\right)^{-\frac{j}{2}-\frac{1}{4\beta}}\\&\times\frac{\Gamma(1+j+\frac{1}{2\beta})\Gamma(1-\frac{1+iu}{2\mu\beta})}{\Gamma(1+j+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta})}\left(cv^{-2\beta}\left(1+j+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta}\right)\right)^{-\frac{1}{4}}\\&\times\exp\left(2\sqrt{cv^{-2\beta}\left(1+j+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta}\right)}\right)\end{aligned}$$

as $u\to+\infty$ uniformly in bounded v-domain. Now, we use (2.59) and that

$$\exp\left(2\sqrt{cv^{-2\beta}\left(1+j+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta}\right)}\right)\sim\exp\left(\sqrt{-\frac{cv^{-2\beta}}{\mu\beta}}(1+i)\sqrt{u}\right),\ \text{as } u\to+\infty$$

uniformly in bounded v-domain, to get

$$\begin{aligned}{}_1F_1(1+j+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta},1+j+\frac{1}{2\beta},cv^{-2\beta})&\sim\frac{\mathrm{e}^{\frac{cv^{-2\beta}}{2}}}{\sqrt{4\pi}}\left(cv^{-2\beta}\left(1+j+\frac{1}{2\beta}-\frac{1+iu}{2\mu\beta}\right)\right)^{-\frac{j}{2}-\frac{1}{4\beta}-\frac{1}{4}}\\&\times\Gamma(1+j+\frac{1}{2\beta})\exp\left(\sqrt{-\frac{cv^{-2\beta}}{\mu\beta}}(1+i)\sqrt{u}\right)\end{aligned}$$

as $u\to+\infty$ uniformly in bounded v-domain. Hence, by (2.60) we deduce that

$$\left|\frac{\partial\Psi}{\partial z}(z,u)\right|\sim\sqrt{-\frac{2\beta c}{\mu}}(X_0)^{\frac{1}{2}(\beta+1)}z^{-\frac{3}{2}(\beta+1)}u^{-\frac{1}{2}}\exp\left(\sqrt{-\frac{c}{\mu\beta}}\sqrt{u}(X_0^{-\beta}-z^{-\beta})\right)\exp\left(\frac{c}{2}(X_0^{-2\beta}-z^{-2\beta})\right),$$

as $u\to+\infty$ uniformly in any bounded z-domain subset of $(0,X_0)$. As c,β and $-\mu$ are positive constants it follows that $\int_1^{+\infty}\left|\mathrm{e}^{(1+iu)t}\,\frac{\partial K}{\partial z}(z,u)\right|du<\infty$ is uniformly convergent in any bounded z-domain in $(0,X_0)$. Besides, for the integral from $-\infty$ to -1, we use similar steps as the ones that led us to get (2.35), and we have

$$\begin{aligned}{}_1F_1(1+j+\frac{1}{2\beta}-\frac{1-iu}{2\mu\beta},1+j+\frac{1}{2\beta},cv^{-2\beta})&\sim\frac{\mathrm{e}^{\frac{1}{2}cv^{-2\beta}}}{\sqrt{\pi}}\left(\frac{cv^{-2\beta}}{-1-j-\frac{1}{2\beta}-\frac{iu-1}{2\mu\beta}}\right)^{-\frac{j}{2}-\frac{1}{4\beta}}\\&\times\frac{\Gamma(1+j+\frac{1}{2\beta})\Gamma(-j-\frac{1}{2\beta}-\frac{iu-1}{2\mu\beta})}{\Gamma(-\frac{iu-1}{2\mu\beta})}\\&\left(cv^{-2\beta}\left(-1-j-\frac{1}{2\beta}-\frac{iu-1}{2\mu\beta}\right)\right)^{-\frac{1}{4}}\\&\times\cos\left(2\sqrt{cv^{-2\beta}\left(-1-j-\frac{1}{2\beta}-\frac{iu-1}{2\mu\beta}\right)}-\frac{\pi}{2}(j+\frac{1}{2\beta})-\frac{\pi}{4}\right)\end{aligned}$$

as $u \to \infty$ uniformly on v-bounded domain. Now, using (2.59) and that

$$\cos\left(2\sqrt{cv^{-2\beta}\left(-1-j-\frac{1}{2\beta}-\frac{iu-1}{2\mu\beta}\right)}-\frac{\pi}{2}(j+\frac{1}{2\beta})-\frac{\pi}{4}\right) \sim \frac{1}{2}e^{i\pi\left(\frac{j}{2}+\frac{1}{4\beta}+\frac{1}{4}\right)}e^{(1-i)\sqrt{-\frac{cv^{-2\beta}}{\mu\beta}u}},$$

as $u \to \infty$ uniformly on v-bounded domain, we get

$$\begin{aligned}{}_1F_1(1+j+\frac{1}{2\beta}-\frac{1-iu}{2\mu\beta},1+j+\frac{1}{2\beta},cv^{-2\beta}) \sim{}& \frac{\mathrm{e}^{\frac{1}{2}cv^{-2\beta}}}{2\sqrt{\pi}}\left(cv^{-2\beta}\right)^{-\frac{j}{2}-\frac{1}{4\beta}-\frac{1}{4}}\Gamma(1+j+\frac{1}{2\beta})\\ &\left(-1-j-\frac{1}{2\beta}-\frac{iu-1}{2\mu\beta}\right)^{-\frac{j}{2}-\frac{1}{4\beta}-\frac{1}{4}}\mathrm{e}^{i\pi\left(\frac{j}{2}+\frac{1}{4\beta}+\frac{1}{4}\right)}\\ &\times\exp\left((1-i)\sqrt{-\frac{cv^{-2\beta}}{\mu\beta}u}\right)\end{aligned}$$

and that

$$|\frac{\partial\Psi}{\partial z}(z,-u)| \sim \sqrt{-\frac{2\beta c}{\mu}}(X_0)^{\frac{1}{2}(\beta+1)}z^{-\frac{3}{2}(\beta+1)}u^{-\frac{1}{2}}\exp\left(\sqrt{-\frac{c}{\mu\beta}}\sqrt{u}(X_0^{-\beta}-z^{-\beta})\right)\exp\left(\frac{c}{2}(X_0^{-2\beta}-z^{-2\beta})\right),$$

as $u \to +\infty$ uniformly in any bounded z-domain subset of $(0, X_0)$. Thus, as c, β and $-\mu$ are positive constants we deduce that $\int_{-\infty}^{-1}\left|\mathrm{e}^{(1+iu)t}\,\frac{\partial\Psi}{\partial z}(z,u)\right|du < \infty$ is uniformly convergent in any bounded z-domain in $(0, X_0)$. We complete the proof by noticing that $(u, z) \in \mathbb{R} \times K \mapsto \mathrm{e}^{(1+iu)t}\,\frac{\partial\Psi}{\partial z}(u, z)$ is a continuous function for any compact set $K \subset (0, X_0)$ (see e.g. [70, Theorem B.3]). □

2.5.3 Numerical tests

For the CEV case, we consider the pricing problem of the quantities introduced in (2.48). More precisely, we approximate $\Pi_{\mathcal{D}}^{\text{U-O},X}$ (resp. $\Pi_{\mathcal{U}}^{\text{D-O},X}$) by the improved MLMC algorithm $\bar{Q}_{\mathcal{D}^{1-\alpha}}$ given in (3.13) (resp. $\bar{P}_{\mathcal{U}^{1-\alpha}}$ given in (3.12)), where we used our interpolated drift implicit scheme

$$\begin{aligned}&\overline{Y}_t^n = \overline{Y}_{t_i}^n + (1-\alpha)\left(\mu\overline{Y}_{t_{i+1}}^n - \alpha\frac{\sigma^2}{2\overline{Y}_{t_{i+1}}^n}\right)(t-t_i) + \gamma(W_t - W_{t_i}), \qquad \text{for } t \in [t_i, t_{i+1}[, 0 \le i \le n-1,\\ &Y_0 = {X_0}^{1-\alpha},\end{aligned}$$

with $\gamma = \sigma(1-\alpha)$. For n large enough, the positive solution to the above implicit scheme is explicit and given by

$$\overline{Y}_{t_{i+1}}^n = \frac{\sqrt{2\sigma^2\alpha(\alpha-1)(1+\mu(\alpha-1)\frac{T}{n})\frac{T}{n} + (\gamma(W_{t_{i+1}}-W_{t_i})+\overline{Y}_{t_i}^n)^2} + \gamma(W_{t_{i+1}}-W_{t_i}) + \overline{Y}_{t_i}^n}{2+2\mu(\alpha-1)\frac{T}{n}}.$$

For the CEV model, we dont have any benchmark price. To illustrate the MLMC complexity performance we choose a set of parameter (2.49), namely $\alpha = 1.2$, $X_0 = 100$, $\mu = 0.1$, $\sigma = 0.2$, $T = 1$. The payoff function $g(x) = \mathrm{e}^{-rT}(x^{\frac{1}{1-\alpha}} - K)_+$ is a discounted call function with $r = 0.1$. For the U-O option the strike is $K = 90$, and the barrier $\mathcal{D} = 150$. For the D-O option the strike is $K = 100$ and the barrier $\mathcal{U} = 90$. The tables and the figures below confirm the high performance of the improved MLMC.

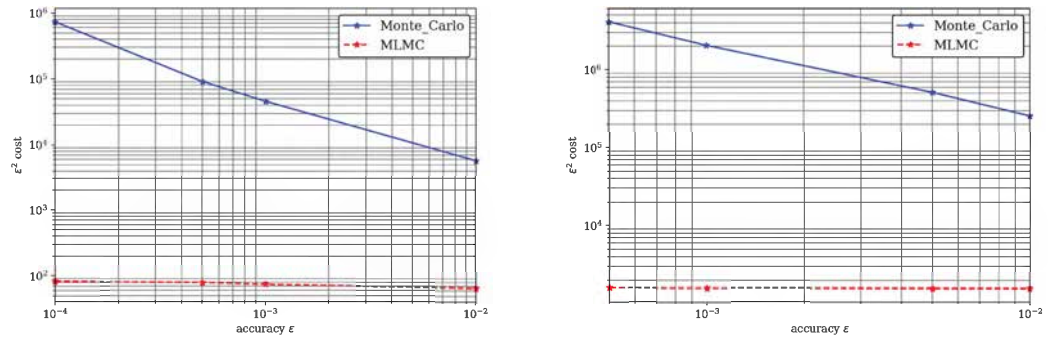
For pricing barrier options under the popular CEV model, the numerical results confirm the supremacy of the improved MLMC algorithm that reaches the optimal time complexity $O(\varepsilon^{-2})$ for a given precision $\varepsilon > 0$.

Accuracy	Price	MLMC cost	MC cost	Saving
10^{-4}	3.0390	8.226×10^{9}	7.34×10^{13}	8922.33
5×10^{-4}	3.0391	3.17×10^{8}	3.67×10^{11}	1155.67
10^{-3}	3.041	7.436×10^{7}	4.587×10^{10}	616.91
10^{-2}	3.0452	6.539×10^{5}	5.734×10^{7}	87.69

Table 2.3: MLMC complexity tests for the U-O barrier option pricing of $\Pi_{\mathcal{D}}^{\text{U-O},X}$

Accuracy	Price	MLMC cost	MC cost	Saving
5×10^{-4}	11.102	6.483×10^{9}	1.642×10^{13}	2532.83
10^{-3}	11.103	1.608×10^{9}	2.053×10^{12}	1276.66
5×10^{-3}	11.106	6.379×10^{7}	2.053×10^{10}	321.77
10^{-2}	11.094	1.587×10^{7}	2.566×10^{9}	161.69

Table 2.4: MLMC complexity tests for the D-O barrier option pricing of $\Pi_{\mathcal{U}}^{\text{D-O},X}$



(a) Approximation of $\Pi_{\mathcal{D}}^{\text{U-O},X}$ **(b)** Approximation of $\Pi_{\mathcal{U}}^{\text{D-O},X}$

Fig 2.2. Comparison for the performances of MLMC vs classical MC algorithm under the CEV model

2.6 Conclusion

In this paper, we proved that the MLMC method for pricing barrier options reaches its optimal time-complexity regime, when the underlying asset has non-Lipschitz coefficients. To apply our theoretical results for the popular CIR and CEV processes, we developed semi-explicit formulas for the densities of the running minimum and running maximum of these processes that are of independent interest. It turns out that under some constraints on the parameters of these models guaranteeing the existence of finite negative moments up to some order, the MLMC method behaves like an unbiased classic Monte Carlo estimator despite the use of approximation schemes. It may be interesting to extend this study by combining this improved version of the MLMC method with importance sampling technics for variance reduction as proposed in [10, 57, 11], we leave this for a possible future work.

Chapter 3

Multi Level Monte Carlo: Application to Barrier Options under Heston Model

Abstract

In this paper, we study the convergence of the variance of the MLMC estimator for the pricing of barrier options in the Heston model. Such results were only studied for one-dimensional models with globally Lipschitz coefficients (see Giles et al. [37]). We introduce a Semi-Exact Euler scheme for the log-Heston model, whose variance process is exactly simulated. We analyse the variance of the associated MLMC method for pricing Barrier option variance. We devise a numerical approximation for which we are able to control the associated MLMC variance. To the best of our knowledge, there is no existing research in the literature examining the validity of extreme path arguments when analyzing the MLMC method for pricing barrier options under a d-dimensional model with $d > 1$. We proved theoretically that the variance of the obtained MLMC method is $O(h_\ell^{\frac{1}{2}-\delta})$ for all $\delta \in (0, \frac{1}{2})$. Several numerical tests were carried out.

3.1 Introduction

The Heston model is a stochastic volatility model introduced by Heston [49]. It is characterized by its volatility which is a stochastic process following a Cox–Ingersoll–Ross (CIR) model. Closed formula to evaluate exotic option prices under the Heston model are rarely available . Therefore, Monte Carlo methods are typically used to estimate these prices. Several discretization schemes exist, such as the standard Euler-Maruyama scheme; see e.g. Glasserman [44]. Nevertheless, this discretization has many drawbacks, the most problematic of which is that the CIR approximation process may become negative with a non-zero probability. In practice, this problem can be solved by setting the process value to zero (absorption fix) or its absolute value (reflection fix) when it becomes negative.

Classical schemes such as the Kloeden and Platen scheme [58] produce erratic results when applied to the Heston model. This is mainly due to the volatility process: the square-root diffusion coefficient of a CIR process is not locally Lipschitz around zero. Therefore, a number of authors have proposed alternative approaches to prove the strong and weak convergence for different discretizations of the square root process. Dereich, Neuenkirch, and Szpruch [28] have established the implicit Euler scheme strong convergence which is of the order $\frac{1}{2}$. Moreover, Berkaoui, Bossy, and Diop [13] showed the same order for a symmetrized Euler scheme in a very restricted parameter regime. Recently, Alfonsi [5] and Neuenkirch and Szpruch [68] have shown that the implicit Euler scheme has a strong convergence rate of 1 that is obtained from the Lamperti transformation of the SDE in the case of an inaccessible boundary point.

Consequently, several corrections have been suggested in the literature to improve the Heston model simulation schemes. Firstly, Broadie and Kaya (see [15]) made an important achievement by showing that the transitions in the Heston model can be simulated exactly, without discretization error for the specific case of pricing European option. They have shown that convergence rate is $O(s^{-\frac{1}{2}})$, where s is the total computational budget whereas the rate of convergence of the Euler discretisation method is $O(s^{-\frac{1}{3}})$. Glasserman and Kim [45] developed a representation of the transition in the Heston model. It can be simulated exactly using sums and mixtures of Gamma random variables and a single normal random variable. The speed and accuracy of this approach has been studied for an European call option and compared to the Broadie–Kaya exact method [15]. It has shown that the first approach is better in terms of the Root Mean Squared Error (RMSE). Other approaches developing alternative approximation schemes exist in the literature (see [4], [50] and [54]).

In this work, we are interested in barrier option price calculations under the log-Heston model X using the MLMC method. We consider an approximation scheme $\overline{X}^n$ in which the log-price process is discretised using an Euler scheme whereas the variance process V is simulated exactly. Based on this approximation we build a MLMC method for pricing barrier options using the Brownian bridge technique. To analyse the MLMC variance, we first demonstrate that the strong convergence of the scheme $\overline{X}^n$ is of order $h^{\frac{1}{2}}$ with $h = \frac{T}{n}$. Next, we develop a detailed control of the trajectories of the processes X, $\overline{X}^n$, and V in order to analyse the contributions of extreme and non-extreme trajectories. Under the condition that $\inf_{t\in[0,T]} X_t$ (resp. $\sup_{t\in[0,T]} X_t$) has a continuous density in the vicinity of the barrier D" for the D-O option (resp. U for the U-O option) and for a Lipschitz payoff function f, we obtain that the variance V_ℓ of a given block ℓ in the MLMC estimator is of order $O(h_\ell^\beta)$ with $\beta > \frac{1}{2}$. For a given precision ε, this leads to a time complexity of order $\varepsilon^{-2-(1-\beta)/\alpha}$, where α denotes the weak convergence order. Of course this is far from the optimal regime. The main reason is that the strong convergence of the scheme $\overline{X}^n$ with $n^{1/2}$ significantly slows down the MLMC method. However, to our knowledge, there is no

study in the literature that analyses the asymptotic behaviour of the MLMC method combined with the Brownian bridge technique for pricing barrier options for d-dimensional models with $d > 1$. Furthermore, we have not found other schemes in the literature that are compatible with the Brownian bridge technique and allow for a detailed analysis of extreme and non-extreme trajectories. We tested numerically our algorithm for the Heston model. We also compared its performance with the MLMC antithetic method. Additionally, we study the problem of moment explosion for the Heston model in continuation of the results obtained by Cozma [23].

3.2 General Framework

We consider the stochastic volatility model of Heston [49] solution to

$$\begin{aligned} dS_t &= rS_t dt + \sqrt{V_t} S_t \;\; d\left(\rho W_t^v + \sqrt{1-\rho^2} W_t^s\right) \\ dV_t &= \kappa(\theta - V_t)dt + \sigma\sqrt{V_t} dW_t^v, \text{ for } t \in [0,T], \end{aligned} \tag{3.1}$$

with $\kappa,\ \theta,\ \sigma,\ S_0,\ V_0 > 0,\ r \in \mathbb{R},\ \rho \in [-1,1],\ T > 0$, and $(W_t^s)_{t\in[0,T]}$, $(W_t^v)_{t\in[0,T]}$ are two independent Brownian motions. It is well known that conditionally to V_u the variance process V_t at time $t > u$, has the density of a scaled non-central chi-squared distribution, that is

$$V_t \overset{d}{=} \frac{\sigma^2(1-\mathrm{e}^{-\kappa(t-u)})}{4\kappa} \chi_d^2\Big(\frac{4\kappa\,\mathrm{e}^{-\kappa(t-u)}}{\sigma^2(1-\mathrm{e}^{-\kappa(t-u)})} V_u\Big),$$

where $\chi_d^2(\lambda)$ denotes a non central chi-squared random variable with $d := \frac{4\theta\kappa}{\sigma^2}$ degrees of freedom and non-centrality parameter λ. The log-Heston model defined by $(X_t = \log(S_t))_{t\in[0,T]}$ is solution to

$$\begin{aligned} dX_t &= (r - \frac{1}{2}V_t)dt + \sqrt{V_t} \;\; d\left(\rho W_t^v + \sqrt{1-\rho^2} W_t^s\right), \\ dV_t &= \kappa(\theta - V_t)dt + \sigma\sqrt{V_t} dW_t^v. \end{aligned} \tag{3.2}$$

So, for $0 \le u \le t \le T$ we can rewrite (3.2) as follows

$$X_t = X_u + r(t-u) - \frac{1}{2}\int_u^t V_s ds + \frac{\rho}{\sigma}\Big(V_t - V_u - \kappa\theta(t-u) + \kappa\int_u^t V_s ds\Big) + \sqrt{1-\rho^2}\int_u^t \sqrt{V_s} dW_s^s, \tag{3.3}$$

where and $\int_u^t \sqrt{V_s} dW_s^v = \frac{1}{\sigma}\Big(V_t - V_u - \kappa\theta(t-u) + \kappa\int_u^t V_s ds\Big)$. In what follows, we assume that the Feller condition $2\kappa\theta > \sigma^2$ holds true to guarantee the positivity of the CIR process.

3.3 CIR and log-Heston models: Some useful properties

In this section, we recall some useful properties of the CIR process.

Lemma 3.3.1. *(See [28] for a proof.) We have*

$$\mathbb{E}\Big[\sup_{t\in[0,T]} V_t^p\Big] < \infty \text{ and } \sup_{t\in[0,T]} \mathbb{E}[V_t^p] < \infty \text{ for } p > -\frac{2\kappa\theta}{\sigma^2}.$$

For all $p \ge 1$, there exists a constant C depending only on p, T, κ, θ, σ, and $V_0 \ge 0$ such that

$$\mathbb{E}|V_t - V_s|^p \le C|t-s|^{p/2}, \;\; s,t \in [0,T].$$

Lemma 3.3.2. *[52].*
Let $V_0 > 0$ and $T > 0$. For all $p \geq 0$, there exists constants $C > 0$, such that

$$\mathbb{E}|\sqrt{V_t} - \sqrt{V_s}|^p \leq C|t-s|^{p/2}, \;\; s,t \in [0,T].$$

The following result plays a key role in determining the variance estimator convergence rate. This will be used later in work.

Lemma 3.3.3. *Let $\kappa\theta > \sigma^2$, for all $p < \frac{\kappa\theta}{\sigma^2}$ there exists a positive constant C depending only on p, T, κ, θ, σ, and $V_0 > 0$ such that*

$$\Big(\mathbb{E}\Big|\frac{1}{V_t} - \frac{1}{V_s}\Big|^p\Big)^{\frac{1}{p}} \leq C|t-s|^{\frac{1}{2}}, \;\; s,t \in [0,T].$$

Proof. We have

$$\mathbb{E}\Big|\frac{1}{V_t} - \frac{1}{V_s}\Big| \leq \mathbb{E}\frac{1}{|V_t V_s|}\mathbb{E}|V_t - V_s|$$

By using Hölder's inequality, for $q, q' > 1$ with $\frac{1}{q} + \frac{1}{q'} = \frac{1}{p}$ yields

$$\Big(\mathbb{E}\Big|\frac{1}{V_t} - \frac{1}{V_s}\Big|^p\Big)^{\frac{1}{p}} \leq \Big(\mathbb{E}\frac{1}{|V_t V_s|^q}\Big)^{\frac{1}{q}}\Big(\mathbb{E}|V_t - V_s|^{q'}\Big)^{\frac{1}{q'}}.$$

Since

$$\mathbb{E}\frac{1}{|V_t V_s|^q} \leq \mathbb{E}\frac{1}{|V_t|^{2q}} + \mathbb{E}\frac{1}{|V_s|^{2q}},$$

and according to Lemma 3.3.1 we get

$$\Big(\mathbb{E}\Big|\frac{1}{V_t} - \frac{1}{V_s}\Big|^p\Big)^{\frac{1}{p}} \leq C|t-s|^{\frac{1}{2}}$$

with $q < \frac{\kappa\theta}{\sigma^2}$, which completes the proof. □

Proposition 3.3.4. *For all $p \geq 1$, we have*

$$\mathbb{E}\Big[\sup_{0\leq t\leq T}|X_t|^p\Big] < \infty.$$

Proof. Similarly to the proof of Lemma 3.3.1, we can prove the following result L^p for the log-price process. □

3.4 Barrier call option

In this section, we study the problem of pricing barrier option. We are particularly interested in pricing a Down and Out (D-O) and up-and-out (U-P) options given respectively by

$$\pi_D = \mathbb{E}\Big[f(X_T)\mathbb{1}_{\{\tau_D > T\}}\Big] \text{ and } \pi_U = \mathbb{E}\Big[f(X_T)\mathbb{1}_{\{\tau_U > T\}}\Big], \tag{3.4}$$

where denotes a given payoff option and τ_D and τ_U denote the first passage time given by:

- $\tau_D = \inf\{t \in [0,T], X_t \leq D\}$ with $X_0 > D > 0$, for a Down-Out (D-O) option,

  and

- $\tau_U = \inf\{t \in [0,T], X_t \geq U\}$ with $0 < X_0 < U$, for an Up-Out (U-O) option.

Now, for a given time grid $t_i = \frac{T}{n}$, $i \in \{1,\dots,n\}$, we approximate the pricing path-dependent options using Brownian interpolation conditional on the two extreme values $\overline{X}^n_{t_i}$ and $\overline{X}^n_{t_{i+1}}$ in the time interval $[t_i, t_{i+1}]$ yielding

$$\begin{aligned}\overline{X}^n_t = \overline{X}^n_{t_i} + (r - \frac{1}{2}V_{t_i})(t-t_i) + \frac{\rho}{\sigma}\Big(V^{int}_t - V_{t_i} - \kappa\theta(t-t_i) + \kappa V_{t_i}(t-t_i)\Big) \\ + \sqrt{1-\rho^2}\sqrt{V_{t_i}}(W^s_t - W^s_{t_i}),\end{aligned}$$

with $V^{int}_t = \frac{n}{T}\big[(t_{i+1}-t)V_{t_i} + (t-t_i)V_{t_{i+1}}\big]$, $t \in [t_i, t_{i+1}]$ and $(V_t)_{t\geq 0}$ is simulated exactly. This can be rewritten as

$$\begin{aligned}\overline{X}^n_t = \overline{X}^n_{t_i} + (r - \frac{1}{2}V_{t_i})(t-t_i) + \frac{\rho}{\sigma}\Big(\frac{n}{T}(V_{t_{i+1}} - V_{t_i})(t-t_i) - \kappa\theta(t-t_i) + \kappa V_{t_i}(t-t_i)\Big) \\ + \sqrt{1-\rho^2}\sqrt{V_{t_i}}(W^s_t - W^s_{t_i}).\end{aligned} \tag{3.5}$$

Then, the above option prices can be approximated respectively by

$$\overline{\pi}_D := \mathbb{E}\Big[f(\overline{X}^n_T)\prod_{i=0}^{n-1}\mathbb{1}_{\{\inf_{t\in[t_i,t_{i+1}]}\overline{X}^n_t > D\}}\Big] \text{ and } \overline{\pi}_U := \mathbb{E}\Big[f(\overline{X}^n_T)\prod_{i=0}^{n-1}\mathbb{1}_{\{\sup_{t\in[t_i,t_{i+1}]}\overline{X}^n_t < U\}}\Big].$$

To get more accurate approximations, we use the Brownian bridge technique to substitute the barrier-crossing indicators by the probabilities that the approximation scheme $\overline{S}_t$ do not cross the barrier in each time interval (t_i, t_{i+1}), $i \in \{1,\dots,n\}$. In what follows, for $x \in \mathbb{R}$, $(x)_+$ stands for $\max(x, 0)$.

Proposition 3.4.1. *Under the above notations, for $h = \frac{T}{n}$, we have*

$$\overline{\pi}_U = \mathbb{E}\Big[f(\overline{X}^n_T)\prod_{i=0}^{n-1}(1-\overline{p}_i)\Big], \text{ where } \overline{p}_i = \exp\Big(\frac{-2(U-\overline{X}^n_{t_i})_+(U-\overline{X}^n_{t_{i+1}})_+}{(1-\rho^2)V_{t_i}h}\Big)$$

and

$$\overline{\pi}_D = \mathbb{E}\Big[f(\overline{X}^n_T)\prod_{i=0}^{n-1}(1-\overline{q}_i)\Big], \text{ where } \overline{q}_i := \exp\Big(\frac{-2(\overline{X}^n_{t_i}-D)_+(\overline{X}^n_{t_{i+1}}-D)_+}{(1-\rho^2)V_{t_i}h}\Big).$$

Proof. We give a proof for the U-O barrier option. At first, noticing that conditionally on $(\overline{X}^n_0, \overline{X}^n_{t_1}, \dots, \overline{X}^n_{t_n}, V_{t_0}, ..., V_{t_n})$, the barrier-crossing indicators $(\mathbb{1}_{\{\sup_{t\in[t_i,t_{i+1}]}\overline{X}^n_t < U\}}, i \in \{1,\dots,n\})$ are independent, we write

$$\begin{aligned}\overline{\pi}_U &= \mathbb{E}\Big[\mathbb{E}\big[f(\overline{X}^n_T)\prod_{i=0}^{n-1}\mathbb{1}_{\{\sup_{t\in[t_i,t_{i+1}]}\overline{X}^n_t < U\}}|\overline{X}^n_0, \overline{X}^n_{t_1}, \dots, \overline{X}^n_{t_n}, V_{t_0}, V_{t_1}, ..., V_{t_n}\big]\Big] \\ &= \mathbb{E}\Big[f(\overline{X}^n_T)\prod_{i=0}^{n-1}\mathbb{E}\big[\mathbb{1}_{\{\sup_{t\in[t_i,t_{i+1}]}\overline{X}^n_t < U\}}|\overline{X}^n_{t_i}, \overline{X}^n_{t_{i+1}}, V_{t_i}, V_{t_{i+1}}\big]\Big] \\ &= \mathbb{E}\Big[f(\overline{X}^n_T)\prod_{i=0}^{n-1}\big(1-\varphi(\overline{X}^n_{t_i}, \overline{X}^n_{t_{i+1}}, V_{t_i}, V_{t_{i+1}})\big)\Big],\end{aligned}$$

where, for $x_i, x_{i+1} \in \mathbb{R}$ and $v_i, v_{i+1} \in \mathbb{R}^*_+$,

$$\varphi(x_i, x_{i+1}, v_i, v_{i+1}) = \mathbb{P}\big(\sup_{t\in[t_i,t_{i+1}]}\overline{X}^n_t \geq U | \overline{X}^n_{t_i} = x_i, \overline{X}^n_{t_{i+1}} = x_{i+1}, V_{t_i} = v_i, V_{t_{i+1}} = v_{i+1}\big).$$

Now, using (3.5), we have

$$\sup_{t\in[t_i,t_{i+1}]}\overline{X}^n_t=\overline{X}^n_{t_i}+\sqrt{1-\rho^2}\sqrt{V_{t_i}}\times\sup_{t\in[t_i,t_{i+1}]}\Bigg[(W^s_t-W^s_{t_i})+\frac{r-\frac{1}{2}V_{t_i}}{\sqrt{1-\rho^2}\sqrt{V_{t_i}}}(t-t_i)$$
$$+\frac{\rho}{\sigma\sqrt{1-\rho^2}\sqrt{V_{t_i}}}\Big(\frac{n}{T}(V_{t_{i+1}}-V_{t_i})(t-t_i)-\kappa\theta(t-t_i)+\kappa V_{t_i}(t-t_i)\Big)\Bigg],\ \text{where}$$
$$W^s_{t_{i+1}}-W^s_{t_i}+\frac{r-\frac{V_{t_i}}{2}}{\sqrt{1-\rho^2}\sqrt{V_{t_i}}}(t_{i+1}-t_i)+\frac{\rho}{\sigma\sqrt{1-\rho^2}\sqrt{V_{t_i}}}$$
$$\times\Big(\frac{n}{T}(V_{t_{i+1}}-V_{t_i})(t_{i+1}-t_i)-\kappa\theta(t_{i+1}-t_i)+\kappa V_{t_i}(t_{i+1}-t_i)\Big)=\frac{\overline{X}^n_{t_{i+1}}-\overline{X}^n_{t_i}}{\sqrt{1-\rho^2}\sqrt{V_{t_i}}}.$$

By the stationarity property of the brownian increments and using a change of probability measure, we easily get that the law of $\sup_{t\in[t_i,t_{i+1}]}\big[W^s_t-W^s_{t_i}+L(v_i,v_{i+1})(t-t_i)\big]$ given $W_{t_{i+1}}-W_{t_i}+L(v_i,v_{i+1})(t_{i+1}-t_i)=\frac{1}{\sqrt{1-\rho^2}\sqrt{v_i}}(x_{i+1}-x_i)$ is equal to the law of $\sup_{t\in[0,t_1]}W_t$ given $W_{t_1}=\frac{1}{\sqrt{1-\rho^2}\sqrt{v_i}}(x_{i+1}-x_i)$ which is given by $\mathbb{P}\Big(\max_{t\in[0,t_1]}W_{t_1}\geq y\Big|W_{t_1}=x\Big)=\mathrm{e}^{\frac{-2(y)_+(y-x)_+}{h}}$ (see e.g. [55, p. 265]). Thus, we get

$$\varphi(x_i,x_{i+1},v_i,v_{i+1})=\mathbb{P}\Big(\sup_{t\in[0,t_1]}W_t\geq\frac{1}{\sqrt{1-\rho^2}\sqrt{v_i}}(U-x_i)|W_{t_1}=\frac{1}{\sqrt{1-\rho^2}\sqrt{v_i}}(x_{i+1}-x_i)\Big)$$
$$=\exp\Big(\frac{-2(U-x_i)_+(U-x_{i+1})_+}{(1-\rho^2)v_ih}\Big).$$

□

The following result gives us the strong convergence rate of our approximation scheme (3.5).

Theorem 3.4.2. *Let $T>0$. For all $p>1$, there exists a constant C_p such that*

$$\max_{1\leq i\leq n}\mathbb{E}|X_{t_i}-\overline{X}^n_{t_i}|^p\leq C_p\Big(\frac{T}{n}\Big)^{p/2}$$

Proof. According to the expression (3.3) we have

$$X_{t_i}=X_0+rt_i-\frac{1}{2}\int_0^{t_i}V_sds+\frac{\rho}{\sigma}(V_{t_i}-V_0-\kappa\theta t_i+\kappa\int_0^{t_i}V_sds)+\sqrt{1-\rho^2}\int_0^{t_i}\sqrt{V_s}dW^s_s.\tag{3.6}$$

In addition, we have from (3.5)

$$\overline{X}^n_{t_i}=X_0+rt_i+(\frac{\kappa\rho}{\sigma}-\frac{1}{2})\frac{T}{n}\sum_{k=0}^{i-1}V_{t_k}+\frac{\rho}{\sigma}\big(V_{t_i}-V_0-\kappa\theta t_i\big)+\sqrt{1-\rho^2}\sum_{k=0}^{i-1}\int_{t_k}^{t_{k+1}}\sqrt{V_{\eta(s)}}dW^s_s,$$

where $\eta(t)=t_k$ if $t\in[t_k,t_{k+1})$. Then,

$$|X_{t_i}-\overline{X}^n_{t_i}|\leq|\frac{\kappa\rho}{\sigma}-\frac{1}{2}|\int_0^{t_i}V_s-V_{\eta(s)}ds+\sqrt{1-\rho^2}\int_0^{t_i}\sqrt{V_s}-\sqrt{V_{\eta(s)}}dW^s_s,$$

which leads to

$$\mathbb{E}|X_{t_i}-\overline{X}^n_{t_i}|^p\leq2^{p-1}\Big(|\frac{\kappa\rho}{\sigma}-\frac{1}{2}|^p\mathbb{E}\Big|\int_0^{t_i}V_s-V_{\eta(s)}ds\Big|^p+\mathbb{E}\Big|\int_0^{t_i}\sqrt{V_s}-\sqrt{V_{\eta(s)}}dW^s_s\Big|^p\Big).$$

By Jensen inequality and Burkholder-Davis-Gundy inequality we get

$$\mathbb{E}|X_{t_i}-\overline{X}^n_{t_i}|^p\leq2^{p-1}\Big(|\frac{\kappa\rho}{\sigma}-\frac{1}{2}|^pT^{p-1}\mathbb{E}\int_0^{t_i}|V_s-V_{\eta(s)}|^pds+(1-\rho^2)^{\frac{p}{2}}\mathbb{E}\Big[\Big(\int_0^{t_i}|\sqrt{V_s}-\sqrt{V_{\eta(s)}}|^2ds\Big)^{\frac{p}{2}}\Big]\Big).$$

Now, applying Jensen inequality and Fubini's theorem we obtain the existence a positive constant $C_{p,T}$ such that

$$\mathbb{E}|X_{t_i} - \overline{X}^n_{t_i}|^p \leq C_{p,T}\Big(\int_0^T \mathbb{E}|V_s - V_{\eta(s)}|^p ds + \int_0^T \mathbb{E}|\sqrt{V_s} - \sqrt{V_{\eta(s)}}|^p ds\Big)$$

We use lemmas 3.3.2 and 3.3.1 to conclude that the convergence error is of order $\Big(\frac{T}{n}\Big)^{\frac{p}{2}}$ which completes the proof. □

3.5 An improved MLMC method with the Brownian bridge technique

In this section, we consider the scheme $(\overline{X}^{2^\ell}_{t_i})_{0\leq i\leq 2^\ell}$ given in (3.5) that approximates $(X_t)_{0\leq t\leq T}$ solution to (3.2) using a time step $h_\ell = 2^{-\ell}T$ for $\ell \in \{0, ..., L\}$, with $L = \dfrac{\log n}{\log 2}$, where n denotes the finest time step number. As the same arguments work for both D-O and U-O barrier options, we give details only for the latter one. To do so, we write

$$\overline{P}_\ell = f(\overline{X}^{2^\ell}_T)\prod_{i=0}^{2^\ell-1} \mathbb{1}_{\{\sup_{t\in[t^\ell_i, t^\ell_{i+1}]}\overline{X}^{2^\ell}_t < U\}}, \qquad \text{where } t^\ell_i = ih_\ell \quad \text{for } \ell \in \{0, ..., L\}. \tag{3.7}$$

It is clear that

$$\overline{\pi}_U = \mathbb{E}\big[\overline{P}_L\big] = \mathbb{E}\big[\overline{P}_0\big] + \sum_{\ell=1}^{L}\mathbb{E}\big[\overline{P}_\ell - \overline{P}_{\ell-1}\big]. \tag{3.8}$$

On the one hand, applying Proposition 3.4.1 yields

$$\mathbb{E}\big[\overline{P}_\ell\big] = \mathbb{E}\big[\overline{P}^f_\ell\big], \text{ where } \overline{P}^f_\ell := f(\overline{X}^{2^\ell}_T)\prod_{i=0}^{2^\ell-1}(1 - \overline{p}^{2^\ell}_i) \text{ with} \tag{3.9}$$

$$\overline{p}^{2^\ell}_i = \exp\Big(\frac{-2(U - \overline{X}^{2^\ell}_{t_i})_+(U - \overline{X}^{2^\ell}_{t_{i+1}})_+}{(1-\rho^2)V_{t^\ell_i}h_\ell}\Big).$$

On the other hand, we first write

$$\mathbb{E}\big[\overline{P}_{\ell-1}\big] = \mathbb{E}\Big[f(\overline{X}^{2^{\ell-1}}_T)\prod_{i=0}^{2^{\ell-1}-1}\mathbb{E}\big[\mathbb{1}_{\{\sup_{t\in[t^{\ell-1}_i, t^{\ell-1}_{i+1}]}\overline{X}^{2^{\ell-1}}_t < U\}}|\overline{X}^{2^{\ell-1}}_{t^{\ell-1}_i}, \overline{X}^{2^{\ell-1}}_{t^\ell_{2i+1}}, \overline{X}^{2^{\ell-1}}_{t^{\ell-1}_{i+1}}, V_{t^\ell_i}, V_{t^{\ell-1}_i}\big]\Big]$$

$$= \mathbb{E}\Big[f(\overline{X}^{2^{\ell-1}}_T)\prod_{i=0}^{2^{\ell-1}-1}\mathbb{E}\big[\mathbb{1}_{\{\sup_{t\in[t^{\ell-1}_i, t^\ell_{2i+1}]}\overline{X}^{2^{\ell-1}}_t < U\}}\mathbb{1}_{\{\sup_{t\in[t^\ell_{2i+1}, t^{\ell-1}_{i+1}]}\overline{X}^{2^{\ell-1}}_t < U\}}|\overline{X}^{2^{\ell-1}}_{t^{\ell-1}_i}, \overline{X}^{2^{\ell-1}}_{t^\ell_{2i+1}}, \overline{X}^{2^{\ell-1}}_{t^{\ell-1}_{i+1}}, V_{t^\ell_i}, V_{t^{\ell-1}_i}\big]\Big],$$

where the coarse scheme $\overline{X}^{2^{\ell-1}}_{t^\ell_{2i+1}}$ is computed using the interpolated scheme (3.5) that is

$$\begin{aligned}\overline{X}^{2^{\ell-1}}_{t^\ell_{2i+1}} &= \overline{X}^{2^{\ell-1}}_{t^{\ell-1}_i} + (r - \frac{1}{2}V_{t^{\ell-1}_i})(t^\ell_{2i+1} - t^{\ell-1}_i)\\ &+ \frac{\rho}{\sigma}\Big(\frac{2^{\ell-1}}{T}(V_{t^{\ell-1}_{i+1}} - V_{t^{\ell-1}_i})(t^\ell_{2i+1} - t^{\ell-1}_i) - \kappa\theta(t^\ell_{2i+1} - t^{\ell-1}_i) + \kappa V_{t^{\ell-1}_i}(t^\ell_{2i+1} - t^{\ell-1}_i)\Big)\\ &+ \sqrt{1-\rho^2}\sqrt{V_{t^{\ell-1}_i}}(W^s_{t^\ell_{2i+1}} - W^s_{t^{\ell-1}_i}).\end{aligned}$$

Thus, we rewrite $\sup_{t\in[t_i^{\ell-1},t_{2i+1}^{\ell}]}\overline{X}_t^{2^{\ell-1}}$ and $\sup_{t\in[t_{2i+1}^{\ell},t_{i+1}^{\ell-1}]}\overline{X}_t^{2^{\ell-1}}$ as follows

$$\sup_{t\in[t_i^{\ell-1},t_{2i+1}^{\ell}]}\overline{X}_t^{2^{\ell-1}} = \overline{X}_{t_i^{\ell-1}}^{2^{\ell-1}} + \sqrt{1-\rho^2}\sqrt{V_{t_i^{\ell-1}}} \times \sup_{t\in[t_i^{\ell-1},t_{2i+1}^{\ell}]}\Bigg[(W_t^s - W_{t_i^{\ell-1}}^s) + \frac{r-\frac{1}{2}V_{t_i^{\ell-1}}}{\sqrt{1-\rho^2}\sqrt{V_{t_i^{\ell-1}}}}(t-t_i^{\ell-1}) + \frac{\rho}{\sigma\sqrt{1-\rho^2}\sqrt{V_{t_i^{\ell-1}}}}\Big(\frac{2^{\ell-1}}{T}(V_{t_{i+1}^{\ell-1}} - V_{t_i^{\ell-1}})(t-t_i^{\ell-1}) - \kappa\theta(t-t_i^{\ell-1}) + \kappa V_{t_i^{\ell-1}}(t-t_i^{\ell-1})\Big)\Bigg]$$

with

$$\begin{aligned}
&W_{t_{2i+1}^{\ell}}^s - W_{t_i^{\ell-1}}^s + \frac{r-\frac{1}{2}V_{t_i^{\ell-1}}}{\sqrt{1-\rho^2}\sqrt{V_{t_i^{\ell-1}}}}(t_{2i+1}^{\ell} - t_i^{\ell-1})\\
&+\frac{\rho}{\sigma\sqrt{1-\rho^2}\sqrt{V_{t_i^{\ell-1}}}}\Big(\frac{2^{\ell-1}}{T}(V_{t_{i+1}^{\ell-1}} - V_{t_i^{\ell-1}})(t_{2i+1}^{\ell} - t_i^{\ell-1}) - \kappa\theta(t_{2i+1}^{\ell} - t_i^{\ell-1}) + \kappa V_{t_i^{\ell-1}}(t_{2i+1}^{\ell} - t_i^{\ell-1})\Big)\\
&=\frac{\overline{X}_{t_{2i+1}^{\ell}}^{2^{\ell-1}} - \overline{X}_{t_i^{\ell-1}}^{2^{\ell-1}}}{\sqrt{1-\rho^2}\sqrt{V_{t_i^{\ell-1}}}}
\end{aligned}$$

and

$$\sup_{t\in[t_{2i+1}^{\ell},t_{i+1}^{\ell-1}]}\overline{X}_t^{2^{\ell-1}} = \overline{X}_{t_{2i+1}^{\ell}}^{2^{\ell-1}} + \sqrt{1-\rho^2}\sqrt{V_{t_i^{\ell-1}}} \times \sup_{t\in[t_{2i+1}^{\ell},t_{i+1}^{\ell-1}]}\Bigg[(W_t^s - W_{t_{2i+1}^{\ell}}^s) + \frac{r-\frac{1}{2}V_{t_i^{\ell-1}}}{\sqrt{1-\rho^2}\sqrt{V_{t_i^{\ell-1}}}}(t-t_{2i+1}^{\ell}) + \frac{\rho}{\sigma\sqrt{1-\rho^2}\sqrt{V_{t_i^{\ell-1}}}}\Big(\frac{2^{\ell-1}}{T}(V_{t_{i+1}^{\ell-1}} - V_{t_i^{\ell-1}})(t-t_{2i+1}^{\ell}) - \kappa\theta(t-t_{2i+1}^{\ell}) + \kappa V_{t_i^{\ell-1}}(t-t_{2i+1}^{\ell})\Big)\Bigg]$$

with

$$\begin{aligned}
&W_{t_{i+1}^{\ell-1}}^s - W_{t_{2i+1}^{\ell}}^s + \frac{r-\frac{1}{2}V_{t_i^{\ell-1}}}{\sqrt{1-\rho^2}\sqrt{V_{t_i^{\ell-1}}}}(t_{i+1}^{\ell-1} - t_{2i+1}^{\ell})\\
&+\frac{\rho}{\sigma\sqrt{1-\rho^2}\sqrt{V_{t_i^{\ell-1}}}}\Big(\frac{2^{\ell-1}}{T}(V_{t_{i+1}^{\ell-1}} - V_{t_i^{\ell-1}})(t_{i+1}^{\ell-1} - t_{2i+1}^{\ell}) - \kappa\theta(t_{i+1}^{\ell-1} - t_{2i+1}^{\ell}) + \kappa V_{t_i^{\ell-1}}(t_{i+1}^{\ell-1} - t_{2i+1}^{\ell})\Big)\\
&=\frac{\overline{X}_{t_{i+1}^{\ell-1}}^{2^{\ell-1}} - \overline{X}_{t_{2i+1}^{\ell}}^{2^{\ell-1}}}{\sqrt{1-\rho^2}\sqrt{V_{t_i^{\ell-1}}}}.
\end{aligned}$$

Then, using the same arguments as in the proof of Proposition 3.4.1, we get

$$\mathbb{E}\big[\overline{P}_{\ell-1}\big] = \mathbb{E}\big[\overline{P}_{\ell-1}^c\big], \text{ where } \overline{P}_{\ell-1}^c := f(\overline{X}_T^{2^{\ell-1}})\prod_{i=0}^{2^{\ell-1}-1}(1-\overline{p}_{i,1}^{2^{\ell-1}})(1-\overline{p}_{i,2}^{2^{\ell-1}}) \text{ with} \tag{3.10}$$

$$\overline{p}_{i,1}^{2^{\ell-1}} = \exp\Big(\frac{-2(U-\overline{X}_{t_i^{\ell-1}}^{2^{\ell-1}})_+(U-\overline{X}_{t_{2i+1}^{\ell}}^{2^{\ell-1}})_+}{(1-\rho^2)V_{t_i^{\ell-1}}h_\ell}\Big),$$

$$\overline{p}_{i,2}^{2^{\ell-1}} = \exp\Big(\frac{-2(U-\overline{X}_{t_{2i+1}^{\ell}}^{2^{\ell-1}})_+(U-\overline{X}_{t_{i+1}^{\ell-1}}^{2^{\ell-1}})_+}{(1-\rho^2)V_{t_i^{\ell-1}}h_\ell}\Big),$$

which can be rewritten as

$$\overline{P}_{\ell-1}^c := f(\overline{X}_T^{2^{\ell-1}})\prod_{i=0}^{2^{\ell}-1}(1-\overline{p}_i^{2^{\ell-1}}) \text{ with } \overline{p}_i^{2^{\ell-1}} = \exp\Big(\frac{-2(U-\overline{X}_{t_i^{\ell}}^{2^{\ell-1}})_+(U-\overline{X}_{t_{i+1}^{\ell}}^{2^{\ell-1}})_+}{(1-\rho^2)V_{t_i^{\ell-1}}h_\ell}\Big). \tag{3.11}$$

Thus, the improved MLMC method approximates $\bar{\pi}_U$ by

$$\bar{P}_U := \frac{1}{N_0}\sum_{k=1}^{N_0}\overline{P}^f_{0,k} + \sum_{\ell=1}^{L}\frac{1}{N_\ell}\sum_{k=1}^{N_\ell}\Big(\overline{P}^f_{\ell,k} - \overline{P}^c_{\ell-1,k}\Big), \tag{3.12}$$

where the condition $\mathbb{E}\big[\overline{P}^f_{\ell-1}\big] = \mathbb{E}\big[\overline{P}^c_{\ell-1}\big]$ is satisfied according to (3.9) and (3.10). Here, $(\overline{P}^f_{\ell,k})_{1\le k\le N_\ell}$ and $(\overline{P}^c_{\ell-1,k})_{1\le k\le N_\ell}$ are respectively independent copies of $\overline{P}^f_\ell$ and $\overline{P}^c_{\ell-1}$ given by (3.9) and (3.11). Similarly, the improved MLMC method approximates $\bar{\pi}_D$ by

$$\bar{Q}_D := \frac{1}{N_0}\sum_{k=1}^{N_0}\overline{Q}^f_{0,k} + \sum_{\ell=1}^{L}\frac{1}{N_\ell}\sum_{k=1}^{N_\ell}\Big(\overline{Q}^f_{\ell,k} - \overline{Q}^c_{\ell-1,k}\Big), \tag{3.13}$$

where $(\overline{Q}^f_{\ell,k})_{1\le k\le N_\ell}$ and $(\overline{Q}^c_{\ell-1,k})_{1\le k\le N_\ell}$ are respectively independent copies of $\overline{Q}^f_\ell$ and $\overline{Q}^c_{\ell-1}$ given by

$$\overline{Q}^f_\ell := f(\overline{X}^{2^\ell}_T)\prod_{i=0}^{2^\ell-1}(1-\overline{q}^{2^\ell}_i) \text{ with } \overline{q}^{2^\ell}_i = \exp\Big(\frac{-2(\overline{X}^{2^\ell}_{t^\ell_i}-D)_+(\overline{X}^{2^\ell}_{t^\ell_{i+1}}-D)_+}{(1-\rho^2)V_{t^\ell_i}h_\ell}\Big)$$

$$\overline{Q}^c_{\ell-1} := f(\overline{X}^{2^{\ell-1}}_T)\prod_{i=0}^{2^\ell-1}(1-\overline{q}^{2^{\ell-1}}_i) \text{ with } \overline{q}^{2^{\ell-1}}_i = \exp\Big(\frac{-2(\overline{X}^{2^{\ell-1}}_{t^\ell_i}-D)_+(\overline{X}^{2^{\ell-1}}_{t^\ell_{i+1}}-D)_+}{(1-\rho^2)V_{t^{\ell-1}_i}h_\ell}\Big).$$

3.5.1 Extreme path events

Regarding the multilevel estimators variance for a barrier option, the analysis follows the extreme path approach introduced in [38]. The next results examine the extreme paths of the asset price, its variance process and their approximation schemes. Lemmas 3.5.1 and 3.5.2 show that certain extreme paths make a negligible contribution to the overall expectation. This result is paramount in the complexity theorem demonstration.

Lemma 3.5.1. *For any $\eta > 0$ and $h_\ell = 2^{-\ell}T$, the following extreme path events satisfy*

$$\max_{0\le i\le 2^\ell}\mathbb{P}\left(\max\left(|\overline{X}^{2^\ell}_{t^\ell_i}|, |\overline{X}^{2^{\ell-1}}_{t^\ell_i}|\right) > h_\ell^{-\eta}\right) = o(h_\ell^q) \tag{3.14}$$

$$\max_{0\le i\le 2^\ell}\mathbb{P}\left(\max\left(\big(|X_{t^\ell_i}-\overline{X}^{2^\ell}_{t^\ell_i}|, |X_{t^\ell_i}-\overline{X}^{2^{\ell-1}}_{t^\ell_i}|, |\overline{X}^{2^\ell}_{t^\ell_i}-\overline{X}^{2^{\ell-1}}_{t^\ell_i}|\big)\right) > h_\ell^{1/2-\eta}\right) = o(h_\ell^q) \tag{3.15}$$

$$\mathbb{P}\left(\sup_{t\in[0,T]}|W_t - W_{t^\ell_i}| > h_\ell^{\frac{1}{2}-\eta}\right) = o(h_\ell^q) \tag{3.16}$$

for all $q > 0$.

Proof. For the first extreme path property, we have

$$\mathbb{P}\left(\max\left(|X_{t^\ell_i}|, |\overline{X}^{2^\ell}_{t^\ell_i}|, |\overline{X}^{2^{\ell-1}}_{t^\ell_i}|\right) > h_\ell^{-\eta}\right) \le \mathbb{P}((|X_{t^\ell_i}| > h_\ell^{-\eta}) + \mathbb{P}((|\overline{X}^{2^\ell}_{t^\ell_i}| > h_\ell^{-\eta}) + \mathbb{P}(|\overline{X}^{2^{\ell-1}}_{t^\ell_i}| > h_\ell^{-\eta}). \tag{3.17}$$

Then, by Markov's inequality we get for $m \ge 1$

$$\mathbb{P}\left(\max\left(|\overline{X}^{2^\ell}_{t^\ell_i}|, |\overline{X}^{2^{\ell-1}}_{t^\ell_i}|\right) > h_\ell^{-\eta}\right) \le h_\ell^{m\eta}\left(\mathbb{E}[|X_{t^\ell_i}|^m] + \mathbb{E}[|\overline{X}^{2^\ell}_{t^\ell_i}|^m] + \mathbb{E}[|\overline{X}^{2^{\ell-1}}_{t^\ell_i}|^m]\right)$$

The result follows using Proposition 3.3.4 for h_ℓ sufficiently small. For the second extreme path property, we proceed in the same way to get

$$\begin{aligned}\mathbb{P}&\left(\max\left(\big(|X_{t^\ell_i}-\overline{X}^{2^\ell}_{t^\ell_i}|, |X_{t^\ell_i}-\overline{X}^{2^{\ell-1}}_{t^\ell_i}|, |\overline{X}^{2^\ell}_{t^\ell_i}-\overline{X}^{2^{\ell-1}}_{t^\ell_i}|\big)\right) > h_\ell^{1/2-\eta}\right)\\ &\le \mathbb{P}(|X_{t^\ell_i}-\overline{X}^{2^\ell}_{t^\ell_i}| > h_\ell^{1/2-\eta}) + \mathbb{P}(|X_{t^\ell_i}-\overline{X}^{2^{\ell-1}}_{t^\ell_i}| > h_\ell^{1/2-\eta}) + \mathbb{P}(|\overline{X}^{2^\ell}_{t^\ell_i}-\overline{X}^{2^{\ell-1}}_{t^\ell_i}| > h_\ell^{1/2-\eta}).\end{aligned} \tag{3.18}$$

Then, by Markov's inequality we get for $m \geq 1$

$$\mathbb{P}\left(\max\left(\left(|X_{t_i^\ell} - \overline{X}^{2^\ell}_{t_i^\ell}|, |X_{t_i^\ell} - \overline{X}^{2^{\ell-1}}_{t_i^\ell}|, |\overline{X}^{2^\ell}_{t_i^\ell} - \overline{X}^{2^{\ell-1}}_{t_i^\ell}|\right)\right.\right.$$
$$\left.\left. \leq h_\ell^{-m(1/2-\eta)}\left(\mathbb{E}[|X_{t_i^\ell} - \overline{X}^{2^\ell}_{t_i^\ell}|^m] + \mathbb{E}[|X_{t_i^\ell} - \overline{X}^{2^{\ell-1}}_{t_i^\ell}|^m] + \mathbb{E}[|\overline{X}^{2^\ell}_{t_i^\ell} - \overline{X}^{2^{\ell-1}}_{t_i^\ell}|^m]\right)\right).$$

Thus, we deduce the result using Theorem 3.4.2 for all $q > 0$ and for h_ℓ sufficiently small. The third property follows using that $\mathbb{E}\left[\sup_{t\in[0,T]} |W_t|^m\right]$ is finite for any positive power m.

□

Lemma 3.5.2. *Let $2\kappa\theta > \sigma^2$, for any $\eta > 0$ and $h_\ell = 2^{-\ell}T$, the following extreme path events satisfy*

$$\max_{0\leq i\leq 2^\ell} \mathbb{P}\left(\max\left(\left(|\frac{1}{V_{t_i^{\ell-1}}}|, |\frac{1}{V_{t_i^\ell}}|\right) > h_\ell^{-\eta}\right)\right. = o(h_\ell^q), \quad \textit{for all } 0 < q < \frac{2\kappa\theta}{\sigma^2}\eta \tag{3.19}$$

$$\max_{0\leq i\leq 2^\ell} \mathbb{P}\Big(\left(|\frac{1}{V_{t_i^\ell}} - \frac{1}{V_{t_i^{\ell-1}}}|\right) > h_\ell^{1/2-\eta}\Big) = o(h_\ell^q), \quad \textit{for all } 0 < q < \frac{\kappa\theta}{\sigma^2}\eta \tag{3.20}$$

$$\max_{0\leq i\leq 2^\ell} \mathbb{P}\left(\int_{t_i^\ell}^{t_{i+1}^\ell} V_s ds > h_\ell^{1-\eta}\right) = o(h_\ell^q), \quad \textit{for all } q > 0 \tag{3.21}$$

$$\max_{0\leq i\leq 2^\ell} \mathbb{P}\left(\sup_{t\in[t_i^\ell, t_{i+1}^\ell]} |\int_{t_i^\ell}^t \sqrt{V_s} dW_s^s| > h_\ell^{\frac{1}{2}-\eta}\right) = o(h_\ell^q), \quad \textit{for all } q > 0. \tag{3.22}$$

Proof. For the first extreme path event we used a similar argument, for $m \geq 1$ we get

$$\mathbb{P}\left(\max\left(|\frac{1}{V_{t_i^{\ell-1}}}|, |\frac{1}{V_{t_i^\ell}}|\right) > h_\ell^{-\eta}\right) \leq h_\ell^{m\eta}\left(\mathbb{E}[\sup_{0\leq t\leq T} |\frac{1}{V_t}|^m] + \mathbb{E}[\sup_{0\leq t\leq T} |\frac{1}{V_t}|^m]\right).$$

Thus, we conclude using Lemma 3.3.1 with choosing $0 < \frac{q}{\eta} < m \leq \frac{2\kappa\theta}{\sigma^2}$, for h_ℓ sufficiently small. For the second extreme path event, we proceed in the same way. Thus, for $m \geq 1$ we have the existence of a positive constant $C > 0$ such that

$$\mathbb{P}\Big(\left(|\frac{1}{V_{t_i^\ell}} - \frac{1}{V_{t_i^{\ell-1}}}|\right) > h_\ell^{1/2-\eta}\Big) \leq h_\ell^{m(\eta-1/2)}(\mathbb{E}\Big|\frac{1}{V_{t_i^\ell}} - \frac{1}{V_{t_i^{\ell-1}}}\Big|^m) \leq C h_\ell^{m\eta},$$

where we used Lemma 3.3.3 with choosing $0 < \frac{q}{\eta} < \frac{\kappa\theta}{\sigma^2}$, for h_ℓ sufficiently small. The third extreme path event is $o(h_\ell^q)$ for all $q > 0$ because, for $m \geq 1$, we have

$$\mathbb{P}\left(\int_{t_i^\ell}^{t_{i+1}^\ell} V_s ds > h_\ell^{1-\eta}\right) \leq h_\ell^{-m(1-\eta)}\mathbb{E}\Big[\Big|\int_{t_i^\ell}^{t_{i+1}^\ell} V_s ds\Big|^m\Big].$$

Using Jensen's inequality and Fubini's theorem we get

$$\begin{aligned}
\mathbb{P}\left(\int_{t_i^\ell}^{t_{i+1}^\ell} V_s ds > h_\ell^{1-\eta}\right) &\leq h_\ell^{-m(1-\eta)}\mathbb{E}\Big[h_\ell^{m-1}\int_{t_i^\ell}^{t_{i+1}^\ell} |V_s|^m ds\Big] \\
&\leq h_\ell^{-m(1-\eta)} h_\ell^{m-1}\int_{t_i^\ell}^{t_{i+1}^\ell} \sup_{s\in[t_i^\ell, t_{i+1}^\ell]} \mathbb{E}|V_s|^m ds \\
&\leq h_\ell^{m\eta} \sup_{s\in[t_i^\ell, t_{i+1}^\ell]} \mathbb{E}|V_s|^m.
\end{aligned}$$

Then we deduce the result using Lemma 3.3.1 by choosing m such that $0 < \frac{q}{\eta} < m$, for h_ℓ sufficiently small. The last property also follows from Lemma 3.3.1, since for $m \geq 1$, we have

$$\mathbb{P}\left(\sup_{t\in[t_i^\ell, t_{i+1}^\ell]} |\int_{t_i^\ell}^t \sqrt{V_s} dW_s^s| > h_\ell^{\frac{1}{2}-\eta}\right) \leq h_\ell^{-m(\frac{1}{2}-\eta)}\mathbb{E}\Big[\sup_{t\in[t_i^\ell, t_{i+1}^\ell]} |\int_{t_i^\ell}^t \sqrt{V_s} dW_s^s|^m\Big].$$

By the Burkholder-Davis-Gundy inequality, Jensen inequality and Fubini's theorem we get

$$\begin{aligned}\mathbb{P}\left(\sup_{t\in[t_i^\ell,t_{i+1}^\ell]}|\int_{t_i^\ell}^{t}\sqrt{V_s}dW_s^s|>h_\ell^{\frac{1}{2}-\eta}\right)&\leq Ch_\ell^{-m(\frac{1}{2}-\eta)}\mathbb{E}\Big|\int_{t_i^\ell}^{t_{i+1}^\ell}V_sds\Big|^{\frac{m}{2}}\\&\leq Ch_\ell^{-m(\frac{1}{2}-\eta)}h_\ell^{\frac{m}{2}-1}\mathbb{E}\Big(\int_{t_i^\ell}^{t_{i+1}^\ell}|V_s|^{\frac{m}{2}}ds\Big)\\&\leq Ch_\ell^{m\eta-1}\int_{t_i^\ell}^{t_{i+1}^\ell}\mathbb{E}|V_s|^{\frac{m}{2}}ds\leq Ch_\ell^{m\eta}\sup_{s\in[t_i^\ell,t_{i+1}^\ell]}\mathbb{E}|V_s|^{\frac{m}{2}},\end{aligned}$$

where C is a positive constant. We complete the proof using Lemma 3.3.1. □

3.5.2 MLMC variance and complexity analysis

In what follows let us assume that there exists a positive constant $a \geq \frac{1}{2}$ such that $\kappa\theta/\sigma^2 > a$ so that the Feller's condition is satisfied. In what follows we analyse the variance of the improved MLMC method given in (3.13) for pricing a D-O barrier option. To do this, we need the following assumptions.
Let f denote a payoff function satisfying:
$\exists C_1 > 0$ s.t. $\forall t \geq 0, x, y \in \mathbb{R}$,

$$|f(x) - f(y)| \leq C_1|x - y|, \tag{3.23}$$

and $\exists C_2 > 0$ s.t. $\forall x \in \mathbb{R}$,

$$|f(x)| \leq C_2(1 + |x|). \tag{3.24}$$

First event A_1. we consider any of the extreme path events given in Lemma 3.5.1 and 3.5.2 that is satisfied. For $\gamma > 0$, we use Hölder's inequality to get

$$\begin{aligned}\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2\mathbb{1}_{A_1}] &\leq \mathbb{E}^{\frac{\gamma}{1+\gamma}}\Big[|\overline{Q}_\ell^f - \overline{Q}_\ell^c|^{\frac{2(1+\gamma)}{\gamma}}\Big]\mathbb{E}^{\frac{1}{1+\gamma}}[\mathbb{1}_{A_1}]\\&\leq 2^{\frac{2+\gamma}{1+\gamma}}\left(\mathbb{E}^{\frac{\gamma}{1+\gamma}}[|\overline{Q}_\ell^f|^{\frac{2(1+\gamma)}{\gamma}}] + \mathbb{E}^{\frac{\gamma}{1+\gamma}}[|\overline{Q}_\ell^c|^{\frac{2(1+\gamma)}{\gamma}}]\right)\Big(\mathbb{P}[A_1]\Big)^{\frac{1}{1+\gamma}}.\end{aligned}$$

As the payoff function f is Lipschitz continuous, we use Theorem 3.4.2, Lemma 3.5.1 and Lemma 3.5.2 to get that

$$\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2\mathbb{1}_{A_1}] = o(h_\ell^{\frac{q}{1+\gamma}}) \text{ for all } q \text{ such that } 0 < \frac{q}{\eta} \leq a. \tag{3.25}$$

Second event A_2. This event corresponds to non-extreme paths satisfying

$$|\inf_{t\in[0,T]} X_t - D| > {h_\ell}^{\frac{1}{2}-\eta(1+\epsilon)} \text{ for } \eta \in (0, 1/2(1+\varepsilon)) \text{ with } \varepsilon > 0.$$

Let us assume that $\inf_{t\in[0,T]} X_t = X_\tau$ with $\tau \in [t_i^\ell, t_{i+1}^\ell]$ for a given $i \in \{0, \ldots, 2^\ell\}$.

- First case: for $X_\tau < D - {h_\ell}^{\frac{1}{2}-\eta(1+\varepsilon)}$, we write

$$\begin{aligned}|\overline{X}_{t_i^\ell}^{2^\ell} - X_\tau| &\leq |\overline{X}_{t_i^\ell}^{2^\ell} - X_{t_i^\ell}| + |X_{t_i^\ell} - X_\tau|\\&\leq |\overline{X}_{t_i^\ell}^{2^\ell} - X_{t_i^\ell}| + rh_\ell + \frac{1}{2}\int_{t_i^\ell}^{t_{i+1}^\ell}V_sds + \sup_{t\in[t_i^\ell,t_{i+1}^\ell]}\left|\int_{t_i^\ell}^{t}\sqrt{V_s}dW_s^s\right|.\end{aligned}$$

Since we are in the case of non-extreme paths events (see Lemma 3.5.1 and Lemma 3.5.2), then for h_ℓ sufficiently small we have that $|\overline{X}_{t_i^\ell}^{2^\ell} - X_\tau| = O({h_\ell}^{\frac{1}{2}-\eta})$ and then $|\overline{X}_{t_i^\ell}^{2^\ell} - X_\tau| < {h_\ell}^{\frac{1}{2}-\eta(1+\varepsilon)}$, which yields that $\overline{X}_{t_i^\ell}^{2^\ell} < D$. For $\overline{X}_{t_i^\ell}^{2^{\ell-1}}$, we have

$$|\overline{X}_{t_i^\ell}^{2^{\ell-1}} - X_\tau| \leq |\overline{X}_{t_i^\ell}^{2^\ell} - \overline{X}_{t_i^\ell}^{2^{\ell-1}}| + |\overline{X}_{t_i^\ell}^{2^\ell} - X_\tau|.$$

Then, as by Lemma 3.5.1 we have $|\overline{X}_{t_i^\ell}^{2^\ell} - \overline{X}_{t_i^\ell}^{2^{\ell-1}}| = O\left(h_\ell{}^{\frac{1}{2}-\eta}\right)$ and since we have proved that $|\overline{X}_{t_i^\ell}^{2^\ell} - X_\tau| = O(h_\ell{}^{\frac{1}{2}-\eta})$, we conclude that $|\overline{X}_{t_i^\ell}^{2^{\ell-1}} - X_\tau| < h_\ell{}^{\frac{1}{2}-\eta(1+\varepsilon)}$, which yields that $\overline{X}_{t_i^\ell}^{2^{\ell-1}} < D$, for a sufficiently small h_ℓ. Hence, we get $\overline{Q}_\ell^f - \overline{Q}_\ell^c = 0$.

• Second case: for $X_\tau > D + h_\ell{}^{\frac{1}{2}-\eta(1+\varepsilon)}$, we proceed in the same way and we check by standard calculations that for h_ℓ sufficiently small $\prod_{i=0}^{2^\ell-1}(1-\overline{q}_i^{2^\ell})$ and $\prod_{i=0}^{2^\ell-1}(1-\overline{q}_i^{2^{\ell-1}})$ are both equal to $1+o(h_\ell{}^a)$ for all $a>0$. Consequently, as the payoff function f is Lipschitz continuous and as we work on the non-extreme paths events, we deduce that $\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_2}] = O(h_\ell{}^{1-2\eta})$.

Third event A_3. This last event corresponds to the rest of the non-extreme path events. We have

$$\begin{aligned}\overline{Q}_\ell^f - \overline{Q}_\ell^c &= f(\overline{X}_T^{2^\ell}) \prod_{i=0}^{2^\ell-1}(1-\overline{q}_i^{2^\ell}) - f(\overline{X}_T^{2^{\ell-1}}) \prod_{i=0}^{2^\ell-1}(1-\overline{q}_i^{2^{\ell-1}}) \\ &= \Big(f(\overline{X}_T^{2^\ell}) - f(\overline{X}_T^{2^{\ell-1}})\Big) \prod_{i=0}^{2^\ell-1}(1-\overline{q}_i^{2^\ell}) + f(\overline{X}_T^{2^{\ell-1}})\Big(\prod_{i=0}^{2^\ell-1}(1-\overline{q}_i^{2^\ell} - \prod_{i=0}^{2^\ell-1}(1-\overline{q}_i^{2^{\ell-1}}\Big)\end{aligned}$$

Since, we still work on the non-extreme paths events (see Lemma 2.3.2 for the extreme paths events), we deduce using the Lipschitz and the linear growth properties of the payoff function f that

$$(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \le \Big(C_1|\overline{X}_T^{2^\ell} - \overline{X}_T^{2^{\ell-1}}| + 2C_2(1+|\overline{X}_T^{2^{\ell-1}}|)\Big)^2 \tag{3.26}$$

$$\le C h_\ell^{-2\eta}, \tag{3.27}$$

where C is a standard positive constant that may change form line to line. Now, we can write

$$\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_3}] \le C h_\ell^{-2\eta} \mathbb{P}(|\inf_{t\in[0,T]} X_t - D| \le h_\ell{}^{\frac{1}{2}-\eta(1+\varepsilon)}).$$

Thus, provided that $\inf_{t\in[0,T]} X_t$ admits a bounded density (see the discussion below), we get

$$\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_3}] = O(h_\ell{}^{\frac{1}{2}-3\eta(1+\varepsilon)}).$$

To complete the proof, we choose

$$\eta = \frac{\delta}{3(1+\varepsilon)},$$

which yields $\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_3}] = O(h^{\frac{1}{2}-\delta})$. Now concerning the second event and for the above choice of η, we easily see that $1-2\eta > \frac{1}{2} - \delta > 0$, which yields $\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_2}] = o(h^{\frac{1}{2}-\delta})$. Finally, for the first event, we choose $q = (1+\gamma)(\frac{1}{2}-\delta)$ to guarantee that $\mathbb{E}[(\overline{Q}_\ell^f - \overline{Q}_\ell^c)^2 \mathbb{1}_{A_1}] = O(h^{\frac{1}{2}-\delta})$ which is satisfied as soon as

$$0 < \frac{q}{\eta} = \frac{3(1+\varepsilon)(1+\gamma)(\frac{1}{2}-\delta)}{\delta} \le a.$$

As γ and ε have to be chosen very close to zero, then the above condition remains valid for δ satisfying

$$\frac{3}{2(a+3)} \le \delta < \frac{1}{2}, \quad \text{with} \quad a > \frac{1}{2}.$$

and therefore for h_ℓ sufficiently small, we get

$$\mathrm{Var}\big(\overline{Q}_\ell^f - \overline{Q}_\ell^c\big) = O(h_\ell^{\frac{1}{2}-\delta}) \text{ for } \delta \in (\frac{3}{2(a+3)}, \frac{1}{2}), \quad \text{with} \quad a > \frac{1}{2}.$$

Of course, if we want to achieve the smallest possible variance, we have to take δ close to zero which suggests to have more restrictions on our parameters by choosing $\kappa\theta/\sigma^2 > a$ for a large $a > \frac{1}{2}$. In our

numerical tests, we observed that $\mathrm{Var}\big(\overline{Q}_\ell^f - \overline{Q}_\ell^c\big)$ is of order h^β where β is close to 1/2 with a less restricted range of parameters satisfying $\kappa\theta/\sigma^2 > \frac{1}{2}$ (see section 3.7). Thus, the complexity theorem 1.3.1 applies and the time complexity of the MLMC method that achives a precision of order ε is of order $\varepsilon^{-2-(1-\beta)/\alpha}$.

It is worth noticing that in our analysis of the event A_3, we used that

$$\left|\prod_{i=0}^{2^\ell-1}(1-\overline{q}_i^{2^\ell}) - \prod_{i=0}^{2^\ell-1}(1-\overline{q}_i^{2^{\ell-1}})\right| = O(1).$$

The reason is that the main contribution in the above error comes from the indices for which none of $\overline{X}^{2^\ell}_{t_i^\ell}, \overline{X}^{2^\ell}_{t_{i+1}^\ell}, \overline{X}^{2^{\ell-1}}_{t_i^\ell}, \overline{X}^{2^{\ell-1}}_{t_{i+1}^\ell}$. Let us call R the set of all these indices. For $i \in R$ we have

$$\big|\log\big(\overline{q}_i^{2^\ell}\big) - \log\big(\overline{q}_i^{2^{\ell-1}}\big)\big| = \frac{2}{h_\ell(1-\rho^2)}\left|\frac{1}{V_{t_i^\ell}}(\overline{X}^{2^\ell}_{t_i^\ell}-D)_+(\overline{X}^{2^\ell}_{t_{i+1}^\ell}-D)_+ - \frac{1}{V_{t_i^{\ell-1}}}(\overline{X}^{2^{\ell-1}}_{t_i^\ell}-D)_+(\overline{X}^{2^{\ell-1}}_{t_{i+1}^\ell}-D)_+\right|.$$

Now, applying twice the relation $f_1g_1 - f_2g_2 = \frac{1}{2}(f_1-f_2)(g_1+g_2) + \frac{1}{2}(f_1+f_2)(g_1-g_2)$ we get

$$\begin{aligned}
\big|\log\big(\overline{q}_i^{2^\ell}\big) - \log\big(\overline{q}_i^{2^{\ell-1}}\big)\big| \le{}& \frac{1}{h_\ell(1-\rho^2)}\times \\
&\Big[\Big|\Big(\frac{1}{V_{t_i^\ell}} - \frac{1}{V_{t_i^{\ell-1}}}\Big)\Big((\overline{X}^{2^\ell}_{t_i^\ell}-D)_+(\overline{X}^{2^\ell}_{t_{i+1}^\ell}-D)_+ + (\overline{X}^{2^{\ell-1}}_{t_i^\ell}-D)_+(\overline{X}^{2^{\ell-1}}_{t_{i+1}^\ell}-D)_+\Big)\Big| \\
&+\Big|\Big(\frac{1}{V_{t_i^\ell}} + \frac{1}{V_{t_i^{\ell-1}}}\Big)\Big((\overline{X}^{2^\ell}_{t_i^\ell}-D)_+(\overline{X}^{2^\ell}_{t_{i+1}^\ell}-D)_+ - (\overline{X}^{2^{\ell-1}}_{t_i^\ell}-D)_+(\overline{X}^{2^{\ell-1}}_{t_{i+1}^\ell}-D)_+\Big)\Big|\Big] \\
\le{}& \frac{1}{h_\ell(1-\rho^2)}\Big[\Big|\Big(\frac{1}{V_{t_i^\ell}} - \frac{1}{V_{t_i^{\ell-1}}}\Big)\Big((\overline{X}^{2^\ell}_{t_i^\ell}-D)_+(\overline{X}^{2^\ell}_{t_{i+1}^\ell}-D)_+ + (\overline{X}^{2^{\ell-1}}_{t_i^\ell}-D)_+(\overline{X}^{2^{\ell-1}}_{t_{i+1}^\ell}-D)_+\Big)\Big| \\
&+\Big|\frac{1}{V_{t_i^\ell}} + \frac{1}{V_{t_i^{\ell-1}}}\Big| \\
&\times\Big|\frac{1}{2}\big((\overline{X}^{2^\ell}_{t_i^\ell}-D)_+ - (\overline{X}^{2^{\ell-1}}_{t_i^\ell}-D)_+\big)\big((\overline{X}^{2^\ell}_{t_{i+1}^\ell}-D)_+ + (\overline{X}^{2^{\ell+1}}_{t_{i+1}^\ell}-D)_+\big) + \frac{1}{2}\big((\overline{X}^{2^\ell}_{t_i^\ell}-D)_+ \\
&+(\overline{X}^{2^{\ell-1}}_{t_i^\ell}-D)_+\big)\big((\overline{X}^{2^\ell}_{t_{i+1}^\ell}-D)_+ - (\overline{X}^{2^{\ell-1}}_{t_{i+1}^\ell}-D)_+\big)\Big|\Big].
\end{aligned}$$

By the Lipschitz property of the map $x \mapsto (x-\mathcal{D})_+$, there exists a positive constant C such that

$$\begin{aligned}
&\big|\log\big(\overline{q}_i^{2^\ell}\big) - \log\big(\overline{q}_i^{2^{\ell-1}}\big)\big| \\
&\le \frac{1}{h_\ell(1-\rho^2)}\Big[2h_\ell^{1-2\eta(1+\varepsilon)}\Big|\frac{1}{V_{t_i^\ell}} - \frac{1}{V_{t_i^{\ell-1}}}\Big| \\
&+\frac{1}{2}\Big|\frac{1}{V_{t_i^\ell}} + \frac{1}{V_{t_i^{\ell-1}}}\Big|\Big|(\overline{X}^{2^\ell}_{t_i^\ell} - \overline{X}^{2^{\ell-1}}_{t_i^\ell})h_\ell^{\frac{1}{2}-\eta(1+\varepsilon)} + (\overline{X}^{2^\ell}_{t_{i+1}^\ell} - \overline{X}^{2^{\ell-1}}_{t_{i+1}^\ell})h_\ell^{\frac{1}{2}-\eta(1+\varepsilon)}\Big|\Big] \\
&\le \frac{1}{h_\ell(1-\rho^2)}\Big[\Big|\frac{1}{V_{t_i^\ell}} - \frac{1}{V_{t_i^{\ell-1}}}\Big|2h_\ell^{1-2\eta(1+\varepsilon)} + \Big|\frac{1}{V_{t_i^\ell}} + \frac{1}{V_{t_i^{\ell-1}}}\Big|h_\ell^{1-2\eta(1+\varepsilon)}\Big].
\end{aligned}$$

Since on the non-extreme paths events we have $\Big|\frac{1}{V_{t_i^\ell}} - \frac{1}{V_{t_i^{\ell-1}}}\Big| \le h_\ell^{\frac{1}{2}-\eta}$ and $\Big|\frac{1}{V_{t_i^\ell}}\Big| \le h_\ell^{-\eta}$, then

$$\big|\log\big(\overline{q}_i^{2^\ell}\big) - \log\big(\overline{q}_i^{2^{\ell-1}}\big)\big| \le Ch_\ell^{-\eta(3+\varepsilon)}.$$

So that, for h_ℓ sufficiently small we have $\big|\log\overline{q}_i^{2^\ell} - \log\overline{q}_i^{2^{\ell-1}}\big| < h_\ell{}^{\frac{1}{2}-4\eta(1+\varepsilon)}$. Thus, we have

$$\begin{aligned}
1-\overline{q}_i^{2^{\ell-1}} &= (1-\overline{q}_i^{2^\ell}) + \overline{q}_i^{2^\ell}\big(1-\exp\big(\log\overline{q}_i^{2^{\ell-1}} - \log\overline{q}_i^{2^\ell}\big)\big) \\
&\le (1-\overline{q}_i^{2^\ell}) + \overline{q}_i^{2^\ell}\big(1-\exp\big(-h_\ell^{-\eta(3+\varepsilon)}\big)\big).
\end{aligned}$$

Consequently, we get

$$\prod_{i\in R}(1-\overline{q}_i^{2^{\ell-1}}) \leq \prod_{i\in R}\Big((1-\overline{q}_i^{2^\ell}) + \overline{q}_i^{2^\ell}\big(1-\exp\Big(-h_\ell^{-\eta(3+\varepsilon)}\Big)\big)\Big)$$
$$\leq \prod_{i\in R}(1-\overline{q}_i^{2^\ell}) + \big(1-\exp\Big(-h_\ell^{-\eta(3+\varepsilon)}\Big)\big),$$

Then using similar arguments as in the third event of the proof of Theorem 2.3.3, we get

$$\Big|\prod_{i=0}^{2^\ell-1}(1-\overline{q}_i^{2^\ell}) - \prod_{i=0}^{2^\ell-1}(1-\overline{q}_i^{2^{\ell-1}})\Big| = O(h_\ell^{-\eta(3+\varepsilon)})$$

which is a trivial and not an accurate upperbound. The reason is that strong convergence order of our semi-exact log-Heston scheme is of order $h_\ell^{\frac{1}{2}}$ and not h_ℓ. and we were not able to find in the literature an alternative first order strong convergence scheme for which we can use similar arguments to prove the corresponding MLMC variance for pricing barrier options in the log-Heston model.

Finally, let us mention that unlike the CIR and CEV models (see sections 2.4 and 2.5), finding the probability density of the running maximum/minimum of the log-Heston model is a challenging problem. Many attempts in the literature failed to obtain this result, for more details see discussions in [62] and [17]. In our context we do not need to have an explicit expression of this density but rather show that it is bounded in the neighbourhood of the barrier. This also is a quite challenging task that is left for a future work.

3.6 Antithetic Multilevel for discretization schemes

The aim of this section is to explore the efficiency of the antithetic MLMC approach using the Milstein discretisation to price barrier option with the Brownian bridge technique.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $W = (W^1, ..., W^q)$ be a q-dimensional Brownian motion defined on the probability space. We consider a multi-dimensional SDE solution to

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t \text{ and } X_0 = x \in \mathbb{R}^d,$$

where $b \in C^2(\mathbb{R}^d, \mathbb{R}^d)$, $\sigma \in C^2(\mathbb{R}^d, \mathbb{R}^{d\times q})$. In the context of barrier options, Giles [42] proved that optimal complexity of order $O(\varepsilon^{-2})$ can be achieved using the MLMC method with the first-order Milstein scheme for one-dimensional SDEs with smooth globally Lipschitz coefficients, using the Brownian bridge technique. Although, in multiple dimension we cannot simulate efficiently the Milstein scheme due to the presence of the Levy area induced by the rectangular terms which come out in the second order term of the scheme.

We consider the general Milstein scheme with a time step $h = \frac{T}{n}$

$$\hat{X}_{k,t_{i+1}} = \hat{X}_{k,t_i} + b_k(\hat{X}_{t_i})h + \sum_{j=1}^{q}\sigma_{kj}(\hat{X}_{t_i})\Delta W_{j,i} + \sum_{j,l=1}^{q} H_{kjl}(\hat{X}_{t_i})(\Delta W_{j,i} - \Delta W_{l,i} - \Omega_{jl}h - A_{jl,i})$$

where the tensor H_{kjl} is given by

$$H_{kjl}(x) = \frac{1}{2}\sum_{m=1}^{d}\sigma_{ml}(x)\frac{\partial\sigma_{kj}}{\partial x_l}(x), \quad k = 1, ..., d \text{ and } j, l = 1, ..., q,$$

$(\Omega_{i,j})_{1\leq i,j\leq q}$ denotes the correlation matrix for the driving Brownian paths and $A_{jl,i}$ corresponds to the Lévy area

$$A_{jl,i} = \int_{t_i}^{t_{i+1}}(W_{j,t} - W_{j,t_i})dW_{l,t} - \int_{t_i}^{t_{i+1}}(W_{l,t} - W_{l,t_i})dW_{j,t}.$$

Now, we introduce the truncated Milstein approximation without the Lévy's areas introduced in [59]

$$\hat{X}_{k,t_{i+1}} = \hat{X}_{k,t_i} + b_k(\hat{X}_{t_i})h + \sum_{j=1}^{q} \sigma_{kj}(\hat{X}_{t_i})\Delta W_{j,i} + \sum_{j,l=1}^{q} h_{kjl}(\hat{X}_{t_i})(\Delta W_{j,i} - \Delta W_{l,i} - \Omega_{jl}h).$$

Clark and Cameron [19], and Müller–Gronbach [67] demonstrated that the strong convergence of a Milstein scheme without the Lévy area is only $\frac{1}{2}$. Hence, when combining the MLMC algorithm with the truncated Milstein scheme, it will not achieve the optimal complexity $O(\varepsilon^{-2})$. To overcome this problem, Giles and Szpruch [41] proposed an antithetic MLMC estimator that neglects the Lévy area term. It achieve the optimal complexity without the need for a high order of convergence. The idea of the antithetic approach is to use the same truncated Milstein scheme but to swap each successive pair of Brownian increments in the scheme. This allows a variance decrease to the order h^2 and this leads to a time complexity $O(\varepsilon^{-2})$. However, for exotic options the time complexity order of the antithetic MLMC method may deteriorate. Giles and Szpruch [40] introduced a modified version of the antithetic MLMC method in which they add a term that sub-samples the Lévy areas. They tested numerically this method for pricing barrier and lookback options under the two-dimensional Clarck-Cameron model. The numerical tests suggests that the time complexity in this case is $O(\varepsilon^{-2}(\log \varepsilon)^2)$ see [40].

Now, let us rather focus on the Heston model defined solution to

$$\begin{aligned} dS_t &= rS_t dt + \sqrt{V_t} S_t d\tilde{W}_t^s \\ dV_t &= \kappa(\theta - V_t)dt + \sigma\sqrt{V_t} dW_t^v \end{aligned} \tag{3.28}$$

where $\tilde{W}^s$ and W^v are two correlated Brownian motions with correlation parameter ρ. We use the antithetic method to estimate the D-O and U-O barrier options. So first, we approximate these prices respectively by

$$\hat{\pi}_D := \mathbb{E}\Big[f(\hat{S}_T^n)\prod_{i=0}^{n-1} \mathbb{1}_{\{\inf_{t\in[t_i,t_{i+1}]}\hat{S}_t^n > D\}}\Big] \text{ and } \hat{\pi}_U := \mathbb{E}\Big[f(\hat{S}_T^n)\prod_{i=0}^{n-1} \mathbb{1}_{\{\sup_{t\in[t_i,t_{i+1}]}\hat{S}_t^n < U\}}\Big],$$

where $\hat{S}_{t_i}^n$ is the truncated Milstein scheme with step $h = \frac{T}{n}$

$$\begin{aligned} \hat{S}_{t_{i+1}}^n &= \hat{S}_{t_i}^n\Big(1 + rh + \sqrt{\hat{V}_{t_i}}(\tilde{W}_{t_{i+1}}^s - \tilde{W}_{t_i}^s) + \frac{1}{2}\hat{V}_{t_i}\big((\tilde{W}_{t_{i+1}}^s - \tilde{W}_{t_i}^s)^2 - h\big) \\ &\quad + \frac{1}{4}\sigma\big((\tilde{W}_{t_{i+1}}^s - \tilde{W}_{t_i}^s)(W_{t_{i+1}}^v - W_{t_i}^v) - \rho h\big)\Big), \\ \hat{V}_{t_{i+1}} &= \hat{V}_{t_i} + \kappa(\theta - V_{t_{i+1}})h + \sigma\sqrt{\hat{V}_{t_i}}(W_{t_{i+1}}^v - W_{t_i}^v) + \frac{1}{2}\sigma^4\big((W_{t_{i+1}}^v - W_{t_i}^v)^2 - h\big). \end{aligned}$$

Following [40] let us introduce the following interpolated scheme

$$\hat{S}_t^n = \hat{S}_{t_i}^n + \lambda_t(\hat{S}_{t_{i+1}}^n - \hat{S}_{t_i}^n) + \sqrt{\hat{V}_{t_i}}\hat{S}_{t_i}^n\Big(\tilde{W}_t^s - \tilde{W}_{t_i}^s - \lambda_t(\tilde{W}_{t_{i+1}}^s - \tilde{W}_{t_i}^s)\Big), \tag{3.29}$$

where $\lambda_t = \frac{t-t_i}{h}$. This interpolation allows us to use the Brownian bridge technique to approximate the barrier options prices.

$$\hat{\pi}_U = \mathbb{E}\Big[f(\hat{S}_T^n)\prod_{i=0}^{n-1}(1-\overline{p}_i)\Big], \text{ where } \overline{p}_i = \exp\Big(\frac{-2(U-\hat{S}_{t_i}^n)_+(U-\hat{S}_{t_{i+1}}^n)_+}{(\hat{S}_{t_i}^n)^2\hat{V}_{t_i}h}\Big)$$

and

$$\hat{\pi}_D = \mathbb{E}\Big[f(\hat{S}_T^n)\prod_{i=0}^{n-1}(1-\overline{q}_i)\Big], \text{ where } \overline{q}_i := \exp\Big(\frac{-2(\overline{S}_{t_i}^n - D)_+(\hat{S}_{t_{i+1}}^n - D)_+}{(\hat{S}_{t_i}^n)^2\hat{V}_{t_i}h}\Big).$$

Now, to set the corresponding antithetic MLMC estimator we define the two successive half time steps of the first fine path approximation $\hat{S}^{2^\ell}$ over the fine time step is given by

$$\begin{aligned}
\hat{S}^{2^\ell}_{t^\ell_{i+1}} &= \hat{S}^{2^\ell}_{t^\ell_i}(1 + r(t^\ell_{i+1} - t^\ell_i) + \sqrt{\hat{V}^{2^\ell}_{t^\ell_i}}(\tilde{W}^s_{t^\ell_{i+1}} - \tilde{W}^s_{t^\ell_i}) + \frac{1}{2}\hat{V}^{2^\ell}_{t^\ell_i}((\tilde{W}^s_{t^\ell_{i+1}} - \tilde{W}^s_{t^\ell_i})^2 - (t^\ell_{i+1} - t^\ell_i)) \\
&\quad + \frac{1}{4}\sigma(\tilde{W}^s_{t^\ell_{i+1}} - \tilde{W}^s_{t^\ell_i})(W^v_{t^\ell_{i+1}} - W^v_{t^\ell_i}) \\
\hat{V}^{2^\ell}_{t^\ell_{i+1}} &= \hat{V}^{2^\ell}_{t^\ell_i} + \kappa(\theta - \hat{V}^{2^\ell}_{t^\ell_{i+1}})(t^\ell_{i+1} - t^\ell_i) + \sigma\sqrt{\hat{V}^{2^\ell}_{t^\ell_i}}(W^v_{t^\ell_{i+1}} - W^v_{t^\ell_i}) + \frac{1}{2}\sigma^4\Big((W^v_{t^\ell_{i+1}} - W^v_{t^\ell_i})^2 - (t^\ell_{i+1} - t^\ell_i)\Big).
\end{aligned}$$

The antithetic path is build following [40] with swapping the Brownian increments. This leads to
Then for the fine path, this yields

$$\begin{aligned}
\mathbb{E}\big[\hat{P}_\ell\big] = \mathbb{E}\big[\hat{P}^f_\ell\big] &= \mathbb{E}\Big[f(\hat{S}^{2^\ell}_T)\prod_{i=0}^{2^\ell-1}\mathbb{1}_{\{\sup_{t\in[t^\ell_i,t^\ell_{i+1}]}\hat{S}^{2^\ell}_t<U\}}\Big] \\
&= \mathbb{E}\Big[\mathbb{E}\big[f(\hat{S}^{2^\ell}_T)\prod_{i=0}^{2^\ell-1}\mathbb{1}_{\{\sup_{t\in[t^\ell_i,t^\ell_{i+1}]}\hat{S}^{2^\ell}_t<U\}}|\hat{S}^{2^\ell}_0, \overline{S}^{2^\ell}_{t_1}, \ldots, \hat{S}^{2^\ell}_{t^\ell_n}, \hat{V}^{2^\ell}_{t^\ell_0}, \hat{V}^{2^\ell}_{t^\ell_1}, ..., \hat{V}^{2^\ell}_{t^\ell_n}\big]\Big] \\
&= \mathbb{E}\big[f(\hat{S}^{2^\ell}_T)\prod_{i=0}^{2^{\ell-1}-1}\mathbb{E}\big[\mathbb{1}_{\{\sup_{t\in[t^\ell_i,t^\ell_{i+1}]}\hat{S}^{2^\ell}_t<U\}}|\hat{S}^{2^\ell}_{t^\ell_i}, \hat{S}^{2^\ell}_{t^\ell_{i+1}}, \hat{V}_{t^\ell_i}, \hat{V}_{t^\ell_{i+1}}\big]\big] \\
&= \mathbb{E}\big[f(\hat{S}^{2^\ell}_T)\prod_{i=0}^{2^\ell-1}(1-\overline{p}^{2^\ell}_i)\big]
\end{aligned}$$

where

$$\hat{p}^{2^\ell}_i = \mathbb{P}\Big(\sup_{t\in[t^\ell_i,t^\ell_{i+1}]}\hat{S}^{2^\ell}_t < U|\hat{S}^{2^\ell}_{t^\ell_i}, \hat{S}^{2^\ell}_{t^\ell_{i+1}}, \hat{V}_{t^\ell_i}, \hat{V}_{t^\ell_{i+1}}\Big) = \exp\Big(\frac{-2(U-\hat{S}^{2^\ell}_{t^\ell_i},)_+(U-\hat{S}^{2^\ell}_{t^\ell_{i+1}})_+}{(\hat{S}^{2^\ell}_{t^\ell_{i+1}})^2\hat{V}_{t^\ell_i}h_\ell}\Big).$$

Then, we can write

$$\begin{aligned}
\mathbb{E}\big[\hat{P}_\ell\big] = \mathbb{E}\big[\hat{P}^f_\ell\big], \text{ where } \hat{P}^f_\ell := f(\hat{S}^{2^\ell}_T)\prod_{i=0}^{2^\ell-1}(1-\hat{p}^{2^\ell}_i) \text{ with} \qquad (3.30)\\
\hat{p}^{2^\ell}_i = \exp\Big(\frac{-2(U-\hat{S}^{2^\ell}_{t_i})_+(U-\hat{S}^{2^\ell}_{t_{i+1}})_+}{(1-\rho^2)\hat{V}_{t^\ell_i}h_\ell}\Big).
\end{aligned}$$

Now, we write the coarse approximation of the trajectory $\hat{S}^{2\ell-1}_{t_i}$ with a time step $h = \frac{T}{n}$ neglecting the Lévy area terms over the fine time step $[t^{\ell-1}_i, t^{\ell-1}_{i+1}]$ is defined by

$$\begin{aligned}
\hat{S}^{2^{\ell-1}}_{t^{\ell-1}_{i+1}} &= \hat{S}^{2^{\ell-1}}_{t^{\ell-1}_i}(1 + r(t^{\ell-1}_{i+1} - t^{\ell-1}_i) + \sqrt{V^{2^{\ell-1}}_{t^{\ell-1}_i}}(\tilde{W}^s_{t^{\ell-1}_{i+1}} - \tilde{W}^s_{t^{\ell-1}_i}) + \frac{1}{2}\hat{V}^{2^{\ell-1}}_{t^{\ell-1}_i}((\tilde{W}^s_{t^{\ell-1}_{i+1}} - \tilde{W}^s_{t^{\ell-1}_i})^2 - (t^{\ell-1}_{i+1} - t^{\ell-1}_i)) \\
&\quad + \frac{1}{4}\sigma(\tilde{W}^s_{t^{\ell-1}_{i+1}} - \tilde{W}^s_{t^{\ell-1}_i})(W^v_{t^{\ell-1}_{i+1}} - W^v_{t^{\ell-1}_i}) \\
\hat{V}^{2^{\ell-1}}_{t^{\ell-1}_{i+1}} &= \hat{V}^{2^{\ell-1}}_{t^{\ell-1}_i} + \kappa(\theta - \hat{V}^{2^{\ell-1}}_{t^{\ell-1}_{i+1}})(t^{\ell-1}_{i+1} - t^{\ell-1}_i) + \sigma\sqrt{\hat{V}^{2\ell-1}_{t^{\ell-1}_i}}(W^v_{t^{\ell-1}_{i+1}} - W^v_{t^{\ell-1}_i}) \\
&\quad + \frac{1}{2}\sigma^4\Big((W^v_{t^{\ell-1}_{i+1}} - W^v_{t^{\ell-1}_i})^2 - (t^{\ell-1}_{i+1} - t^{\ell-1}_i)\Big).
\end{aligned}$$

Then, we obtain

$$\mathbb{E}\big[\hat{P}_{\ell-1}\big] = \mathbb{E}\Big[f(\hat{S}_T^{2^{\ell-1}}) \prod_{i=0}^{2^{\ell-1}-1} \mathbb{1}_{\{\sup_{t\in[t_i^{\ell-1}, t_{i+1}^{\ell-1}]} \hat{S}_t^{2^{\ell-1}} < U\}}\Big]$$
$$= \mathbb{E}\Big[\mathbb{E}\big[f(\hat{S}_T^{2^{\ell-1}}) \prod_{i=0}^{2^{\ell-1}-1} \mathbb{1}_{\{\sup_{t\in[t_i^{\ell-1}, t_{i+1}^{\ell-1}]} \hat{S}_t^{2^{\ell-1}} < U\}} | \hat{S}_{t_i^{\ell-1}}^{2^{\ell-1}}, \hat{S}_{t_{2i+1}^{\ell}}^{2^{\ell-1}}, \hat{S}_{t_{i+1}^{\ell-1}}^{2^{\ell-1}}, \hat{V}_{t_i^{\ell-1}}, \hat{V}_{t_{i+1}^{\ell-1}}, \hat{V}_{t_{2i+1}^{\ell-1}}\big]\Big]$$
$$= \mathbb{E}\big[f(\hat{S}_T^{2^{\ell-1}}) \prod_{i=0}^{2^{\ell-1}-1} \mathbb{E}\big[\mathbb{1}_{\{\sup_{t\in[t_i^{\ell-1}, t_{2i+1}^{\ell}]} \hat{S}_t^{2^{\ell-1}} < U\}} \mathbb{1}_{\{\sup_{t\in[t_{2i+1}^{\ell}, t_{i+1}^{\ell-1}]} \hat{S}_t^{2^{\ell-1}} < U\}} | \hat{S}_{t_i^{\ell-1}}^{2^{\ell-1}}, \hat{S}_{t_{2i+1}^{\ell}}^{2^{\ell-1}}, \hat{S}_{t_{i+1}^{\ell-1}}^{2^{\ell-1}}, \hat{V}_{t_i^{\ell-1}}, \hat{V}_{t_{i+1}^{\ell-1}}, \hat{V}_{t_{2i+1}^{\ell-1}}\big]\big].$$

where the coarse scheme $\hat{S}_{t_{2i+1}^{\ell}}^{2^{\ell-1}}$ is computed using the interpolation scheme (3.29) that is

$$\hat{S}_{t_{2i+1}^{\ell-1}}^{2^{\ell-1}} = \hat{S}_{t_i^{\ell-1}}^{2^{\ell-1}} + \frac{(t_{2i+1}^{\ell} - t_i^{\ell-1}))}{h}(\hat{S}_{t_{2i+1}^{\ell-1}}^{2^{\ell-1}} - \hat{S}_{t_i^{\ell-1}}^{2^{\ell-1}}) + \sqrt{\hat{V}_{t_i^{\ell-1}}^{2^{\ell-1}}}\hat{S}_{t_i^{\ell-1}}^{2^{\ell-1}}\big((\tilde{W}_{t_{2i+1}^{\ell}}^{s} - \tilde{W}_{t_i^{\ell}}^{s}) - \frac{(t_{2i+1}^{\ell} - t_i^{\ell-1})}{h}(\tilde{W}_{t_{i+1}^{\ell-1}}^{s} - \tilde{W}_{t_i^{\ell-1}}^{s})\big)$$
$$\hat{V}_{t_{2i+1}^{\ell}}^{2^{\ell-1}} = \hat{V}_{t_i^{\ell}}^{2^{\ell-1}} + \kappa(\theta - \hat{V}_{t_{2i+1}^{\ell}}^{2^{\ell-1}})(t_{2i+1}^{\ell} - t_i^{\ell}) + \sigma\sqrt{\hat{V}_{t_i^{\ell}}^{2^{\ell-1}}}(W_{t_{2i+1}^{\ell}}^{v} - W_{t_i^{\ell}}^{v}) + \frac{1}{2}\sigma^4\Big((W_{t_{2i+1}^{\ell}}^{v} - W_{t_i^{\ell}}^{v})^2 - (t_{2i+1}^{\ell} - t_i^{\ell})\Big).$$

Then, by the independence of the Brownian increments and using the same arguments as in the proof of proposition 3.4.1, we have

$$\mathbb{E}\big[\hat{P}_{\ell-1}\big] = \mathbb{E}\big[\hat{P}_{\ell-1}^{c}\big], \text{ where } \hat{P}_{\ell-1}^{c} := f(\hat{S}_T^{2^{\ell-1}}) \prod_{i=0}^{2^{\ell-1}-1} (1 - \hat{p}_{i,1}^{2^{\ell-1}})(1 - \hat{p}_{i,2}^{2^{\ell-1}}) \text{ with} \tag{3.31}$$
$$\hat{p}_{i,1}^{2^{\ell-1}} = \exp\Big(\frac{-2(U - \hat{S}_{t_i^{\ell-1}}^{2^{\ell-1}})_+(U - \hat{S}_{t_{2i+1}^{\ell}}^{2^{\ell-1}})_+}{(\hat{S}_{t_i^{\ell}}^{2^{\ell-1}})^2 \hat{V}_{t_i^{\ell-1}} h_\ell}\Big),$$
$$\hat{p}_{i,2}^{2^{\ell-1}} = \exp\Big(\frac{-2(U - \hat{S}_{t_{2i+1}^{\ell}}^{2^{\ell-1}})_+(U - \hat{S}_{t_{i+1}^{\ell-1}}^{2^{\ell-1}})_+}{(\hat{S}_{t_i^{\ell-1}}^{2^{\ell-1}})^2 \hat{V}_{t_i^{\ell-1}} h_\ell}\Big),$$

which can be rewritten as

$$\hat{P}_{\ell-1}^{c} := f(\hat{S}_T^{2^{\ell-1}}) \prod_{i=0}^{2^{\ell}-1} (1 - \hat{p}_i^{2^{\ell-1}}) \text{ with } \hat{p}_i^{2^{\ell-1}} = \exp\Big(\frac{-2(U - \hat{S}_{t_i^{\ell}}^{2^{\ell-1}})_+(U - \hat{S}_{t_{i+1}^{\ell}}^{2^{\ell-1}})_+}{(\hat{S}_{t_i^{\ell-1}}^{2^{\ell-1}})^2 \hat{V}_{t_i^{\ell-1}} h_\ell}\Big), \tag{3.32}$$

Thus, the improved MLMC method approximates $\hat{\pi}_U$ by

$$\hat{P}_U := \frac{1}{N_0}\sum_{k=1}^{N_0} \hat{P}_{0,k}^{f} + \sum_{\ell=1}^{L} \frac{1}{N_\ell}\sum_{k=1}^{N_\ell} \Big(\frac{1}{2}(\hat{P}_{\ell,k}^{f} + \hat{P}_{\ell,k}^{a}) - \hat{P}_{\ell-1,k}^{c}\Big), \tag{3.33}$$

where, $(\hat{P}_{\ell,k}^{f})_{1\le k\le N_\ell}$ and $(\hat{P}_{\ell-1,k}^{c})_{1\le k\le N_\ell}$ are respectively independent copies of $\hat{P}_\ell^f$ and $\hat{P}_{\ell-1}^c$ given by (3.30) and (3.32). Similarly, the improved MLMC method approximates $\bar{\pi}_D$ by

$$\hat{Q}_D := \frac{1}{N_0}\sum_{k=1}^{N_0} \hat{Q}_{0,k}^{f} + \sum_{\ell=1}^{L} \frac{1}{N_\ell}\sum_{k=1}^{N_\ell} \Big(\frac{1}{2}(\overline{Q}_{\ell,k}^{f} + \hat{Q}_{\ell,k}^{a}) - \hat{Q}_{\ell-1,k}^{c}\Big), \tag{3.34}$$

where $(\hat{Q}_{\ell,k}^{f})_{1\le k\le N_\ell}$ and $(\hat{Q}_{\ell-1,k}^{c})_{1\le k\le N_\ell}$ are respectively independent copies of $\hat{Q}_\ell^f$ and $\hat{Q}_{\ell-1}^c$ given by

$$\hat{Q}_\ell^f := f(\hat{S}_T^{2^\ell}) \prod_{i=0}^{2^\ell-1} (1 - \hat{q}_i^{2^\ell}) \text{ with } \hat{q}_i^{2^\ell} = \exp\Big(\frac{-2(\hat{S}_{t_i^{\ell}}^{2^\ell} - D)_+(\hat{S}_{t_{i+1}^{\ell}}^{2^\ell} - D)_+}{(1-\rho^2)\hat{V}_{t_i^{\ell}} h_\ell}\Big)$$
$$\hat{Q}_{\ell-1}^c := f(\hat{S}_T^{2^{\ell-1}}) \prod_{i=0}^{2^\ell-1} (1 - \hat{q}_i^{2^{\ell-1}}) \text{ with } \hat{q}_i^{2^{\ell-1}} = \exp\Big(\frac{-2(\hat{S}_{t_i^{\ell}}^{2^{\ell-1}} - D)_+(\hat{S}_{t_{i+1}^{\ell}}^{2^{\ell-1}} - D)_+}{(1-\rho^2)\hat{V}_{t_i^{\ell-1}} h_\ell}\Big).$$

The aim of using the antithetic method is to compare it numerically with our MLMC semi-exact Heston approximation. It is worth noticing that for the antithetic MLMC scheme the arguments to control the associated variance order does not work since we do not have control of the negative moments for the drift implicit Milstein scheme $\hat{V}$.

3.7 Numerical results

In this section, we consider a D-O Call option with strike $K = 100$ and barrier $B = 75$. We suppose that the stock price is following a Heston model (3.1) with parameters $S_0 = 100$, $V_0 = 0.1$, $\sigma = 0.1$, $\theta = 0.1$, $\kappa = 2$, $\rho = -0.5$ and $r = -0.02$. In practice, $\overline{\pi}_D$ is computed using formula (3.8). In the following table, we obtain an estimation of the time complexity of the MLMC and Antithetic MLMC algorithm.

Estimation	α	β	MLMC Complexity
MLMC Semi Exact	1.173726	0.962958	$\varepsilon^{2.03}$
Antithetic MLMC	0.724667	0.662456	$\varepsilon^{2.47}$

Table 3.1: Time complexity of MLMC and antithetic MLMC methods

Figure 3.1 shows the behaviour of the sample size N_ℓ for several precisions and for both the MLMC and the AMLMC method.

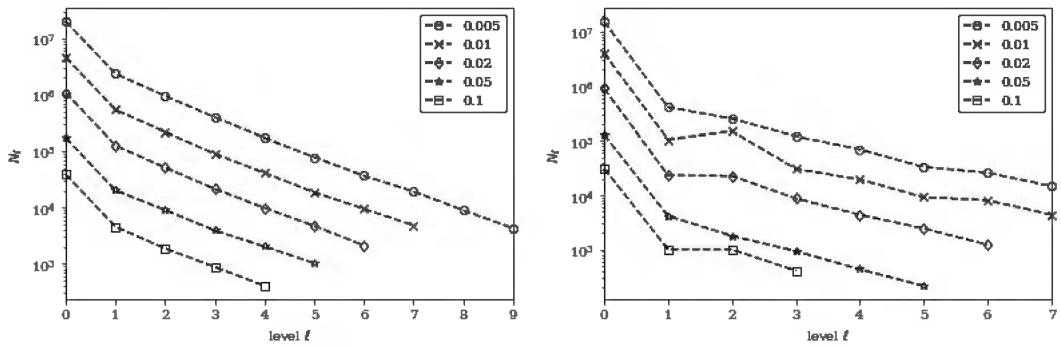


Fig 3.1. Sample size N_ℓ for several process

Figure 3.2 illustrates the computational complexity C which is of order $\epsilon^{-2-(1-\beta)/\alpha}$. Here, we recall that the computational complexity C is defined as $C = \sum_{\ell=0}^{L} 2^\ell N_\ell$. We observe that the map $\epsilon \mapsto \epsilon^{2+(1-\beta)/\alpha} C$ is constant. It turns out that our MLMC algorithm achieves a better performance than the AMLMC method.

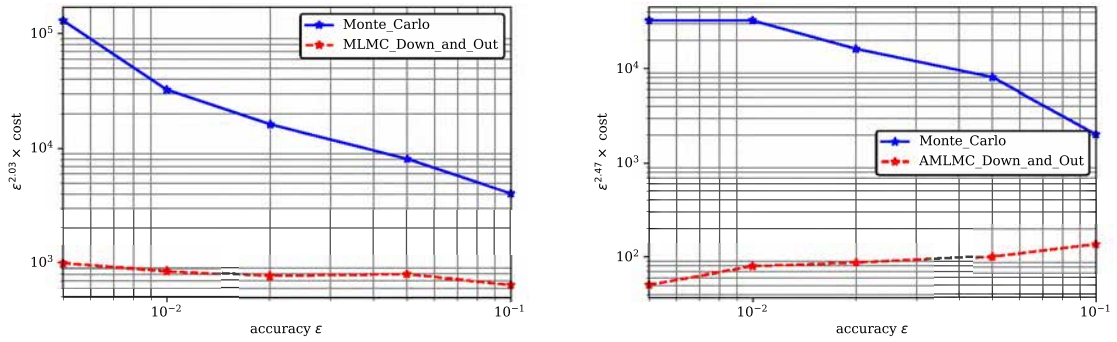


Fig 3.2. Comparison for the performances of MLMC vs classical MC algorithm under the Heston model

3.8 Conclusions and future work

In this work, we devise a numerical scheme for the approximation of the log-Heston model for which we have been able to analyse the associated MLMC method that approximates barrier option prices using the Brownian bridge technique. We proved that the obtained MLMC method has a variance of order $O(h_\ell^\beta)$ with $\beta < \frac{1}{2}$. A possible way to improve this rate, is to combine our MLMC method with the Richardson extrapolation technique as proposed by Lemaire and Pagès [60] for MLMC methods. This is left to a future work.

3.9 Appendix

In order to prove the boundedness of the Heston model moments, we need to examine the stability for the moments of order higher than 1.

Lemma 3.9.1 (See [23, Proposition 3.5] for a proof.)**.** *For $\lambda > 0$ we define the stochastic process*

$$\theta_t = \exp\{\lambda \int_0^t V_u du\}, \quad t > 0. \tag{3.35}$$

If $T < T^*$, then the first moment of θ_T is bounded, i.e.,

$$\mathbb{E}[\theta_T] < \infty, \tag{3.36}$$

where T^* is given by:

- *for $\kappa^2 < 2\lambda\sigma^2$,*

$$T^* = \frac{2}{\sqrt{2\lambda\sigma^2 - \kappa^2}} \Big[\frac{\pi}{2} + \arctan\Big(\frac{\kappa}{\sqrt{2\lambda\sigma^2 - \kappa^2}}\Big)\Big], \tag{3.37}$$

- *for $\kappa^2 > 2\lambda\sigma^2$,*

$$T^* = \infty. \tag{3.38}$$

Using the previous lemma, it is possible to obtain the integrability of a random variable of the form

$$\exp\left(\lambda_1 \int_0^t V_u du + \lambda_2 \int_0^t \sqrt{V_u} dW_u^v + \lambda_3 \int_0^t \sqrt{V_u} dW_u^s\right),$$

where W^s is a Brownian motion independent from the CIR process V. Let $(S_t)_{t\geq 0}$ be the solution of the stochastic differential equation (3.1)

$$S_t = S_0 \exp\Big(rt - \frac{1}{2}\int_0^t V_u du + \sqrt{1-\rho^2}\int_0^t \sqrt{V_u} dW_u^s + \rho\int_0^t \sqrt{V_u} dW_u^v\Big).$$

The next result proves that the asset price moments of order higher than 1 can be obtained.

Theorem 3.9.2. *Let $\alpha \geq 1$. If $T < T_\alpha^*$, then there exists $w_1 > \alpha$ such that, for all $w \in [1, w_1)$, the following holds:*

$$\sup_{t\in[0,T]} \mathbb{E}[S_t^w] < \infty, \tag{3.39}$$

where T_α^* is given by:

- *for $\kappa < \varphi(\alpha)\sigma$,*

$$T_\alpha^* = \frac{2}{\sqrt{\varphi(\alpha)^2\sigma^2 - \kappa^2}}\Big[\frac{\pi}{2} + \arctan\Big(\frac{\kappa}{\sqrt{\varphi(\alpha)^2\sigma^2 - \kappa^2}}\Big)\Big], \tag{3.40}$$

  where

$$\varphi(\alpha) = \alpha|\rho| + \sqrt{(\alpha-1)\alpha}$$

- *for $\kappa > \varphi(\alpha)\sigma$*

$$T_\alpha^* = \infty. \tag{3.41}$$

Proof. The proof is based on similar arguments as in [23]. Conditioning on the σ-algebra $\mathcal{G}_T^v = \sigma(W_t^v; 0 \leq t \leq T)$, we get

$$\mathbb{E}[S_t^w] = S_0^w \mathbb{E}\Big[\exp\Big(rwt + [\frac{1}{2}(w(w-1)) - \frac{1}{2}w^2\rho^2]\int_0^t V_u du + w\rho\int_0^t \sqrt{V_u} dW_u^v\Big)\Big] \tag{3.42}$$

Now suppose that $T < T_\alpha^*$, with T_α^* defined in (3.40) - (3.41). First, if $\kappa < \varphi(\alpha)\sigma$ since φ is increasing on $[1,\infty)$ and by a continuity argument, we can find $w_1 > \alpha$ such that, for all $w \in (w_1, \alpha)$,

$$\kappa < \varphi(w)\sigma \;\text{ and }\; T < \frac{2}{\sqrt{\varphi(w)^2\sigma^2 - \kappa^2}}\left[\frac{\pi}{2} + \arctan\Big(\frac{\kappa}{\sqrt{\varphi(w)^2\sigma^2 - \kappa^2}}\Big)\right] \tag{3.43}$$

Second, if $\kappa = \varphi(\alpha)\sigma$ since φ is strictly increasing on $[1,\infty)$ and

$$\lim_{w\longrightarrow\alpha+} \frac{2}{\sqrt{\varphi(w)^2\sigma^2 - \kappa^2}}\left[\frac{\pi}{2} + \arctan\Big(\frac{\kappa}{\sqrt{\varphi(w)^2\sigma^2 - \kappa^2}}\Big)\right] = \infty \tag{3.44}$$

we can find $w_1 > \alpha$ such that (3.43) holds for all $w \in (\alpha, w_1)$. Finally, if $\kappa > \varphi(\alpha)\sigma$, by a continuity argument, we can find $w_1 > \alpha$ such that, for all $w \in (\alpha, w_1)$, $\kappa > \varphi(w)\sigma$.

For $p \geq 1$, we can rewrite (3.42) as

$$\begin{aligned}
\mathbb{E}[S_t^w] &= S_0^w\, \mathrm{e}^{rwt}\, \mathbb{E}\left[\exp\left(\frac{1}{p}\left(p\Big[\frac{1}{2}\big(w(w-1)\big) - \frac{1}{2}w^2\rho^2\Big]\int_0^t V_u du + pw\rho\int_0^t \sqrt{V_u} dW_u^v\right)\right)\right] \\
&= S_0^w\, \mathrm{e}^{rwt}\, \mathbb{E}\Bigg[\exp\Bigg(\frac{1}{p}\Big(p\Big[\frac{1}{2}\big(w(w-1)\big) + \frac{1}{2}w^2\rho^2(p-1)\Big]\int_0^t V_u du \\
&\quad + pw\rho\int_0^t \sqrt{V_u} dW_u^v - \frac{1}{2}p^2w^2\rho^2\int_0^t V_u du\Big)\Bigg)\Bigg] \\
&= S_0^w\, \mathrm{e}^{rwt}\, \mathbb{E}\left[\exp\left(\frac{1}{p}\left(M_t - \frac{1}{2} < M >_t\right) + \frac{1}{p}\left(p\Big[\frac{1}{2}\big(w(w-1)\big) + \frac{1}{2}w^2\rho^2(p-1)\Big]\int_0^t V_u du\right)\right)\right]
\end{aligned}$$

with $M_t = pw\rho \int_0^t \sqrt{V_u} dW_u^v$.

By applying Hölder's inequality with (p, q) satisfying $q = \dfrac{p}{p-1}$ such that $p = 1 + \dfrac{\sqrt{w\rho^2(w-1)}}{w\rho^2}$, $q = 1 + \dfrac{w\rho^2}{\sqrt{w\rho^2(w-1)}}$, we get

$$\mathbb{E}[S_t^w] \leq S_0^w \, \mathrm{e}^{rwt} \, \mathbb{E}^{\frac{1}{p}} \left[\exp\left(M_t - \frac{1}{2} < M >_t \right) \right] \mathbb{E}^{\frac{1}{q}} \left[\exp\left(\frac{q}{p} \left(p\big(\frac{1}{2}(w(w-1)) + \frac{1}{2} w^2\rho^2(p-1)\big] \int_0^t V_u du \right) \right) \right]$$

and thus

$$\begin{aligned} \sup_{t\in[0,T]} \mathbb{E}[S_t^w] \leq & S_0^w (1 \vee \mathrm{e}^{rwT}) \sup_{t\in[0,T]} \mathbb{E}^{\frac{1}{p}} \left[\exp\left(M_t - \frac{1}{2} < M >_t \right) \right] \\ & \times \mathbb{E}^{\frac{1}{q}} \left[\exp\left(\frac{q}{p} \left(p\Big[\frac{1}{2}\big(w(w-1)\big) + \frac{1}{2} w^2\rho^2(p-1)\Big] \int_0^T V_u du \right) \right) \right]. \end{aligned}$$

The stochastic exponential is a martingale if Novikov's condition is satisfied, i.e.

$$\mathbb{E}\Big[\exp\Big(\frac{1}{2} < M >_T \Big) \Big] = \mathbb{E}\Big[\exp\big(\frac{1}{2} w^2 p^2 \sigma^2 \int_0^T V_u du \big) \Big] < \infty.$$

According to Theorem 3.9.1 we get that this expectation is finite. Thus,

$$\begin{aligned} \sup_{t\in[0,T]} \mathbb{E}[S_t^w] &\leq S_0^w \, \mathrm{e}^{|r|wT} \, \mathbb{E}^{\frac{1}{q}} \left[q \left([\frac{1}{2}(w(w-1)) + \frac{1}{2} w^2\rho^2(p-1)] \int_0^T V_u du \right) \right] \\ &\leq S_0^w \, S_0^w \, \mathrm{e}^{|r|wT} \, \mathbb{E}^{\frac{1}{q}} \left[\frac{p}{p-1} \left(\frac{1}{2}(w(w-1)) + \frac{1}{2} w^2\rho^2(p-1) \right) \int_0^T V_u du \right] \\ &\leq S_0^w \, \mathrm{e}^{|r|wT} \, \mathbb{E}^{\frac{1}{q}} \left[\left(\frac{q}{2} w(w-1) + \frac{1}{2} w^2\rho^2 p \right) \int_0^T V_u du \right]. \end{aligned}$$

Finally, by applying Theorem 3.9.1 for $\lambda = \frac{q}{2} w(w-1) + \frac{1}{2} w^2\rho^2 p = \frac{1}{2}\varphi(w)^2$, the above sup is finite, completing the proof. □

Chapter 4

Gaussian Process Regression

Abstract

In this paper, we study the Gaussian process regression (GPR) method, known for speeding-up classical pricing routines while keeping a reasonable precision, to compute the prices and sensitivities of a barrier option under the popular Heston model. We show that the GPR method's performance improves effectively when trained with a Monte Carlo (MC) method that uses the Brownian bridge technique instead of a crude MC method. It turns out that the drift implicit scheme achieves the best performance among all the tested pricing methods. Besides, we study the efficiency of the analytical formulas provided systematically by the trained GPR to compute the sensitivities of barrier options. To do so, we develop new formulas for computing the Delta and Vega sensitivities using pathwise MC methods for different approximation schemes under the log-Heston model. The GPR method seems quite reliable for the computation of the Delta, except on the neighbourhood of the barrier, but less efficient for the Vega computation. Several numerical tests were processed to assess our obtained results.

4.1 Introduction

Derivative products play a key role in financial markets. Among other uses, it allows market participants to manage their risks. In particular, their pricing and hedging are important. In order to compute derivative prices, stochastic models are needed. Due to the complexity of these models, options prices cannot be found in closed form. So, different numerical approximation methods are explored to estimate derivative prices. However, some of them, like the Monte Carlo methods, are very time-consuming. Several researches has been performed to speed up calculations, but financial institutions still face certain limitations. Machine learning techniques allows to compute option prices as well as hedge ratios faster while keeping a reasonable accuracy.

De Spiegeleer et al. [26] apply Gaussian Process Regression (GPR) to estimate and fit derivatives products. They show that the use of GPR for pricing purposes speeds up the computations compared to other methods such as the Fast Fourier Transform (FFT) algorithm and the binomial tree model. Yet, the loss of precision is negligible. Their analysis is performed for vanilla options in the Heston model. For barrier type options they use a classical Monte Carlo method without the Brownian bridge technique but rather with keeping the indicator function with a discrete running minimum/maximum.

Crépey and Dixon [24] extend [26] by applying GPR to a derivative portfolio. A multi-Gaussian process regression is performed to approximate the portfolio value and its Credit Value Adjustment (CVA). A comparison is given between GPR Monte Carlo (MC-GP) and the Nested Monte Carlo. It has been concluded that the (MC-GP) method performs well in terms of calculation time for CVA.

Ludkovski [64] explores the use of a GPR model to adjust the continuation value of Bermudan options. This methodology is based on a backward dynamic programming algorithm that employs Monte Carlo simulations and GPR. In a similar way, Goudenège et al. [48] propose the GPR-MC to compute the price of American Basket put options.

In this chapter, we study the GPR method known for speeding-up classical routines while keeping a reasonable precision to compute the prices and sensitivities of a barrier option under the popular Heston model. Our algorithm learns from a training data set of parameters and returns the option prices or their sensitivities with respect to a given parameter. After the training phase, the reliability of the model is measured on a set of new parameters, called testing data set. In this context, we show that the GPR method trained by a MC based on the Brownian bridge technique have a better accuracy than a GPR method trained with the basic MC as proposed in [26]. Besides, we develop new formulas for computing the sensitivities of the barrier option that allow us to build new pathwise MC estimators (see e.g. [43] or [69] for more details on the pathwise method), in order to quantify the uncertainty in using the analytic derivatives provided systematically by the GPR method (see e.g. [29, section 5.1]). It turns out that when using our pathwise MC estimators as Benchmark values, the Delta computed by the GPR method is quite accurate except when the spot price is close to the barrier. However, the Vega provided by this latter method seems less precise (see Section 4.4 for more details).

The remainder of the chapter is presented as follows. In section 4.2, we introduce the GPR method based on Bayesian techniques and the Gaussian processes. We examine the fitting of the GPR hyperparameters for an optimal performance. In section 4.3, we introduce the various simulation schemes for the log-Heston model and we present the pricing method of a Down and out barrier option based on the conditional piecewise Brownian interpolation. Then, we analyse the sensitivity computation for the the Delta (Δ) and Vega (ν). To do so, we develop

new analytical formulas to compute barrier options sensitivities using the pathwise method for the different discretization schemes. In section 4.4, we first show that the GPR procedure trained by the MC method with the Brownian bridge technique outperforms the GPR trained by a crude MC method. For the Greeks computation, we compute the RMSE between the Greeks given automatically via the GPR method and our pathwise estimators developed in section 4.3. Finally, additional numerical tests were processed.

4.2 Gaussian process regression

In this section, we introduce the GPR method and some theoretical results that will be useful in the rest of this chapter. At the beginning, we discuss the Bayesian theory regarding metrics estimation. Then, we define the Gaussian processes and the Kernel matrices which play a crucial role in the GPR method. Finally, the GPR methodology is detailed by presenting how Gaussian processes and Bayesian theory are used to estimate deriavtives prices.

4.2.1 Bayesian method

The Bayesian technique is a statistical approach based on the Bayes Theorem. It relies on updating *a priori* knowledge based on new observed data in order to obtain a more accurate estimate, which is the *a posteriori* distribution.

Let $(X, Y) = \{(X_i, Y_i)|i = 1, ..., n\}$ with $n \geq 1$ be our dataset. The output variable Y is defined in the classical regression model as

$$Y_i = \theta X_i + \epsilon_i,$$

where ϵ_i are i.i.d Gaussian noise with variance $\sigma_n > 0$.

We are interested in the Bayesian linear regression parametric model. It is an extension of the classical linear regression in which it is assumed *a priori* that $\theta \sim \mathcal{N}(0, \tau)$, independently of the ϵ_i, $1 \leq i \leq n$. Therefore, training consists in finding the optimal random variable θ that minimises the differences between the output of the linear model and the output data Y. According to Bayesian techniques, we can find the posterior distribution Y_* given new input test points X_* as follows

$$p(Y_*|X_*, X, Y) = \int p(Y_*|X_*, \theta)p(\theta|X, Y)d\theta.$$

4.2.2 Gaussian processes

When applying the GPR, it is assumed that the *a priori* distribution is a Gaussian distribution. The Gaussian distribution is characterized by the mean μ and variance σ^2. It is denoted by $\mathcal{N}(\mu, \sigma^2)$ and the probability density function is given by:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \, \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} .$$

To further generalise the Gaussian distribution, let X be a random normal vector on $\mathbb{R}^n$. The random vector X is said to follow a multivariate Gaussian distribution if it has a density $p(x)$ given by:

$$p(x) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \exp\Big(-\frac{1}{2}(x-\mu)^\top \Sigma (x-\mu)\Big), \quad \text{with } x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n,$$

where μ is an $n \times 1$ mean vector and Σ is an $n \times n$ covariance matrix of the random vector X.

Now let's detail some multivariate distribution properties that will be useful later. Let x and y be a jointly Gaussian random vector with a joint distribution given by

$$\begin{bmatrix} x \\ y \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \begin{bmatrix} A & C \\ C^\top & B \end{bmatrix} \right).$$

Let us recall that the marginal distribution is $x \sim \mathcal{N}(\mu_x, A)$, and the conditional density of x given y is:

$$x|y \sim \mathcal{N}(\mu_x + CB^{-1}(y - \mu_y), A - CB^{-1}C^\top).$$

(See e.g. [80])

Definition 4.2.1. *[14, Section 6.4] Let us consider $\mathcal{X} \subseteq \mathbb{R}^p$, a Gaussian process is defined as a distribution over functions $f : \mathcal{X} \to \mathbb{R}$ such that for any $x_1, ..., x_n \in \mathcal{X}$, the random vector $(f(x_1), ..., f(x_n))$ has a multivariate Gaussian distribution.*

It is characterised by the mean function $\mathcal{X} \ni x \mapsto m(x) \in \mathbb{R}$ and the covariance function $\mathcal{X} \times \mathcal{X} \ni (x, x') \mapsto K_{x,x'} \in \mathbb{R}$ also known as the "kernel". We note

$$f(x) \sim \mathbf{G}\mathcal{P}(m(x), K_{x,x'}),$$

where $m(x) = \mathbb{E}[f(x)]$ and $K_{x,x'} = cov(f(x), f(x')) = \mathbb{E}[(f(x) - m(x))(f(x') - m(x'))]$, for any $x, x' \in \mathcal{X}$.

The choice of a kernel is important for a GP. It is necessary to check that the matrix K is symmetric positive semi-definite.

4.2.2.1 Reproducing kernel Hilbert space

In this subsection, we will present a brief introduction to the reproduction of kernel Hilbert spaces. The theory was developed by Aronszajn [7] and Saitoh [73].

Definition 4.2.2. *(Reproducing kernel Hilbert space) Let $\mathcal{H}$ be a Hilbert space of complex valued functions f defined on an index set $\mathcal{X}$. Then $\mathcal{H}$ is said to be Reproducing Kernel Hilbert Space with an inner product and the norm $||.||_{\mathcal{H}} = \sqrt{\langle f, f \rangle_{\mathcal{H}}}$ if there exists a function $K : \mathcal{X} \times \mathcal{X} \longrightarrow \mathbb{R}$ with the following properties:*

- *for every x, $K(x, x')$ as a function of x' belongs to $\mathcal{H}$.*
- *The matrix K has the reproducing property $\langle f(.), K(., x) \rangle_{\mathcal{H}} = f(x)$.*

The latter showed that the reproducible kernel of the RKHS is a positive semi-definite symmetric Hermitian form. Moore's Theorem proves the converse he showed that any positive semidefinite symmetric kernel defines a Hilbert space with a unique reproducible kernel.

Theorem 4.2.3. *(Moore-Aronszajn Theorem, [7]) Suppose K is a symmetric, semidefinite positive kernel on a set $\mathcal{X}$. Then there is a unique Hilbert space of functions on $\mathcal{X}$ for which K is a reproducing kernel.*

Based on Moore's Theorem, the RKHS can be obtained from the construction of the reproductive kernel function. We consider the space of functions f defined as

$$\{f(x) = \sum_{i=1}^{n} \alpha_i K(x, x_i) : n \in \mathbb{N}, x_i \in \mathcal{X}, \alpha_i \in \mathbb{R}\}$$

Let $g(x) = \sum_{j=1}^{n} \alpha'_j K(x, x'_j)$. The inner product is defined as follows

$$\langle f, g \rangle_{\mathcal{H}} = \sum_{i=1}^{n} \sum_{j=1}^{n'} \alpha_i \alpha'_j K(x_i, x_j).$$

We have also

$$\langle f(.), K(., x) \rangle_{\mathcal{H}} = \sum_{i=1}^{n} \alpha_i K(x, x_i) = f(x).$$

Now, we state the representer Theorem in the case of the quadratic error introduced by [79].

Theorem 4.2.4 (Representer Theorem (See section 6.2 e.g. [80]))**.** *Consider the functional* $J(f) = \frac{\lambda}{2}||f||_{\mathcal{H}} + \frac{1}{c}\sum_{i=1}^{n}(y_i - f(x_i))^2$ *with* $\lambda > 0$ *and* $c > 0$*. Then, if* $f \in \mathcal{H}$ *is a minimizer of* $J(f)$ *then* f *writes uniquely* $f(x) = \sum_{i=1}^{n} \alpha_i K(x, x_i)$ *for* $n \geq 1$*,* $(\alpha_i)_{i=1}^{n} \in \mathbb{R}^n$ *and* $K \in RKHS$*.*

4.2.2.2 The kernel choice

Several families of kernels can be considered, depending on the problematic that we are trying to solve. Regarding option pricing and hedging, the squared-exponential family of kernels (also called the family of Automatic Relevance Determination kernels) is one of the most frequently used. Such kernel is defined by

$$K_{SE}(x, x'; \sigma_f, \ell) = \sigma_f^2 \exp\Big\{(-\frac{1}{2\ell^2}|x - x'|^2)\Big\}, \tag{4.1}$$

where σ_f and ℓ are the so-called hyperparameters. The length scale parameter ℓ determines how far apart the input data are considered to mutually influence the predictions. Another well-known kernel family is the Matern kernel family, see [80].

In the rest of this chapter, we are going to use the squared-exponential family of kernels in order to compute prices and greeks for Down and Out call options under the Heston model.

4.2.3 Gaussian process regression

We consider $(X, Y) = \{(X_i, Y_i)|i = 1, ..., n\}$ a training data set consisting of observations. Here, for each $1 \leq i \leq n$, X_i is a d-dimensional vector of input variables and Y_i is the (potentially noisy) output. The Gaussian Process Regression (GPR) method is a machine learning method which is based on a Gaussian Process (GP) combined with Bayesian techniques to model the dependence of Y on X. In GPR, we assume a classical model regression which represents the output as follows:

$$Y_i = f(X_i) + \epsilon_i, \tag{4.2}$$

where ϵ_i are i.i.d Gaussian noise with variance σ_n and $f(X)$ is a Gaussian process. Using Bayesian methods with knowledge of the *a priori* Gaussian distribution and updating with the observed data points (X, Y), we get *a posteriori* distribution. We consider X_* with n_* inputs of dimension $\mathbb{R}^d$, we set $f_* = f(X_*)$. Then, the joint distribution is:

$$\begin{bmatrix} Y \\ f_* \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} K_{X,X} + \sigma_n^2 I & K_{X,X_*} \\ K_{X_*,X} & K_{X_*,X_*} \end{bmatrix} \right).$$

and the conditional distribution of f_* given Y is given by

$$f_*|X, Y, X_* \sim \mathcal{N}(\mathbb{E}[f_*|X, Y, X_*], cov[f_*|X, Y, X_*])$$

where

$$\mathbb{E}[f_*|X, Y, X_*] = K_{X_*,X}[K_{X,X} + \sigma_n^2 I]^{-1} Y \tag{4.3}$$

and

$$cov[f_*|X, Y, X_*] = K_{X_*,X_*} - K_{X_*,X}[K_{X,X} + \sigma_n^2 I]^{-1} K_{X,X_*}. \tag{4.4}$$

We denote by $K = K_{X,X}$ and $K_* = K_{X,X_*}$. We use the notation $k(x_*) = k_*$ for the case where there is only one test point x_* to represent the covariance vector between the test point and the n training points. Now, we can write the equations (4.3) and (4.4) for a single test point x_* and using the new notation as follows

$$\mathbb{E}[f_*|X, Y, x_*] = \bar{f}_* = k_*^T (K + \sigma_n^2 I)^{-1} Y \tag{4.5}$$

$$\mathbb{V}[f_*] = k(x_*, x_*) - k_*^T (K + \sigma_n^2 I)^{-1} k_*. \tag{4.6}$$

We can write the mean prediction as a linear combination of n kernel functions, each one centred on a learning point, given by

$$f_{x_*} = \sum_{i=1}^{n} \alpha_i k(x_i, x_*), \tag{4.7}$$

with $\alpha = k_*^T (K + \sigma_n^2 I)^{-1} Y$.

Thanks to the representer Theorem 4.2.4, the prediction mean can be written in the form of the equation (4.7) i.e. any function predicted by a regression model using a kernel can be expressed as a linear combination of the kernel values evaluated at the training points, weighted by optimal coefficients. These coefficients are determined by optimising the model using training data.

The reconstruction of a function f from limited and potentially noisy data can be a challenge. The problem is difficult to solve without making additional assumptions. In the case of a Bayesian approach, the hypotheses are characterised by an *a priori* probability distribution of the prices function, and after observation of the data, this distribution is updated to obtain an *a posteriori* distribution. The problem of introducing prior assumptions has also been discussed in the context of regularization, where these assumptions are in terms of the regularity of the function f.

Now, if we consider $\lambda = 1$ and $c = 2\sigma_n^2$, where σ_n is the variance of the Gaussian noise model. The remarkable thing is that by minimising J given in Theorem 4.2.4, we get $\hat{\alpha} = (K + \sigma_n^2 I)^{-1} Y$. Therefore the prediction for a test point x_* is $f(x_*) = k(x_*)^T (K + \sigma_n^2 I)^{-1} Y$ which is exactly the form of the GPR predictive mean defined in (4.5).

The key is to find the length-scale parameter of the associated Kernel that gives the most optimised model performance. Ludkovski and Saporito [65] define for each type of parameter in a given range $[a_i, b_i]$, with i in the given set I, a kernel with an associated window size ℓ_i and thus the optimization is performed over $(\ell_i)_{i \in I}$ instead of one single ℓ.

The drawback is that considering a set $(\ell_i)_{i \in I}$ to be optimized makes the maximum likelihood step (while training the GPR) much more difficult to perform, implying a larger computation time or a smaller precision. An alternative proposed in [29] consists in reparametrizing the interval of GPR inputs parameters to be $[0, 1]$, which then gives only one ℓ to optimise instead of several ones.

4.2.4 Hyperparameter Selection

The technique used to determine the optimal parameters of our model is maximum likelihood estimation. We look to maximise the likelihood function for the observed data $(X_i, Y_i)_{1\leq i\leq n}$ to obtain the optimal model hyperparameter. The likelihood is given by

$$L(\sigma_f, \ell, \sigma_n) = \mathbb{P}(Y|\sigma_f, \ell, \sigma_n) = \frac{1}{(2\pi)^{\frac{n}{2}}\det(K(\sigma_f, \ell) + \sigma_n^2 I)^{\frac{1}{2}}} \exp\Big(-\frac{1}{2}Y^T(K(\sigma_f, \ell) + \sigma_n^2 I)^{-1}Y\Big),$$

where $K(\sigma_f, \ell) := (K_{SE}(X_i, X_j; \sigma_f, \ell))_{1\leq i\leq n}$ is defined in (4.1). For a more simplified calculation of the derivative, we use the (log-)likelihood which is defined as follow

$$l(\sigma_f, \ell, \sigma_n) := \left(-\frac{1}{2}\log\Big(\det(K(\sigma_f, \ell) + \sigma_n^2 I)\Big) - \frac{1}{2}Y^T(K(\sigma_f, \ell) + \sigma_n^2 I)^{-1}Y + C\right), \tag{4.8}$$

where $C = -\frac{n}{2}\log(2\pi)$.
Let $C_n = [K(\sigma_f, \ell) + \sigma_n^2 I]$ and $\theta = (\sigma_f, \ell, \sigma_n)$. In order to get the maximum of (4.8), we need to compute the derivatives of the log-likelihood function given by

$$\frac{\partial}{\partial\theta_i} l(\sigma_f, \ell, \sigma_n) = -\frac{1}{2}Tr\Big(C_n^{-1}\frac{\partial C_n}{\partial\theta_i}\Big) + \frac{1}{2}Y^\top C_n^{-1}\frac{\partial C_n}{\partial\theta_i}C_n^{-1}Y, \tag{4.9}$$

and then solve this equation.

To solve (4.8), methods such as gradient descent and Quasi-Newton methods can be used (for more details see [66]). However, the function $l(\sigma_f, \ell, \sigma_n)$ is not necessarily convex. As a result, there may be several local maxima and it is difficult to obtain the global maximum. We risk falling into a local minimum rather than the absolute one. To address this issue in practice, we will initialize our built-in optimizer from a variety of initialization points to increase the probability of approaching the log-likelihood function's absolute maximum value. Here is a detailed explanation of the local maxima solution(see [9]).

4.3 Pathwise MC method for estimating barrier options sensitivities under the log-Heston model

In this section, we are interested in the problem of calculating barrier options and their sensitivities under Heston model using the GPR method. First of all, we introduce different approximation schemes based on the Brownian bridge technique for pricing barrier options that will be used for training the GPR method. Besides, we develop explicit formulas for computing the Greeks namely : the Δ and ν. We use the pathwise method (see e.g. [43]) for the different approximation schemes developed in Chapter 3 under the Heston model in order to get benchmark values for the various sensitivities that allow us to measure the efficiency of the GPR method. Finally, the numerical results show that the obtained GPR model outperforms the basic GPR model trained on a crude approximation involving the indicator function of the barrier option payoff. Let us consider the log-Heston model

$$\begin{aligned} dX_t &= (\mu - \frac{1}{2}V_t)dt + \sqrt{V_t}\ \ d\left(\rho W_t^v + \sqrt{1-\rho^2}W_t^s\right), \\ dV_t &= \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dW_t^v. \end{aligned} \tag{4.10}$$

with $\kappa,\ \theta,\ \sigma,\ X_0,\ V_0 > 0$, $\mu \in \mathbb{R}$, $\rho \in [-1, 1]$, $T > 0$, and $(W_t^s)_{t\in[0,T]}$, $(W_t^v)_{t\in[0,T]}$ are two independent Brownian motions. Let us introduce three different approximation schemes for

the log-Heston model. The time interval is divided into a uniform grid of step $h = \frac{T}{n}$, so that we have the following time points $t_i = \frac{iT}{n}$ for all $i \in \{0, 1, ..., n\}$. Different schemes have been proposed in the literature to approximate the log-Heston model. First, we consider a drift implicit Milstein scheme for log-Heston's stochastic volatility model. In this model, the stock process is approximated by using a Euler scheme and the volatility process $(V_t)_{t\in[0,T]}$ is approximated by a drift implicit Milstein scheme that preserves the positivity of the original SDE and has a strong convergence rate equal to one as proved by Neuenkirch and Szpruch in [68]. It is given by the following system:

$$\begin{aligned}\hat{X}^n_{t_{i+1}} &= \hat{X}^n_{t_i} + (\mu - \frac{1}{2}\hat{V}^n_{t_i})(t_{i+1} - t_i) + \sqrt{\hat{V}^n_{t_i}}(\tilde{W}^s_{t_{i+1}} - \tilde{W}^s_{t_i}), \\ \hat{V}^n_{t_{i+1}} &= \hat{V}^n_{t_i} + \kappa(\theta - \hat{V}^n_{t_{i+1}})h + \sigma\sqrt{\hat{V}^n_{t_i}}(W^v_{t_{i+1}} - W^v_{t_i}) + \frac{1}{2}\sigma^4((W^v_{t_{i+1}} - W^v_{t_i})^2 - h),\end{aligned} \tag{4.11}$$

for $0 \leq i \leq n-1$, with $\tilde{W}^s_t = \rho W^v_t + \sqrt{1-\rho^2}W^s_t$, $\hat{X}^n_{t_0} = X_0$ and $\hat{V}^n_{t_0} = V_0$. The second approximation scheme is the drift implicit Euler scheme for log-Heston's stochastic volatility model. Here, the stock process is still approximated by using a Euler scheme but the volatility process is given by a drift implicit Euler scheme whose strong convergence rate is also one as proved by Alfonsi [3]. It is given by the following system:

$$\begin{aligned}\tilde{X}^n_{t_{i+1}} &= \tilde{X}^n_{t_i} + (\mu - \frac{1}{2}(\hat{Y}^n_{t_i})^2)(t_{i+1} - t_i) + \hat{Y}^n_{t_i}(\tilde{W}^s_{t_{i+1}} - \tilde{W}^s_{t_i}), \\ \hat{Y}^n_{t_{i+1}} &= \frac{\hat{Y}^n_{t_i} + \gamma(W^v_{t_{i+1}} - W^v_{t_i}) + \sqrt{\left(\hat{Y}^n_{t_i} + \gamma(W^v_{t_{i+1}} - W^v_{t_i})\right)^2 + 2(1 + \frac{\kappa}{2}h)(\kappa\theta - \frac{\sigma^2}{4})h}}{2(1 + \frac{\kappa}{2}h)},\end{aligned} \tag{4.12}$$

for $0 \leq i \leq n-1$, with $\tilde{X}^n_{t_0} = X_0$ and $\hat{Y}^n_{t_0} = Y_0 = \sqrt{V_0}$. The last approximation is the Semi-exact log-Heston approximation scheme studied in Chapter 3. It is given by

$$\begin{aligned}\overline{X}^n_{t_{i+1}} &= \overline{X}^n_{t_i} + (\mu - \frac{1}{2}V_{t_i})(t_{i+1} - t_i) + \frac{\rho}{\sigma}\Big(V_{t_{i+1}} - V_{t_i} - \kappa\theta(t_{i+1} - t_i) + \kappa V_{t_i}(t_{i+1} - t_i)\Big) \\ &\quad + \sqrt{1-\rho^2}\sqrt{V_{t_i}}(W^s_{t_{i+1}}. - W^s_{t_i}),\end{aligned} \tag{4.13}$$

with $\overline{X}_0 = X_0$ and where we estimate the volatility using the exact simulation of a χ^2 distribution: conditionally to V_u, we have the variance process V_t at time t, $t > u$ given by

$$V_t \overset{d}{=} \frac{\sigma^2(1 - \mathrm{e}^{-\kappa(t-u)})}{4\kappa}\chi^2_d\Big(\frac{4\kappa\,\mathrm{e}^{-\kappa(t-u)}}{\sigma^2(1 - \mathrm{e}^{-\kappa(t-u)})}V_u\Big), \tag{4.14}$$

where $\chi^2_d(\lambda)$ denotes a non central χ^2 random variable with $d := \frac{4\theta\kappa}{\sigma^2}$ degrees of freedom and non-centrality parameter λ.

We consider a Down and Out barrier call option with a strike K and a barrier level D. For $f : \mathbb{R} \to \mathbb{R}$ a C^1 function with bounded derivative, we consider the payoff given by

$$P = f(X_T)\mathbb{1}_{\{\inf_{t\in[0,T]} X_t > D\}}. \tag{4.15}$$

Therefore, to approximate the barrier option price we use the numerical approximation scheme studied in Chapter 3 based on the simulation of the running minimum at each time step using a Brownian bridge interpolation (See [43]). In order to use this technique we introduce the continuous version of the above approximation schemes of the log-Heston price process given respectively by

$$\hat{X}^n_t = \hat{X}^n_{t_i} + (\mu - \frac{1}{2}\hat{V}^n_{t_i})(t - t_i) + \sqrt{\hat{V}^n_{t_i}}(\tilde{W}^s_t - \tilde{W}^s_{t_i}), \text{ for all } t \in [t_i, t_{i+1}],$$

$$\tilde{X}^n_t = \tilde{X}^n_{t_i} + (\mu - \frac{1}{2}(\hat{Y}^n_{t_i})^2)(t-t_i) + \hat{Y}^n_{t_i}(\tilde{W}^s_t - \tilde{W}^s_{t_i}), \text{ for all } t \in [t_i, t_{i+1}],$$

and

$$\begin{aligned}\overline{X}^n_t = \overline{X}^n_{t_i} + (\mu - \frac{1}{2}V_{t_i})(t-t_i) + \frac{\rho}{\sigma}\Big(\frac{n}{T}(V_{t_{i+1}} - V_{t_i})(t-t_i) - \kappa\theta(t-t_i) + \kappa V_{t_i}(t-t_i)\Big) \\ + \sqrt{1-\rho^2}\sqrt{V_{t_i}}(W^s_t - W^s_{t_i}), \text{ for all } t \in [t_i, t_{i+1}].\end{aligned}$$

Thus applying the Brownian bridge technique, we approximate the barrier option payoff by either

$$\hat{P} = f(\hat{X}^n_T)_+ \prod_{i=0}^{n-1}(1-\hat{q}_i), \text{ where } \hat{q}_i := \exp\Big(\frac{-2(\hat{X}^n_{t_i} - D)_+(\hat{X}^n_{t_{i+1}} - D)_+}{\hat{V}^n_{t_i} h}\Big) \tag{4.16}$$

or

$$\tilde{P} = f(\tilde{X}^n_T)_+ \prod_{i=0}^{n-1}(1-\tilde{q}_i), \text{ where } \tilde{q}_i := \exp\Big(\frac{-2(\tilde{X}^n_{t_i} - D)_+(\tilde{X}^n_{t_{i+1}} - D)_+}{(\hat{Y}^n_{t_i})^2 h}\Big) \tag{4.17}$$

or

$$\overline{P} = f(\overline{X}^n_T) \prod_{i=0}^{n-1}(1-\tilde{q}_i), \text{ where } \overline{q}_i := \exp\Big(\frac{-2(\overline{X}^n_{t_i} - D)_+(\overline{X}^n_{t_{i+1}} - D)_+}{(1-\rho^2)V_{t_i} h}\Big). \tag{4.18}$$

In order to test the efficiency of the GPR method in approximating barrier option prices and their associated sensitivities, we need to develop accurate benchmark values. To do so, we run for each approximation scheme presented above a Monte Carlo method with a high sample size $N = 200000$, (see Section 4.4). For the Greeks computation we develop pathwise Monte Carlo method also for each approximation schemes. For this purpose, let us note that by standard computations we obtain a general formula for the partial derivative of $\hat{P}$ with respect to any parameter θ

$$\frac{\partial \hat{P}}{\partial \theta} = f'(\hat{X}^n_T)\frac{\partial \hat{X}^n_T}{\partial \theta}\prod_{i=0}^{n-1}(1-\hat{q}_i) - f(\hat{X}^n_T)\sum_{i=0}^{n-1}\left[\prod_{k=0,k\neq i}^{n-1}(1-\hat{q}_k)\frac{\partial \hat{q}_i}{\partial \theta}\right], \tag{4.19}$$

where

$$\frac{\partial \hat{q}_i}{\partial \theta} = \mathbb{1}_{(\hat{X}^n_{t_i}, \hat{X}^n_{t_{i+1}} > D)}\hat{q}_i\left[\frac{-2\delta_{t_i}(\hat{X}^n_{t_{i+1}} - D)}{b^2_{t_i} h} + \frac{-2\delta_{t_{i+1}}(\hat{X}^n_{t_i} - D)}{b^2_{t_i} h} + \frac{4(\hat{X}^n_{t_i} - D)(\hat{X}^n_{t_{i+1}} - D)\dot{b}_{t_i}}{b^3_{t_i} h}\right] \tag{4.20}$$

with $\delta_{t_i} = \dfrac{\partial \hat{X}^n_{t_i}}{\partial \theta}$, and $\dot{b}_{t_i} = \dfrac{\partial b_{t_i}}{\partial \theta}$ where $b_{t_i} = \sqrt{\hat{V}^n_{t_i}}$. (see e.g. [16]). Of course the same formula remains valid for $\frac{\partial \overline{P}}{\partial \theta}$ and $\frac{\partial \tilde{P}}{\partial \theta}$ with replacing the process $\hat{X}^n$ respectively by $\tilde{X}^n$ and $\overline{X}^n$, $\hat{q}_i$ respectively by $\overline{q}_i$ and $\tilde{q}_i$ and b_{t_i} resectively by $\hat{Y}^n_{t_i}$ and $\sqrt{(1-\rho^2)V_{t_i}}$. In what follows, we are mainly interested in two sensitivities $\Delta = \frac{\partial P}{\partial X_0}$ and $\nu = \frac{\partial P}{\partial V_0}$.

4.3.1 The drift implicit Milstein scheme for log-Heston's stochastic volatility model: pathwise sensitivities

In this subsection, we study the calculation of the Delta (Δ) and Vega (ν) for a Down and Out call option by using the drift implicit Milstein scheme for log-Heston's stochastic volatility model (4.11).

Theorem 4.3.1. *Let $X_0 > D$ and consider the drift implicit Milstein scheme defined in* (4.11). *We have*

$$\frac{\partial \hat{P}}{\partial X_0} = f'(\hat{X}_T^n)\prod_{i=0}^{n-1}(1-\hat{q}_i) - f(\hat{X}_T^n)\sum_{i=0}^{n-1}\left[\prod_{k=0,k\neq i}^{n-1}(1-\hat{q}_k)\frac{\partial \hat{q}_i}{\partial X_0}\right],$$

where

$$\frac{\partial \hat{q}_i}{\partial X_0} = \mathbb{1}_{\left(\hat{X}_{t_i}^n, \hat{X}_{t_{i+1}}^n > D\right)}\hat{q}_i\left[\frac{-2(\hat{X}_{t_{i+1}}^n - D)}{\hat{V}_{t_i}^n h} + \frac{-2(\hat{X}_{t_i}^n - D)}{\hat{V}_{t_i}^n h}\right],$$

Proof. For all $0 \leq i \leq n$, we have

$$\hat{X}_{t_i}^n = X_0 + \mu t_i - \frac{1}{2}\frac{T}{n}\sum_{k=0}^{i-1}\hat{V}_{t_k}^n + \sum_{k=0}^{i-1}\int_{t_k}^{t_{k+1}}\sqrt{\hat{V}_{t_k}^n}d\tilde{W}_s^s.$$

Then, for all $1 \leq i \leq n-1$, we have

$$\frac{\partial \hat{X}_{t_i}^n}{\partial X_0} = 1$$

As $b_{t_i} = \sqrt{\hat{V}_{t_i}^n}$ we have $\dot{b}_{t_i} = 0$. Replacing in (4.20) we get

$$\frac{\partial \hat{q}_i}{\partial X_0} = \mathbb{1}_{\left(\hat{X}_{t_i}^n, \hat{X}_{t_{i+1}}^n > D\right)}\hat{q}_i\left[\frac{-2(\hat{X}_{t_{i+1}}^n - D)}{\hat{V}_{t_i}^n h} + \frac{-2(\hat{X}_{t_i}^n - D)}{\hat{V}_{t_i}^n h}\right],$$

which completes the proof.

□

We proceed similarly for the Vega calculation.

Theorem 4.3.2. *Let $X_0 > D$ and consider the drift implicit Milstein scheme defined in* (4.11). *We have*

$$\frac{\partial \hat{P}}{\partial V_0} = f'(\hat{X}_T^n)\frac{\partial \hat{X}_T^n}{\partial V_0}\prod_{i=0}^{n-1}(1-\hat{q}_i) - f(\hat{X}_T^n) + \sum_{i=0}^{n-1}\left[\prod_{k=0,k\neq i}^{n-1}(1-\hat{q}_k)\frac{\partial \hat{q}_i}{\partial V_0}\right],$$

where

$$\frac{\partial \hat{q}_i}{\partial V_0} = \mathbb{1}_{\left(\hat{X}_{t_i}^n, \hat{X}_{t_{i+1}}^n > D\right)}\hat{q}_i\left[\frac{-2\delta_{t_i}(\hat{X}_{t_{i+1}}^n - D)}{\hat{V}_{t_i}^n h} + \frac{-2\delta_{t_{i+1}}(\hat{X}_{t_i}^n - D)}{\hat{V}_{t_i}^n h} + \frac{4\dot{b}_{t_i}(\hat{X}_{t_i}^n - D)(\hat{X}_{t_{i+1}}^n - D)}{\sqrt{\hat{V}_{t_i}^n}^3 h}\right],$$

with

$$\delta_{t_{i+1}} = -\frac{1}{2}\frac{T}{n}\sum_{k=0}^{i}\Big(\prod_{j=1}^{k}\varphi_j\Big) + \sum_{k=0}^{i}\frac{1}{2\sqrt{\hat{V}_{t_k}^n}}\Big(\prod_{j=1}^{k}\varphi_j\Big)(\tilde{W}_{t_{k+1}}^s - \tilde{W}_{t_k}^s),$$

and $\dot{b}_{t_i} = \frac{1}{2\sqrt{\hat{V}_{t_i}^n}}\prod_{k=1}^{i}\varphi_k$, *where*

$$\varphi_k = \left(1 + \frac{\sigma}{2\sqrt{\hat{V}_{k-1}^n}}(W_{t_k}^v - W_{t_{k-1}}^v)\right)\frac{1}{1+\kappa h}, \text{ where } k = 1, ..., n.$$

Proof. By (4.19) that

$$\frac{\partial \hat{P}}{\partial V_0} = f'(\hat{X}_T^n)\frac{\partial \hat{X}_T^n}{\partial V_0}\prod_{i=0}^{n-1}(1-\hat{q}_i) - f(\hat{X}_T^n)\sum_{i=0}^{n-1}\left[\prod_{k=0,k\neq i}^{n-1}(1-\hat{q}_k)\frac{\partial \hat{q}_i}{\partial V_0}\right],$$

where

$$\frac{\partial \hat{q}_i}{\partial V_0} = \mathbb{1}_{(\hat{X}_{t_i}^n, \hat{X}_{t_{i+1}}^n > D)}\hat{q}_i\left[\frac{-2\delta_{t_i}(\hat{X}_{t_{i+1}}^n - D)}{b_{t_i}^2 h} + \frac{-2\delta_{t_{i+1}}(\hat{X}_{t_i}^n - D)}{b_{t_i}^2 h} + \frac{4(\hat{X}_{t_i}^n - D)(\hat{X}_{t_{i+1}}^n - D)\dot{b}_{t_i}}{b_{t_i}^3 h}\right], \tag{4.21}$$

with $\delta_{t_i} = \dfrac{\partial \hat{X}_{t_i}^n}{\partial V_0}$ and $\dot{b}_{t_i} = \dfrac{\partial b_{t_i}}{\partial V_0}$, where $b_{t_i} = \sqrt{\hat{V}_{t_i}^n}$. We recall that the implicit Milstein scheme is defined as follows for all $0 \leq i \leq n$

$$\hat{X}_{t_i}^n = X_0 + \mu t_i - \frac{1}{2}\frac{T}{n}\sum_{k=0}^{i-1}\hat{V}_{t_k}^n + \sum_{k=0}^{i-1}\sqrt{\hat{V}_{t_k}^n}(\tilde{W}_{t_{k+1}}^s - \tilde{W}_{t_k}^s).$$

We have for all $1 \leq i \leq n-1$,

$$\delta_{t_{i+1}} = \frac{\partial \hat{X}_{t_{i+1}}^n}{\partial V_0} = -\frac{1}{2}\frac{T}{n}\sum_{k=0}^{i}\frac{\partial \hat{V}_{t_k}^n}{\partial V_0} + \sum_{k=0}^{i}\frac{1}{2\sqrt{\hat{V}_{t_k}^n}}\frac{\partial \hat{V}_{t_k}^n}{\partial V_0}(\tilde{W}_{t_{k+1}}^s - \tilde{W}_{t_k}^s),$$

with $\delta_{t_0} = 0$. Now, by using the chain rule we get

$$\frac{\partial \hat{V}_{t_k}^n}{\partial V_0} = \frac{\partial \hat{V}_{t_k}^n}{\partial \hat{V}_{t_{k-1}}^n}\frac{\partial \hat{V}_{t_{k-1}}^n}{\partial V_0} = \prod_{j=1}^{k}\frac{\partial \hat{V}_{t_j}^n}{\partial \hat{V}_{t_{j-1}}^n},$$

and according to formula (4.11),

$$(1+\kappa h)\hat{V}_{t_j}^n = \hat{V}_{t_{j-1}}^n + \kappa\theta h + \sigma\sqrt{\hat{V}_{t_{j-1}}^n}(W_{t_j}^v - W_{t_{j-1}}^v) + \frac{1}{2}\sigma^4((W_{t_j}^v - W_{t_{j-1}}^v)^2 - h)$$

so

$$(1+\kappa h)\frac{\partial \hat{V}_{t_j}^n}{\partial \hat{V}_{t_{j-1}}^n} = 1 + \frac{\sigma}{2\sqrt{\hat{V}_{t_{j-1}}^n}}(W_{t_j}^v - W_{t_{j-1}}^v)$$

which implies that

$$\frac{\partial \hat{V}_{t_k}^n}{\partial V_0} = \prod_{j=1}^{k}\left[\left(1 + \frac{\sigma}{2\sqrt{\hat{V}_{t_{j-1}}^n}}(W_{t_j}^v - W_{t_{j-1}}^v)\right)\frac{1}{1+\kappa h}\right] =: \prod_{j=1}^{k}\varphi_j,$$

where

$$\varphi_j = \left(1 + \frac{\sigma}{2\sqrt{\hat{V}_{t_{j-1}}^n}}(W_{t_j}^v - W_{t_{j-1}}^v)\right)\frac{1}{1+\kappa h}.$$

Then, we have that $\delta_{t_0} = 0$, and

$$\delta_{t_{i+1}} = -\frac{1}{2}\frac{T}{n}\sum_{k=0}^{i}\Big(\prod_{j=1}^{k}\varphi_j\Big) + \sum_{k=0}^{i}\frac{1}{2\sqrt{\hat{V}_{t_k}^n}}\Big(\prod_{j=1}^{k}\varphi_j\Big)(\tilde{W}_{t_{k+1}}^s - \tilde{W}_{t_k}^s).$$

Besides, we have

$$\dot{b}_{t_i} = \frac{1}{2\sqrt{\hat{V}_{t_i}^n}}\prod_{k=1}^{i}\varphi_k,$$

which complete the proof. □

4.3.2 The drift implicit Euler scheme for log-Heston's stochastic volatility model : pathwise sensitivities

The aim of this part is to study the calculation of the Delta (Δ) and Vega (ν) for a Down and Out call option by using the drift implicit Euler scheme for log-Heston's stochastic volatility model.

Theorem 4.3.3. *Let $X_0 > D$ and consider the drift implicit Euler scheme defined in* (4.12)*. We have*

$$\frac{\partial \tilde{P}}{\partial X_0} = f'(\tilde{X}_T^n)\prod_{i=0}^{n-1}(1-\tilde{q}_i) - f(\overline{X}_T^n)\sum_{i=0}^{n-1}\left[\prod_{k=0,k\neq i}^{n-1}(1-\tilde{q}_k)\frac{\partial \tilde{q}_i}{\partial X_0}\right],$$

where

$$\frac{\partial \tilde{q}_i}{\partial X_0} = \mathbb{1}_{\left(\overline{X}_{t_i}^n, \overline{X}_{t_{i+1}}^n > D\right)}\tilde{q}_i\left[\frac{-2(\overline{X}_{t_{i+1}}^n - D)}{(\hat{Y}_{t_i}^n)^2 h} + \frac{-2(\overline{X}_{t_i}^n - D)}{(\hat{Y}_{t_i}^n)^2 h}\right].$$

Proof. For all $0 \leq i \leq n$, we have the following scheme

$$\tilde{X}_{t_i}^n = X_0 + \mu t_i - \frac{1}{2}\frac{T}{n}\sum_{k=0}^{i-1}(\hat{Y}_{t_k}^n)^2 + \sum_{k=0}^{i-1}\hat{Y}_{t_k}^n(\tilde{W}_{t_{k+1}}^s - \tilde{W}_{t_k}^s).$$

Then,

$$\frac{\partial \tilde{X}_{t_i}^n}{\partial X_0} = 1.$$

Since $b_{t_i} = \hat{Y}_{t_i}^n$, then $\dot{b}_{t_i} = 0$. Replacing in (4.20) we obtain

$$\frac{\partial \overline{q}_i}{\partial X_0} = \mathbb{1}_{\left(\overline{X}_{t_i}^n, \overline{X}_{t_{i+1}}^n > D\right)}\overline{q}_i\left[\frac{-2(\overline{X}_{t_{i+1}}^n - D)}{(\hat{Y}_{t_i}^n)^2 h} + \frac{-2(\overline{X}_{t_i}^n - D)}{(\hat{Y}_{t_i}^n)^2 h}\right].$$

which establishes the formula. □

Theorem 4.3.4. *Let $X_0 > D$ and we consider the drift implicit Euler scheme defined in* (4.12)*. We have*

$$\frac{\partial \tilde{P}}{\partial V_0} = f'(\tilde{X}_T^n)\frac{\partial \tilde{X}_T^n}{\partial V_0}\prod_{i=0}^{n-1}(1-\tilde{q}_i) - f(\tilde{X}_T^n)\sum_{i=0}^{n-1}\left[\prod_{k=0,k\neq i}^{n-1}(1-\tilde{q}_k)\frac{\partial \tilde{q}_i}{\partial V_0}\right],$$

where

$$\frac{\partial \tilde{q}_i}{\partial V_0} = \mathbb{1}_{\left\{\tilde{X}_{t_i}^n, \tilde{X}_{t_{i+1}}^n > D\right\}}\tilde{q}_i\left[\frac{-2\delta_{t_i}(\tilde{X}_{t_{i+1}}^n - D)}{(\hat{Y}_{t_i}^n)^2 h} + \frac{-2\delta_{t_{i+1}}(\tilde{X}_{t_i}^n - D)}{(\hat{Y}_{t_i}^n)^2 h} + \frac{4\tilde{b}_{t_i}(\tilde{X}_{t_i}^n - D)(\tilde{X}_{t_{i+1}}^n - D)}{(\hat{Y}_{t_i}^n)^3 h}\right],$$

with $\delta_{t_{i+1}} = -\frac{1}{2}\frac{T}{n}\sum_{k=0}^{i}\frac{\hat{Y}_{t_k}^n}{Y_0}\left(\prod_{j=1}^{k}\xi_j\right) + \sum_{k=0}^{i}\left(\prod_{j=1}^{k}\xi_j\right)(\tilde{W}_{t_{k+1}}^s - \tilde{W}_{t_k}^s)$ *and* $\dot{b}_{t_i} = \frac{1}{2Y_0}\left(\prod_{j=1}^{k}\xi_j\right)$,

$$\xi_k = \frac{1}{2(1+\frac{\kappa}{2}h)}\left(1 + \frac{Y_{t_{k-1}} + \gamma(W_{t_k}^v - W_{t_{k-1}}^v)}{\sqrt{\left(Y_{t_{k-1}} + \gamma(W_{t_k}^v - W_{t_{k-1}}^v)\right)^2 + 2\left(1+\frac{\kappa}{2}h\right)(\kappa\theta - \frac{\sigma^2}{4})h}}\right) \text{ for all } k = 1, .., n.$$

Proof. We know from (4.19) that

$$\frac{\partial \tilde{P}}{\partial V_0} = f'(\tilde{X}_T^n)\frac{\partial \tilde{X}_T^n}{\partial V_0}\prod_{i=0}^{n-1}(1-\tilde{q}_i) - f(\tilde{X}_T^n)\sum_{i=0}^{n-1}\left[\prod_{k=0,k\neq i}^{n-1}(1-\tilde{q}_k)\frac{\partial \tilde{q}_i}{\partial V_0}\right],$$

where

$$\frac{\partial \tilde{q}_i}{\partial V_0} = \mathbb{1}_{(\tilde{X}_{t_i}^n, \tilde{X}_{t_{i+1}}^n > D)}\tilde{q}_i\left[\frac{-2\delta_{t_i}(\tilde{X}_{t_{i+1}}^n - D)}{b_{t_i}^2 h} + \frac{-2\delta_{t_{i+1}}(\tilde{X}_{t_i}^n - D)}{b_{t_i}^2 h} + \frac{4(\tilde{X}_{t_i}^n - D)(\tilde{X}_{t_{i+1}}^n - D)\dot{b}_{t_i}}{b_{t_i}^3 h}\right], \tag{4.22}$$

with $\delta_{t_i} = \dfrac{\partial \tilde{X}_{t_i}^n}{\partial V_0}$ and $\dot{b}_{t_i} = \dfrac{\partial b_{t_i}}{\partial V_0}$ where $b_{t_i} = \hat{Y}_{t_i}^n$. Let us recall that the drift implicit Euler scheme is defined as follows for all $0 \leq i \leq n$

$$\tilde{X}_{t_i}^n = X_0 + \mu t_i - \frac{1}{2}\frac{T}{n}\sum_{k=0}^{i-1}(\hat{Y}_{t_k}^n)^2 + \sum_{k=0}^{i-1}\hat{Y}_{t_k}^n(\tilde{W}_{t_{k+1}}^s - \tilde{W}_{t_k}^s).$$

then,

$$\delta_{t_{i+1}} = \frac{\partial \tilde{X}_{t_{i+1}}^n}{\partial V_0} = -\frac{1}{2}\frac{T}{n}\sum_{k=0}^{i}\frac{\partial(\hat{Y}_{t_k}^n)^2}{\partial V_0} + \sum_{k=0}^{i}\frac{\partial \hat{Y}_{t_k}^n}{\partial V_0}(\tilde{W}_{t_{k+1}}^s - \tilde{W}_{t_k}^s),$$

with $\delta_{t_0} = 0$ and as $Y_0 = \sqrt{V_0}$, we get

$$\frac{\partial(\hat{Y}_{t_k}^n)^2}{\partial V_0} = 2\hat{Y}_{t_k}^n\frac{\partial \hat{Y}_{t_k}^n}{\partial V_0} = 2\hat{Y}_{t_k}^n\frac{\partial \hat{Y}_{t_k}^n}{\partial Y_0}\frac{\partial Y_0}{\partial V_0} = \frac{\hat{Y}_{t_k}^n}{Y_0}\frac{\partial \hat{Y}_{t_k}^n}{\partial Y_0}. \tag{4.23}$$

By applying the chain rule, we obtain

$$\begin{aligned}\frac{\partial \hat{Y}_{t_k}^n}{\partial Y_0} &= \prod_{j=1}^{k}\frac{\partial \hat{Y}_{t_j}^n}{\partial \hat{Y}_{t_{j-1}}^n} \\ &= \prod_{j=1}^{k}\frac{1}{2(1+\frac{\kappa}{2}h)}\left(1 + \frac{\hat{Y}_{t_{j-1}}^n + \gamma(W_{t_j}^v - W_{t_{j-1}}^v)}{\sqrt{\left(\hat{Y}_{t_{j-1}}^n + \gamma(W_{t_j}^v - W_{t_{j-1}}^v)\right)^2 + 2\big(1+\frac{\kappa}{2}h\big)(\kappa\theta - \frac{\sigma^2}{4})h}}\right),\end{aligned}$$

then, we can rewrite (4.23) as

$$\begin{aligned}\frac{\partial(\hat{Y}_{t_k}^n)^2}{\partial V_0} &= \frac{\hat{Y}_{t_k}^n}{Y_0}\prod_{j=1}^{k}\frac{1}{2(1+\frac{\kappa}{2}h)}\left(1 + \frac{\hat{Y}_{t_{j-1}}^n + \gamma(W_{t_j}^v - W_{t_{j-1}}^v)}{\sqrt{\left(\hat{Y}_{t_{j-1}}^n + \gamma(W_{t_j}^v - W_{t_{j-1}}^v)\right)^2 + 2\big(1+\frac{\kappa}{2}h\big)(\kappa\theta - \frac{\sigma^2}{4})h}}\right) \\ &= \frac{\hat{Y}_{t_k}^n}{Y_0}\prod_{j=1}^{k}\xi_j,\end{aligned}$$

with $\xi_j = \dfrac{1}{2(1+\frac{\kappa}{2}h)}\left(1 + \dfrac{\hat{Y}_{t_{j-1}}^n + \gamma(W_{t_j}^v - W_{t_{j-1}}^v)}{\sqrt{\left(\hat{Y}_{t_{j-1}}^n + \gamma(W_{t_j}^v - W_{t_{j-1}}^v)\right)^2 + 2\big(1+\frac{\kappa}{2}h\big)(\kappa\theta - \frac{\sigma^2}{4})h}}\right)$. Therefore, we have that $\delta_{t_0} = 0$,

$$\delta_{t_{i+1}} = -\frac{1}{2}\frac{T}{n}\sum_{k=0}^{i}\frac{\hat{Y}_{t_k}^n}{Y_0}\Big(\prod_{j=1}^{k}\xi_j\Big) + \sum_{k=0}^{i}\Big(\prod_{j=1}^{k}\xi_j\Big)(\tilde{W}_{t_{k+1}}^s - \tilde{W}_{t_k}^s).$$

Additionally,

$$\dot{b}_{t_i} = \frac{\partial \hat{Y}^n_{t_i}}{\partial V_0} = \frac{\partial \hat{Y}^n_{t_i}}{\partial Y_0}\frac{\partial Y_0}{\partial V_0} = \frac{1}{2Y_0}\Big(\prod_{j=1}^{k}\xi_j\Big),$$

We complete the proof using formula (4.19) and (4.20). □

4.3.3 The Semi-exact log-Heston scheme: pathwise sensitivities

The goal of this section is to study the calculation of the Delta (Δ) and Vega (ν) for a Down and Out call option by using the Semi-Exact log-Heston's stochastic volatility model

Theorem 4.3.5. *Let $X_0 > D$ and consider the log-Heston scheme $(\overline{X}^n_{t_i})_{1\le i\le n}$ defined in* (4.13)*. We have*

$$\frac{\partial \overline{P}}{\partial X_0} = f'(\overline{X}^n_T)\prod_{i=0}^{n-1}(1-\tilde{q}_i) - f(\overline{X}^n_T)\sum_{i=0}^{n-1}\left[\prod_{k=0,k\neq i}^{n-1}(1-\overline{q}_k)\frac{\partial \overline{q}_i}{\partial X_0}\right],$$

$$\frac{\partial \overline{q}_i}{\partial X_0} = \mathbb{1}_{(\overline{X}^n_{t_i},\overline{X}^n_{t_{i+1}}>D)}\overline{q}_i\left[\frac{-2(\overline{X}^n_{t_{i+1}}-D)}{(1-\rho^2)V_{t_i}h} + \frac{-2(\overline{X}^n_{t_i}-D)}{(1-\rho^2)V_{t_i}h}\right]. \tag{4.24}$$

Proof. According to (4.13), we can rewrite $\overline{X}^n_{t_i}$ as

$$\overline{X}^n_{t_i} = X_0 + \mu t_i + (\frac{\kappa\rho}{\sigma}-\frac{1}{2})\frac{T}{n}\sum_{k=0}^{i-1}V_{t_k} + \frac{\rho}{\sigma}(V_{t_i}-V_0-i\kappa\theta\frac{T}{n}) + \sqrt{1-\rho^2}\sum_{k=0}^{i-1}\int_{t_k}^{t_{k+1}}\sqrt{V_{t_k}}dW^s_s. \tag{4.25}$$

We have that $\dfrac{\partial \overline{X}^n_{t_i}}{\partial X_0} = 1$. In addition, we have that $b_{t_i} = \sqrt{1-\rho^2}\sqrt{V_{t_i}}$. Therefore,

$$\dot{b}_{t_i} = \frac{\partial b_{t_i}}{\partial X_0} = 0.$$

As a consequence,

$$\frac{\partial \overline{q}_i}{\partial X_0} = \mathbb{1}_{(\overline{X}^n_{t_i},\overline{X}^n_{t_{i+1}}>D)}\overline{q}_i\left[\frac{-2(\overline{X}^n_{t_{i+1}}-D)}{(1-\rho^2)V_{t_i}h} + \frac{-2(\overline{X}^n_{t_i}-D)}{(1-\rho^2)V_{t_i}h}\right].$$

□

Theorem 4.3.6. *Let $X_0 \in (X_{0,min}, X_{0,max})$ where $X_{0,min} > D$ and let $2\kappa\theta > \sigma^2$. Then, for $\overline{P}$ defined in* (4.18)*, we have*

$$\frac{\partial \mathbb{E}(\overline{P})}{\partial X_0} = \mathbb{E}\Big(\frac{\partial \overline{P}}{\partial X_0}\Big).$$

Proof. To prove the equality, we use Theorem 4.6.1. The first two conditions are easy to verify, and we are left to demonstrate (3). For every $X_0 \in (X_{0,min}, X_{0,max})$, as f' is bounded and f is with linear growth, we use Theorem 4.3.5 to get the existence of a positive constant C that may change from line to line such that

$$\Big|\frac{\partial \overline{P}}{\partial X_0}\Big| \le C\Big(1 + (1+|\overline{X}^n_T|)\sum_{i=0}^{n-1}\Big|\frac{\partial \overline{q}_i}{\partial X_0}\Big|\Big).$$

By using (4.24) we have for all $0 \leq i \leq n-1$

$$\Big|\frac{\partial \overline{q}_i}{\partial X_0}\Big| \leq C\Big(\frac{|\overline{X}^n_{t_{i+1}}| + |\overline{X}^n_{t_i}| + D}{V_{t_i}}\Big).$$

Then,

$$\Big|\frac{\partial \overline{P}}{\partial X_0}\Big| \leq C\Big(1 + (1+|\overline{X}^n_T|)\sum_{i=0}^{n-1}\Big(|\overline{X}^n_{t_{i+1}}| + |\overline{X}^n_{t_i}| + D\Big)\frac{1}{V_{t_i}}\Big).$$

Since the approximation scheme (4.25) can be written as $\overline{X}^n_{t_i} = X_0 + \zeta_i$, where $\zeta_i = \mu t_i + (\frac{\kappa\rho}{\sigma} - \frac{1}{2})\frac{T}{n}\sum_{k=0}^{i-1} V_{t_k} + \frac{\rho}{\sigma}(V_{t_i} - V_0 - i\kappa\theta\frac{T}{n}) + \sqrt{1-\rho^2}\sum_{k=0}^{i-1}\int_{t_k}^{t_{k+1}}\sqrt{V_{t_k}}dW^s_s$, then we get

$$\Big|\frac{\partial \overline{P}}{\partial X_0}\Big| \leq C\Bigg(1 + \sum_{i=0}^{n-1}(1 + |X_{0,max}| + |\zeta_n|)\Big(|X_{0,max}| + |\zeta_i| + D\Big)\frac{1}{V_{t_i}}\Bigg).$$

Thus, by using Hölder's inequality, for $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$ yields

$$\mathbb{E}\Big|\frac{\partial \overline{P}}{\partial X_0}\Big| \leq C\Big(1 + \sum_{i=0}^{n-1}\mathbb{E}^{\frac{1}{p}}\Big|(1 + |X_{0,max}| + |\zeta_n|)\Big(|X_{0,max}| + |\zeta_i| + D\Big)\Big|^p\mathbb{E}^{\frac{1}{q}}\Big(\Big|\frac{1}{V_{t_i}}\Big|^q\Big)\Big).$$

Finally, we use lemmas 3.3.1, 3.3.3 and 3.3.4 to conclude that $\mathbb{E}\Big|\frac{\partial \overline{P}}{\partial X_0}\Big| < \infty$ by choosing $1 < q < \frac{2\kappa\theta}{\sigma^2}$. □

Theorem 4.3.7. *Let $X_0 > D$ and consider the log-Heston scheme $(\overline{X}^n_{t_i})_{1\leq i\leq n}$ defined in* (4.13)*. We have*

$$\frac{\partial \overline{P}}{\partial V_0} = f'(\overline{X}^n_T)\frac{\partial \overline{X}^n_T}{\partial V_0}\prod_{i=0}^{n-1}(1-\overline{q}_i) - f(\overline{X}^n_T)\sum_{i=0}^{n-1}\Bigg[\prod_{k=0,k\neq i}^{n-1}(1-\overline{q}_k)\frac{\partial \overline{q}_i}{\partial V_0}\Bigg], \tag{4.26}$$

$$\begin{aligned}\frac{\partial \overline{q}_i}{\partial V_0} = \mathbb{1}_{(\overline{X}^n_{t_i}, \overline{X}^n_{t_{i+1}} > D)}\overline{q}_i\Bigg[&\frac{-2\delta_{t_i}(\overline{X}^n_{t_{i+1}} - D)}{(1-\rho^2)V_{t_i}h} + \frac{-2\delta_{t_{i+1}}(\overline{X}^n_{t_i} - D)}{(1-\rho^2)V_{t_i}h} \\ &+ \frac{4(\overline{X}^n_{t_i} - D)(\overline{X}^n_{t_{i+1}} - D)\dot{b}_{t_i}}{(1-\rho^2)^{\frac{3}{2}}V_{t_i}^{\frac{3}{2}}h}\Bigg]\end{aligned}$$

with

$$\delta_{t_{i+1}} = \frac{\partial \overline{X}^n_{t_i}}{\partial V_0} = T(\frac{\rho\kappa}{\sigma} - \frac{1}{2}) + \frac{\rho}{\sigma}(\prod_{k=1}^{i}\varphi_k - 1) + \frac{\sqrt{1-\rho^2}}{2}\sum_{k=0}^{i-1}\frac{1}{\sqrt{V_{t_k}}}\prod_{k=1}^{i}\varphi_k(W^s_{t_{i+1}} - W^s_{t_i})$$

and $\dot{b}_{t_i} = \frac{\sqrt{1-\rho^2}}{2\sqrt{V_{t_i}}}\prod_{k=1}^{i}\varphi_k$, *where*

$$\varphi_k = \mathrm{e}^{-\kappa(t_{k+1}-t_k)}\Big[1 + \frac{Z_k}{\sqrt{\frac{4\kappa\,\mathrm{e}^{-\kappa(t_{k+1}-t_k)}}{\sigma^2(1-\mathrm{e}^{-\kappa(t_{k+1}-t_k)})}V_{t_k}}}\Big] \text{ for all } k = 1, .., n,$$

and $(Z_k)_{1\leq k\leq n}$ are independent standard normal random variables.

Proof. The main of the proof is based on the properties of the Chi-squared distribution since the conditional distribution of V_t given V_u, $t > u$ follows a scaled non-central chi-squared distribution. According to Broadie and Kaya [15], $\chi^2_d(\lambda)$ a non central chi-squared random variable with d, $d > 1$ degrees of freedom and noncentrality parameter λ can be generated using χ^2_{d-1} a chi-squared random variable with $d-1$, $d > 1$ degrees of freedom and an independent standard normal random variable Z. We set

$$\chi^2_d(\lambda) = (Z + \sqrt{\lambda})^2 + \chi^2_{d-1}.$$

Then, we can write (4.14) as

$$V_{t_{i+1}} \overset{d}{=} \frac{\sigma^2(1 - \mathrm{e}^{-\kappa(t_i - t_{i+1})})}{4\kappa}\Big(\Big(Z_i + \sqrt{\frac{4\kappa\,\mathrm{e}^{-\kappa(t_{i+1}-t_i)}}{\sigma^2(1-\mathrm{e}^{-\kappa(t_{i+1}-t_i)})}V_{t_i}}\Big)^2 + \chi^2_{d-1}\Big), \quad \text{for} \quad d = \frac{4\kappa\theta}{\sigma^2} > 1.$$

First, we start by computing $\delta_{t_{i+1}}$. For $0 \le i \le n-1$, we have

$$\begin{aligned}\overline{X}^n_{t_{i+1}} &= \overline{X}^n_{t_i} + (\mu - \frac{1}{2}V_{t_i})(t_{i+1} - t_i) + \frac{\rho}{\sigma}\Big(V_{t_{i+1}} - V_{t_i} - \kappa\theta(t_{i+1} - t_i) + \kappa V_{t_i}(t_{i+1} - t_i)\Big) \\ &\quad + \sqrt{1-\rho^2}\int_{t_i}^{t_{i+1}} \sqrt{V_{t_i}} dW^s_s\end{aligned}$$

and

$$\frac{\partial V_{t_{i+1}}}{\partial V_{t_i}} = \mathrm{e}^{-\kappa(t_{i+1}-t_i)}\Big[1 + \frac{Z_i}{\sqrt{\frac{4\kappa\,\mathrm{e}^{-\kappa(t_{i+1}-t_i)}}{\sigma^2(1-\mathrm{e}^{-\kappa(t_{i+1}-t_i)})}V_{t_i}}}\Big].$$

Therefore,

$$\frac{\partial \overline{X}^n_{t_{i+1}}}{\partial V_0} = \frac{\partial \overline{X}^n_{t_i}}{\partial V_0} + \frac{\partial V_{t_i}}{\partial V_0}\Big(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\Big)h + \frac{\rho}{\sigma}\Big(\frac{\partial V_{t_{i+1}}}{\partial V_0} - \frac{\partial V_{t_i}}{\partial V_0}\Big) + \frac{\sqrt{1-\rho^2}}{2\sqrt{V_{t_i}}}\frac{\partial V_{t_i}}{\partial V_0}(W^s_{t_{i+1}} - W^s_{t_i}).$$

By induction, we get

$$\begin{aligned}\frac{\partial \overline{X}^n_{t_i}}{\partial V_0} &= \sum_{k=0}^{i-1}\Big(\frac{\partial \overline{X}^n_{t_{k+1}}}{\partial V_0} - \frac{\partial \overline{X}^n_{t_k}}{\partial V_0}\Big) \\ &= T(\frac{\rho\kappa}{\sigma} - \frac{1}{2}) + \frac{\rho}{\sigma}(\frac{\partial V_{t_i}}{\partial V_0} - 1) + \frac{\sqrt{1-\rho^2}}{2}\sum_{k=0}^{i-1}\frac{1}{\sqrt{V_{t_k}}}\frac{\partial V_{t_k}}{\partial V_0}(W^s_{t_{i+1}} - W^s_{t_i}).\end{aligned}$$

Moreover,

$$\frac{\partial V_{t_i}}{\partial V_0} = \prod_{k=1}^{i}\frac{\partial V_{t_k}}{\partial V_{t_{k-1}}} = \prod_{k=1}^{i}\varphi_k,$$

where $\varphi_k = \mathrm{e}^{-\kappa(t_{k+1}-t_k)}\Big[1 + \frac{Z_k}{\sqrt{\frac{4\kappa\,\mathrm{e}^{-\kappa(t_{k+1}-t_k)}}{\sigma^2(1-\mathrm{e}^{-\kappa(t_{k+1}-t_k)})}V_{t_k}}}\Big]$. Therefore,

$$\delta_{t_i} = \frac{\partial \overline{X}^n_{t_i}}{\partial V_0} = T(\frac{\rho\kappa}{\sigma} - \frac{1}{2}) + \frac{\rho}{\sigma}(\prod_{k=1}^{i}\varphi_k - 1) + \frac{\sqrt{1-\rho^2}}{2}\sum_{k=0}^{i-1}\frac{1}{\sqrt{V_{t_k}}}\prod_{k=1}^{i}\varphi_k(W^s_{t_{i+1}} - W^s_{t_i}).$$

Besides, as $b_{t_i} = \sqrt{1-\rho^2}\sqrt{V_{t_i}}$ then,

$$\dot{b}_{t_i} = \frac{\partial b_{t_i}}{\partial V_0} = \frac{\sqrt{1-\rho^2}}{2\sqrt{V_{t_i}}}\frac{\partial V_{t_i}}{\partial V_0} = \frac{\sqrt{1-\rho^2}}{2\sqrt{V_{t_i}}}\prod_{k=1}^{i}\varphi_k.$$

We complete the proof using formula (4.19) and (4.20). □

Note that the pathwise method using the Chi-squared distribution is easy to implement, however applying Theorem 4.6.1 to exchange expectation and derivative does not seem possible for the above scheme due to the product involving the term $\frac{\partial \overline{q}_i}{\partial V_0}$ in relation (4.26). More precisely, we are not able to control that term since we do not have the finiteness of all negative moments for the process V. However, we can rely on the flow process theory to overcome this issue.

In what follows, we provide an alternative approach to calculate the Vega of the Semi-exact scheme, which will be useful later on when applying Theorem 4.26. To do so, we consider the flow process $V(0,V_0) = (V_t(0,V_0), t \in [0,T])$, $V_0 > 0$ and with initial condition $V_0(0,V_0) = V_0$ satisfying

$$V_t(0,V_0) = V_0 + \int_0^t \kappa(\theta - V_u(0,V_0))du + \int_0^t \sigma\sqrt{V_u(0,V_0)}dW_u^v.$$

Recall that condition $\sigma^2 < 2\kappa\theta$ ensures that V remains strictly positive almost surely. Then, by [72, Theorem V.39]), the process $(V_t(0,V_0), t \in [0,T])$ is differentiable w.r.t. V_0.

Theorem 4.3.8. *Let $X_0 > D$ and consider the log-Heston scheme $(\overline{X}^n_{t_i})_{1\le i\le n}$ defined in* (4.13)*. We have*

$$\frac{\partial \overline{P}}{\partial V_0} = f'(\overline{X}^n_T)\frac{\partial \overline{X}^n_T}{\partial V_0}\prod_{i=0}^{n-1}(1-\overline{q}_i) - f(\overline{X}^n_T)\sum_{i=0}^{n-1}\left[\prod_{k=0,k\neq i}^{n-1}(1-\overline{q}_k)\frac{\partial \overline{q}_i}{\partial V_0}\right],$$

$$\begin{aligned}\frac{\partial \overline{q}_i}{\partial V_0} = \mathbb{1}_{(\overline{X}^n_{t_i},\overline{X}^n_{t_{i+1}}>D)}\overline{q}_i\Bigg[&\frac{-2\delta_{t_i}(\overline{X}^n_{t_{i+1}}-D)}{(1-\rho^2)V_{t_i}(0,V_0)h} + \frac{-2\delta_{t_{i+1}}(\overline{X}^n_{t_i}-D)}{(1-\rho^2)V_{t_i}(0,V_0)h}\\ &+ \frac{4(\overline{X}^n_{t_i}-D)(\overline{X}^n_{t_{i+1}}-D)\dot{b}_{t_i}}{(1-\rho^2)^{\frac{3}{2}}V_{t_i}^{\frac{3}{2}}(0,V_0)h}\Bigg]\end{aligned}$$

with

$$\delta_{t_i} = \frac{\partial \overline{X}^n_{t_i}}{\partial V_0} = T(\frac{\rho\kappa}{\sigma}-\frac{1}{2})+\frac{\rho}{\sigma}(\frac{\partial V_{t_i}(0,V_0)}{\partial V_0}-1)+\frac{\sqrt{1-\rho^2}}{2}\sum_{k=0}^{i-1}\frac{1}{\sqrt{V_{t_k}(0,V_0)}}\frac{\partial V_{t_k}(0,V_0)}{\partial V_0}(W^s_{t_{k+1}}-W^s_{t_k})$$

and $\dot{b}_{t_i} = \frac{\sqrt{1-\rho^2}}{2\sqrt{V_{t_i}(0,V_0)}}\frac{\partial V_{t_i}(0,V_0)}{\partial V_0}$, *where*

$$\frac{\partial V_{t_i}(0,V_0)}{\partial V_0} = \exp\Big(-\frac{\kappa}{2}t_i - (\frac{\kappa\theta}{2}-\frac{\sigma^2}{8})\int_0^{t_i}\frac{du}{V_u(0,V_0)} + \frac{1}{2}(\log V_{t_{i+1}}(0,V_0) - \log V_0)\Big) \text{ for all } i = 1,..,n. \tag{4.27}$$

Proof. First, we start by applying the Itô formula to the CIR process. We have for all $i \in \{0,...,n-1\}$

$$\log V_{t_i} = \log V_0 + \int_0^{t_i}\frac{1}{V_u}\kappa(\theta - V_u)du + \sigma\int_0^{t_i}\frac{dW^v_u}{\sqrt{V_u}} - \frac{\sigma^2}{2}\int_0^{t_i}\frac{1}{V_u}du. \tag{4.28}$$

To calculate $\frac{\partial V_{t_i}(0,V_0)}{\partial V_0}$, we consider the flow process $(V_t(0,V_0), t \in [0,T])$ given by

$$V_t(0,V_0) = V_0 + \int_0^t \kappa(\theta - V_u(0,V_0))du + \int_0^t \sigma\sqrt{V_u(0,V_0)}dW_u^v.$$

Then, by [72, Theorem V.39]), the process $(V_t(0,V_0), t\in[0,T])$ is differentiable w.r.t. V_0 and for all $i\in\{0,...,n\}$ we have

$$\frac{\partial V_{t_i}(0,V_0)}{\partial V_0} = \exp\Big(-\kappa t_i - \frac{\sigma^2}{8}\int_0^{t_i}\frac{du}{V_u(0,V_0)} + \frac{\sigma}{2}\int_0^{t_i}\frac{dW_u^v}{\sqrt{V_u(0,V_0)}}\Big).$$

By using (4.28) we get

$$\begin{aligned}\frac{\partial V_{t_i}(0,V_0)}{\partial V_0} &= \exp\Big(-\kappa t_i - \frac{\sigma^2}{8}\int_0^{t_i}\frac{du}{V_u(0,V_0)} - \frac{\kappa\theta}{2}\int_0^{t_i}\frac{du}{V_u(0,V_0)} + \frac{\kappa}{2}t_i\\ &\quad + \frac{\sigma^2}{4}\int_0^{t_i}\frac{du}{V_u(0,V_0)} + \frac{1}{2}(\log V_{t_i} - \log V_0)\Big)\\ &= \exp\Big(-\frac{\kappa}{2}t_i - (\frac{\kappa\theta}{2} - \frac{\sigma^2}{8})\int_0^{t_i}\frac{du}{V_u(0,V_0)} + \frac{1}{2}(\log V_{t_{i+1}}(0,V_0) - \log V_0)\Big).\end{aligned}$$

Similarly, to the proof of the Theorem 4.3.7 we have

$$\frac{\partial \overline{X}^n_{t_{i+1}}}{\partial V_0} = T(\frac{\rho\kappa}{\sigma} - \frac{1}{2}) + \frac{\rho}{\sigma}(\frac{\partial V_{t_i}(0,V_0)}{\partial V_0} - 1) + \frac{\sqrt{1-\rho^2}}{2}\sum_{k=0}^{i-1}\frac{1}{\sqrt{V_{t_k}(0,V_0)}}\frac{\partial V_{t_k}(0,V_0)}{\partial V_0}(W^s_{t_{i+1}} - W^s_{t_i}).$$

Besides, as $b_{t_i} = \sqrt{1-\rho^2}\sqrt{V_{t_i}(0,V_0)}$ then,

$$\dot{b}_{t_i} = \frac{\partial b_{t_i}}{\partial V_0} = \frac{\sqrt{1-\rho^2}}{2\sqrt{V_{t_i}}}\frac{\partial V_{t_i}(0,V_0)}{\partial V_0}.$$

We complete the proof using formula (4.19) and (4.20). □

Theorem 4.3.9. *Let $V_0\in(V_{0,min}, V_{0,max})$ with $V_{0,min}>0$ and let $2\kappa\theta>\sigma^2$. Then, for $\overline{P}$ defined in* (4.18)*, we have*

$$\frac{\partial \mathbb{E}(\overline{P})}{\partial V_0} = \mathbb{E}\Big(\frac{\partial \overline{P}}{\partial V_0}\Big).$$

Proof. At first, let us observe that the first term of $\frac{\partial\overline{P}}{\partial V_0}$ is bounded by $\frac{\partial \overline{X}^n_T}{\partial V_0}$ since f' is bounded and the product is less than 1 then we get

$$\frac{\partial \overline{X}^n_T}{\partial V_0} = T(\frac{\rho\kappa}{\sigma} - \frac{1}{2}) + \frac{\rho}{\sigma}(\frac{\partial V_T(0,V_0)}{\partial V_0} - 1) + \frac{\sqrt{1-\rho^2}}{2}\sum_{k=0}^{n-1}\frac{1}{\sqrt{V_{t_k}(0,V_0)}}\frac{\partial V_{t_k}(0,V_0)}{\partial V_0}(W^s_{t_{k+1}} - W^s_{t_k}).$$

By (4.27) $x\mapsto V_{t_i}(0,x)$ is increasing so that as $\kappa\theta\geq\frac{\sigma^2}{4}$, we get

$$\frac{\partial V_{t_i}(0,V_0)}{\partial V_0} \leq \frac{\sqrt{V_{t_i}(0,V_{0,max})}}{\sqrt{V_{0,min}}}.$$

Then, there exist positive constants C_1, C_2 and C_3 that may change from line to line, such that

$$\Big|\frac{\partial \overline{X}^n_T}{\partial V_0}\Big| \leq C_1 + C_2\Big|\sqrt{V_{t_i}(0,V_{0,max})}\Big| + C_3\sum_{k=0}^{n-1}\Big|\frac{\sqrt{V_{t_k}(0,V_{0,max})}}{\sqrt{V_{t_k}(0,V_{0,min})}}(W^s_{t_{k+1}} - W^s_{t_k})\Big|,$$

and

$$\mathbb{E}\Big|\frac{\partial \overline{X}_T^n}{\partial V_0}\Big| \leq C_1 + C_2\mathbb{E}\Big|\sqrt{V_{t_i}(0,V_{0,max})}\Big| + C_3\sum_{k=0}^{n-1}\mathbb{E}\Big|\frac{\sqrt{V_{t_k}(0,V_{0,max})}}{\sqrt{V_{t_k}(0,V_{0,min})}}(W^s_{t_{k+1}} - W^s_{t_k})\Big|.$$

Using the independence between the processes W^s and V we get

$$\begin{aligned}\mathbb{E}\Big|\frac{\partial \overline{X}_T^n}{\partial V_0}\Big| &\leq C_1 + C_2\mathbb{E}\Big|\sqrt{V_{t_i}(0,V_{0,max})}\Big| + C_3\sum_{k=0}^{n-1}\mathbb{E}\Big|\frac{\sqrt{V_{t_k}(0,V_{0,max})}}{\sqrt{V_{t_k}(0,V_{0,min})}}\Big|\mathbb{E}\Big|W^s_{t_{k+1}} - W^s_{t_k}\Big| \\ &\leq C_1 + C_2\mathbb{E}\Big|\sqrt{V_{t_i}(0,V_{0,max})}\Big| + C_3\sqrt{\frac{T}{n}}\sum_{k=0}^{n-1}\mathbb{E}\Big|\frac{\sqrt{V_{t_k}(0,V_{0,max})}}{\sqrt{V_{t_k}(0,V_{0,min})}}\Big|.\end{aligned}$$

By Hölder's inequality, for $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$, we get

$$\mathbb{E}\Big|\frac{\partial \overline{X}_T^n}{\partial V_0}\Big| \leq C_1 + C_2\mathbb{E}\Big|\sqrt{V_{t_i}(0,V_{0,max})}\Big| + C_3\sqrt{\frac{T}{n}}\sum_{k=0}^{n-1}\Big(\mathbb{E}\Big|\sqrt{V_{t_k}(0,V_{0,max})}\Big|^p\Big)^{\frac{1}{p}}\Big(\mathbb{E}\Big|\frac{1}{\sqrt{V_{t_k}(0,V_{0,min})}}\Big|^q\Big)^{\frac{1}{q}}.$$

Then, using Lemma 3.3.1 we easily get $\mathbb{E}\Big|\frac{\partial \overline{X}_T^n}{\partial V_0}\Big| < \infty$ for $1 < q < \frac{2\kappa\theta}{\sigma^2}$.

Now, concerning the second term of $\frac{\partial \overline{P}}{\partial V_0}$ we only treat the third term therein, since the two first ones can be treated easily using similar arguments as above. So, we have

$$\begin{aligned}&\Big|f(\overline{X}_T^n)\sum_{i=0}^{n-1}\frac{2(\overline{X}^n_{t_i} - D)(\overline{X}^n_{t_{i+1}} - D)}{(1-\rho^2)V^2_{t_i}(0,V_0)h}\frac{\partial V_{t_i}(0,V_0)}{\partial V_0}\Big| \\ &\qquad\leq \Big|f(\overline{X}_T^n)\sum_{i=0}^{n-1}\frac{2(\overline{X}^n_{t_i} - D)(\overline{X}^n_{t_{i+1}} - D)}{h(1-\rho^2)V^2_{t_i}(0,V_{0,min})}\frac{\sqrt{V_{t_i}(0,V_{0,max})}}{\sqrt{V_{0,min}}}\Big| \\ &\qquad\leq C(1+|\overline{X}_T^n|)\sum_{i=0}^{n-1}\Big|(\overline{X}^n_{t_i} - D)(\overline{X}^n_{t_{i+1}} - D)\frac{\sqrt{V_{t_i}(0,V_{0,max})}}{V^2_{t_i}(0,V_{0,min})}\Big|,\end{aligned}$$

where C is a positive constant that may change from line to line. Then, we use Hölder's inequality, for $p_1, p_2, p_3 > 1$ with $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 1$ yields

$$\begin{aligned}&\mathbb{E}\Big|f(\overline{X}_T^n)\sum_{i=0}^{n-1}\frac{2(\overline{X}^n_{t_i} - D)(\overline{X}^n_{t_{i+1}} - D)}{(1-\rho^2)V^2_{t_i}(0,V_0)h}\frac{\partial V_{t_i}(0,V_0)}{\partial V_0}\Big| \leq \\ &C\sum_{i=0}^{n-1}\Big(1 + (\mathbb{E}|\overline{X}_T^n|^{p_1})^{\frac{1}{p_1}}\Big)\Big(\mathbb{E}\Big|(\overline{X}^n_{t_i} - D)(\overline{X}^n_{t_{i+1}} - D)\Big|^{p_2}\Big)^{\frac{1}{p_2}}\Big(\mathbb{E}\Big|\frac{\sqrt{V_{t_i}(0,V_{0,max})}}{V^2_{t_i}(0,V_{0,min})}\Big|^{p_3}\Big)^{\frac{1}{p_3}}.\end{aligned}$$

Again by using Hölder's inequality, for $q_1, q_2 > 1$ with $\frac{1}{q_1} + \frac{1}{q_2} = \frac{1}{p_3}$ yields and by the finiteness of all positive moment of the process $\overline{X}^n$ we get

$$\begin{aligned}&\mathbb{E}\Big|f(\overline{X}_T^n)\sum_{i=0}^{n-1}\frac{2(\overline{X}^n_{t_i} - D)(\overline{X}^n_{t_{i+1}} - D)}{(1-\rho^2)V^2_{t_i}(0,V_0)h}\frac{\partial V_{t_i}(0,V_0)}{\partial V_0}\Big| \leq \\ &C\sum_{i=0}^{n-1}\Big(\mathbb{E}\Big|\sqrt{V_{t_i}(0,V_{0,max})}\Big|^{p_3q_1}\Big)^{\frac{1}{p_3q_1}}\Big(\mathbb{E}\Big|V^2_{t_i}(0,V_{0,min})\Big|^{-p_3q_2}\Big)^{\frac{1}{p_3q_2}}.\end{aligned}$$

We complete the proof by applying Lemma 3.3.1 with choosing $1 < p_3q_2 < \frac{2\kappa\theta}{\sigma^2}$. □

Note that the result of the above theorem is only valid for our Semi-exact log-Heston scheme (4.18) because the other approximation schemes (4.17) and (4.16) involve respectively negative powers of $\hat{V}_{t_i}^n$ and $\hat{Y}_{t_i}^n$, and up to our knowledge we lack in the literature results regarding the negative moments of these approximation schemes. This is another main advantage of our Semi-exact log-Heston scheme, making it more suitable for justifying the computation of pathwise sensitivities for barrier options with the Brownian bridge technique.

4.4 Numerical results using GPR

We consider (X, Y) a training set with $X = (X_i)_{i=1,\dots,n} = (K^{(i)}, V_0^{(i)}, \sigma, T^{(i)}, r^{(i)}, \theta^{(i)}, \kappa^{(i)}, \rho^{(i)})$ being an $n \times 8$ matrix and Y an $n \times 1$ vector representing the price simulated by the Monte Carlo method. We set $X_* = (K, V_0, \sigma, T, r, \theta, \kappa, \rho)$ a testing point, then the estimated price $f(X_*)$ of a Down and Out barrier call option is given by (4.3) with $m(X_*) = 0$.

The derivative with respect to X_* can be computed using the following formula:

$$\partial_{X_*}\mathbb{E}[f_*|X, Y, X_*] = \partial_{X_*} m(X_*) + \partial_{X_*} K_{X_*,X}\alpha, \tag{4.29}$$

with $\partial_{X_*} K_{X_*,X} = \frac{1}{\ell^2}(X - X_*)K_{X_*,X}$ and $\alpha = [K_{X,X} + \sigma_n^2 I]^{-1}Y$, (for more details see [24]). We obtain the second order sensitivities by differentiating once again with respect to X_*. We notice that α is already computed once the training procedure for pricing is done. It is computed by the Cholesky factorization method of $[K_{X,X} + \sigma_n^2 I]$ with a complexity of order $O(n^3)$(for more details see [29]).

In what follows, we develop a numerical study in order to test the GPR methodology and its efficiency under the Heston model. The study focuses on Down and Out (D-O) call barrier option. First, D-O option price will be estimated using GPR algorithm. Then, we will compute the associated Δ and ν using (4.29).

4.4.1 Down and Out pricing using GPR

The input values $X = (X_i)_{1\le i\le n} = (K^{(i)}, V_0^{(i)}, \sigma, T^{(i)}, r^{(i)}, \theta^{(i)}, \kappa^{(i)}, \rho^{(i)})$, for the training set with $n = 5000$ is obtained via a uniform random sampling, where the parameters range can be found in Table 4.1.

Parameter	Minimum	Maximum
K	0.6	1.4
κ	1.4	2.6
V_0	0.01	0.16
σ	0.35	0.75
T	$11M$	$1Y$
r	0.015	0.025
θ	0.01	0.16
ρ	-0.85	-0.45

Table 4.1: Training parameters for D-O barrier options under Heston model.

The output values $Y = (Y_i)_{1\leq i\leq n}$ used for training are obtained by approximating the D-O barrier call price for the model specifications given by $X = (X_i)_{1\leq i\leq n}$ and the spot price fixed to $S_0 = 1$, where the price is effectively computed using Monte Carlo simulations with 100000 sample paths and a discretization time step equal to $N = 250$ for the schemes introduced in (4.11) ("Drift implicit Milstein"), (4.12) ("Drift implicit Euler") and (4.13) ("Semi-exact") with a call payoff function $f : x \in \mathbb{R} \mapsto f(x) = (\exp(x) - K)_+$. We fix the spot $S_0 = \exp(X_0) = 1$ and we take two possible values for the barrier $B = \exp(D) \in \{0.75, 0.85\}$. Besides, we also train the GPR method using again the approximation scheme (4.12) without the Brownian bridge technique but rather with keeping the indicator function with a discrete running minimum

$$\left(\exp\left(\tilde{X}_T^n\right) - K\right)_+ \mathbb{1}_{\{\min_{1\leq i\leq n} \tilde{X}_{t_i}^n > D\}} \tag{4.30}$$

as tested in [26]. This method is referred to as "MC with indicator". Once the training phase is complete, the model is ready for prediction. We tested the model on 10000 input parameters. The test points are constructed uniformly over small parameter ranges (see Table 4.2).

Parameter	Minimum	Maximum
K	0.5	1.5
κ	1.5	2.5
V_0	0.02	0.16
σ	0.4	0.7
T	$11M$	$1Y$
r	0.015	0.025
θ	0.02	0.16
ρ	-0.8	-0.5

Table 4.2: Test parameters for D-O barrier options under the Heston model

The pricing performance is measured in terms of maximum and average absolute errors,

$$MAE = \max_{1\leq i\leq n} \{|E_{Benchmark}(i) - E_{GPR}(i)|\},$$

and

$$AAE = \frac{1}{n}\sum_{i=1}^{n} |E_{Benchmark}(i) - E_{GPR}(i)|,$$

where the benchmark price $E_{Benchmark}$ corresponds to the price calculated using the drift implicit Euler scheme (4.16) with a Monte Carlo sample size of $N = 200000$, since it gave us a better performance for pricing then the other schemes. Besides, the quality of the model fit is assessed by the mean square error(RMSE).

In Table 4.3 and 4.4, we observe that the GPR trained with the "MC indicator" method (4.30) is slightly less accurate than the other schemes trained using the Brownian bridge technique. Both tables 4.3 and 4.4 show that there is a reasonable fit between the GPR pricing method and the price. The GPR constructed by the various approximation methods show that the MAE remains reasonable and both AAE and RMSE are very minimal.

	Semi-exact	Drift implicit Milstein	Drift implicit Euler	MC indicator
MAE	0.0137	0.0134	0.0136	0.0147
AAE	0.0025	0.0020	0.0018	0.0037
RMSE	0.0031	0.0026	0.0024	0.0046

Table 4.3: The GPR performance for a D-O barrier option with $B = 0.75$

	Semi-exact	Drift implicit Milstein	Drift implicit Euler	MC indicator
MAE	0.0211	0.0167	0.0147	0.0263
AAE	0.0033	0.0063	0.0018	0.0033
RMSE	0.0041	0.0028	0.0024	0.0081

Table 4.4: The GPR performance for a D-O barrier option with $B = 0.85$

• **Semi-exact Heston model** In Figure 4.1 and 4.2, we observe that the barrier option prices predicted by the GPR method trained by the Semi-exact is acceptable and the accuracy is reasonable.

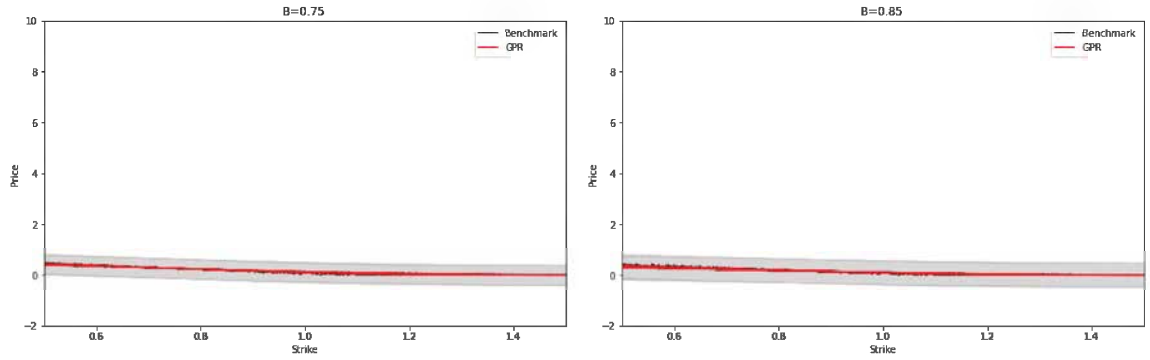


Fig 4.1. GPR: Semi-exact Simulation Method vs Benchmark Method with two barrier 0.75 and 0.85.

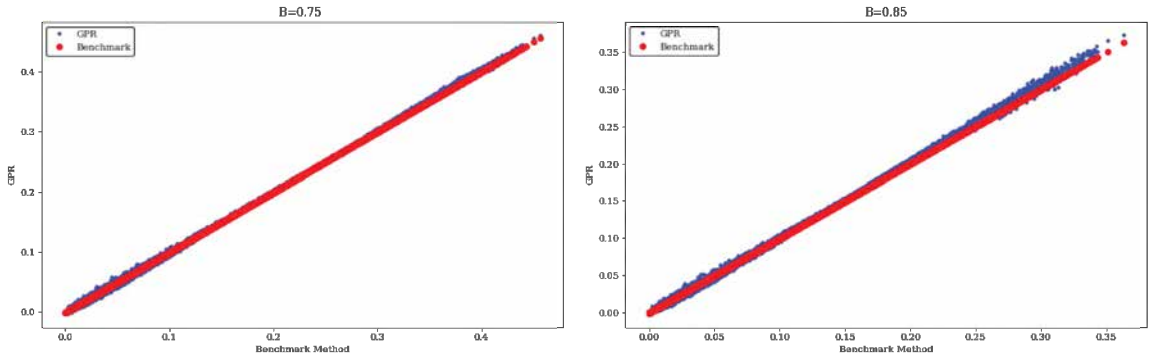


Fig 4.2. Benchmark Method vs GPR predicted for pricing of Down and Out barrier option.

• **Drift implicit Milstein** In Figure 4.3 and 4.4 we provide similar plot except that the GPR is trained using the drift implicit Milstein scheme.

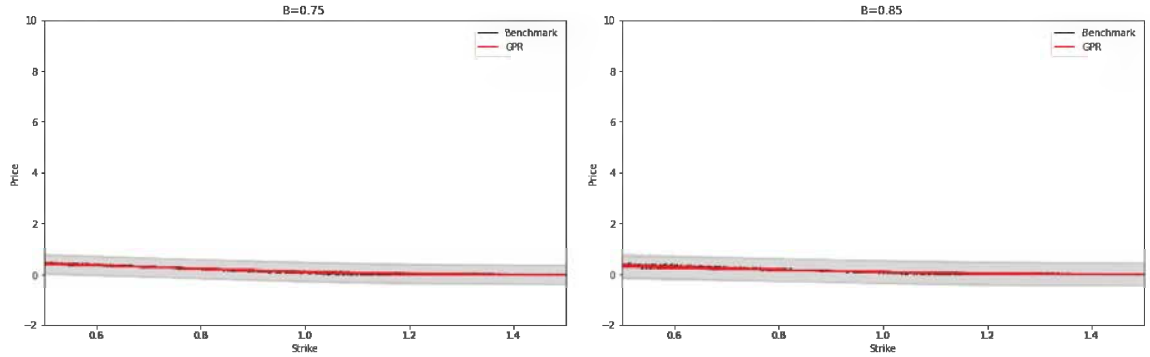


Fig 4.3. GPR: Drift implicit Milstein scheme vs Benchmark Method with two barrier 0.75 and 0.85.

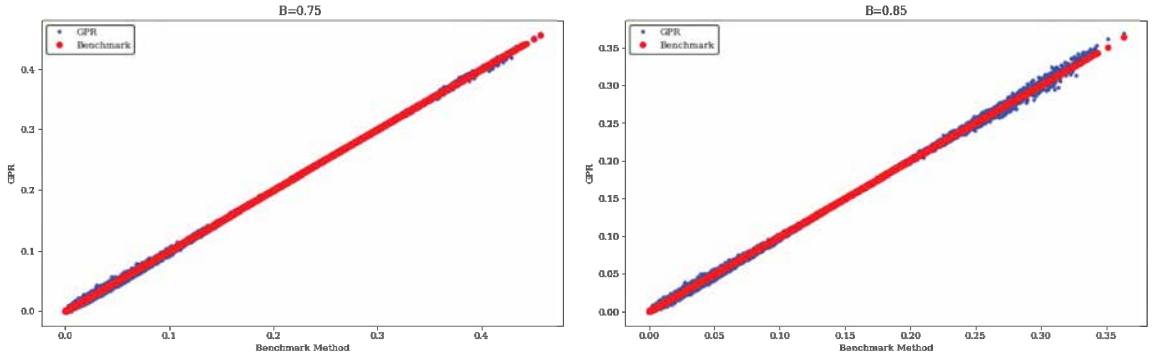


Fig 4.4. Benchmark Method vs GPR predicted for pricing of Down and Out barrier option.

• **Drift implicit Euler** Similarly to the above schemes, we plot the GPR pricing prediction using the drift implicit Euler scheme compared to the Benchmark price.

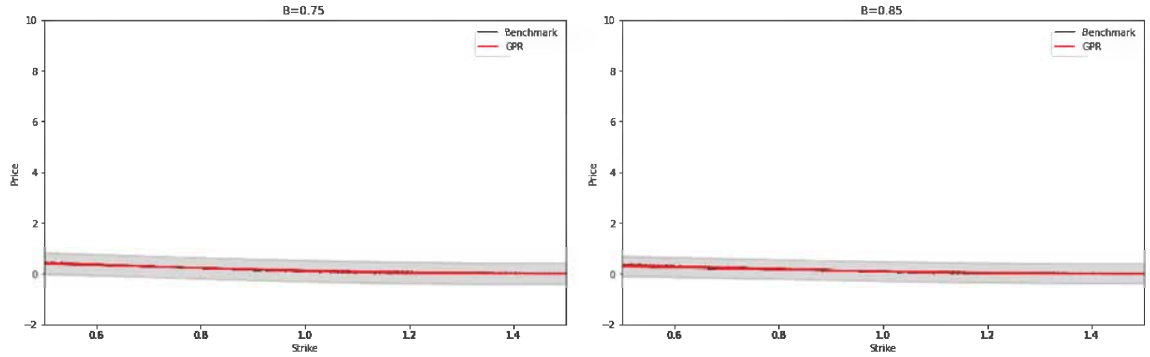


Fig 4.5. GPR: Drift implicit Euler scheme vs Benchmark Method with two barrier 0.75 and 0.85.

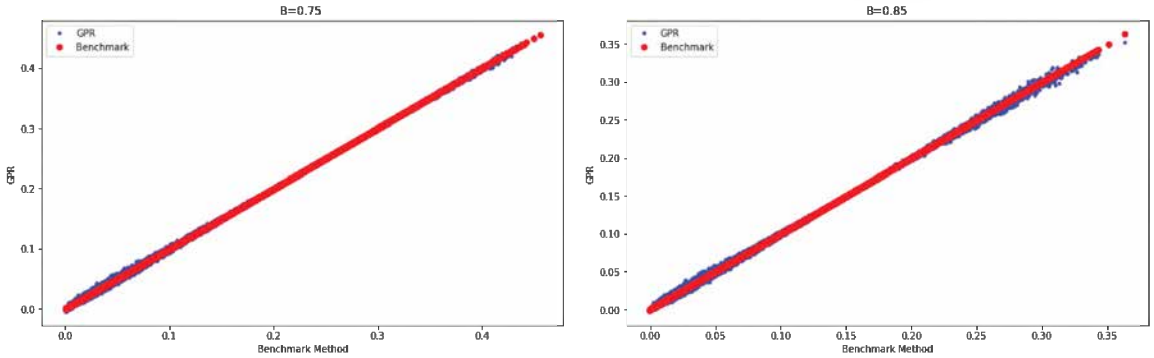


Fig 4.6. Benchmark Method vs GPR predicted for pricing of Down and Out barrier option.

Remark 4.4.1. *If we use the MC indicator* (4.30) *to train the GPR, we can see from Figure 4.7 that the match between the GPR prices and the benchmark ones is not perfect. Indeed, as shown in tables 4.3 and 4.4 the method is slightly imprecise compared to the other schemes with Brownian bridge technique.*

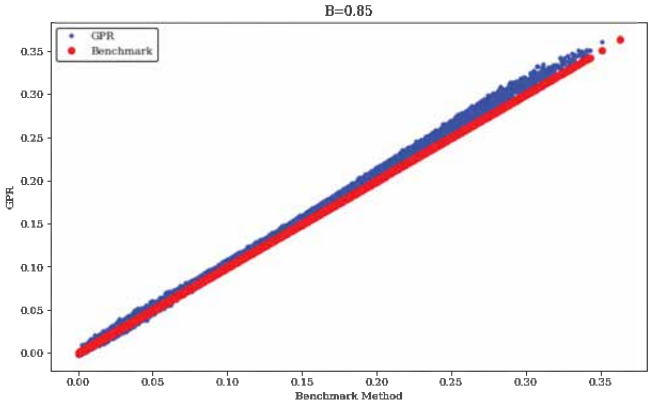


Fig 4.7. MC Benchmark vs MC indicator (4.30) for the price of a Down and Out barrier option.

4.4.2 Delta Sensitivity

One of the GPR method advantages is that once it is trained for the pricing, the sensitivities computation becomes straightforward. The formula defined in (4.29) can be used for any differentiable kernel to calculate price sensitivities without any extra training step. This allows us to reduce significantly the computation time. In this section, we present the results obtained for Δ calculation of a Down and Out barrier options under the Heston model using the GPR methodology.

To do so, we consider a couple (X, Y), where the input $X = (X_i)_{i=1,\dots,n} = (S_0^{(i)})$ is an $n \times 1$ vector containing spot values randomly selected between 70 and 130, and the output Y is an $n \times 1$ vector representing the price simulated by the Monte Carlo method. The rest of the parameters are taken as constant as given in Table 4.5.

Parameter	Value
K	100
κ	2
V_0	0.1
σ	0.5
T	$1Y$
r	0.02
θ	0.1
ρ	-0.5
B	80

Table 4.5: Parameters values for Down and Out barrier options under Heston model Δ Computation

In our study, we used a set of 5000 prices to train the GPR model. In this procedure, we do not need any additional train since the GPR was already trained during the price calculations process. Then, we compare the Delta GPR with the pathwise MC methods developed in Section 4.3. Table 4.6, confirms the good performance of the GPR method for computing the Delta when compared with the pathwise MC methods.

	Semi-exact	Drift implicit Milstein	Drift implicit Euler
MAE	0.4323	0.2598	0.2581
AAE	0.0175	0.0157	0.0148
RMSE	0.0438	0.0346	0.0326

Table 4.6: Difference between the GPR Delta estimate and MC formula Greek for Heston model

the predicted Delta GPR curve and the pathwise MC Delta for each considered scheme as a function of spot values. We note that the GPR Delta is generally close to the real values, with a deviation in the region where the option becomes illiquid (when the option is in the money or out of the money). Additionally, we note a deviation when the spot value is close to the barrier value. This is probably due to the discontinuity of the Down and Out option Delta around the barrier value.

• **Semi-exact Heston model** In Figure 4.8, we plot the Delta GPR prices together with the pathwise MC Delta using the Semi-exact scheme as function of the spot. In Figure 4.9, we plot the pathwise MC Delta versus the Delta GPR predictions for the Down and Out barrier options. We see that our scatter plot of 10000 predictions of Delta with a GPR model hardly deviates from this straight line.

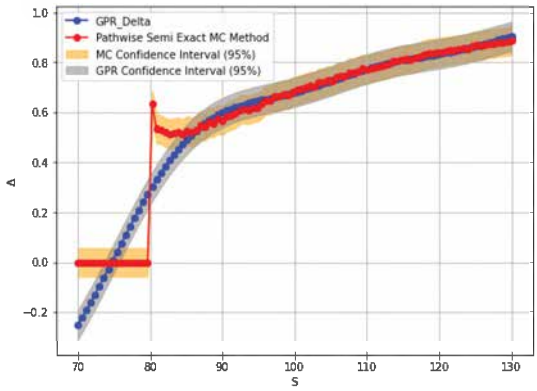


Fig 4.8. Delta: Pathwise MC method using the Semi-exact scheme and the GPR Method.

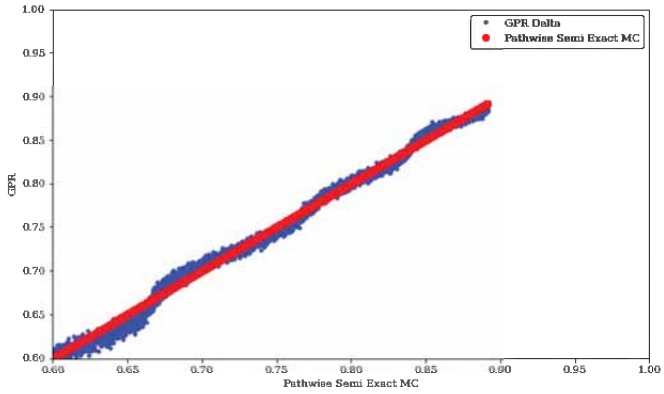


Fig 4.9. Delta pathwise MC method with the Semi-exact scheme vs Delta GPR prediction.

• **Drift implicit Milstein** We provide similar plot with the pathwise MC using the drift implicit Milstein scheme.

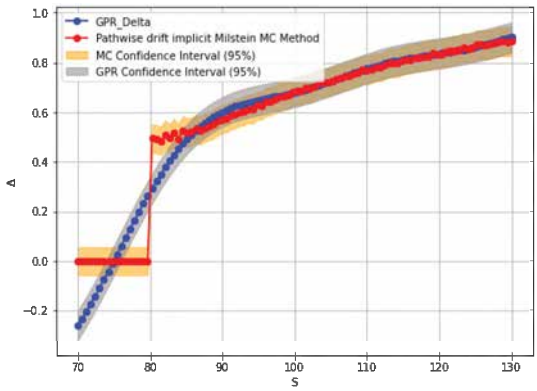


Fig 4.10. Delta: Pathwise MC method using the drift implicit Milstein scheme and the GPR Method.

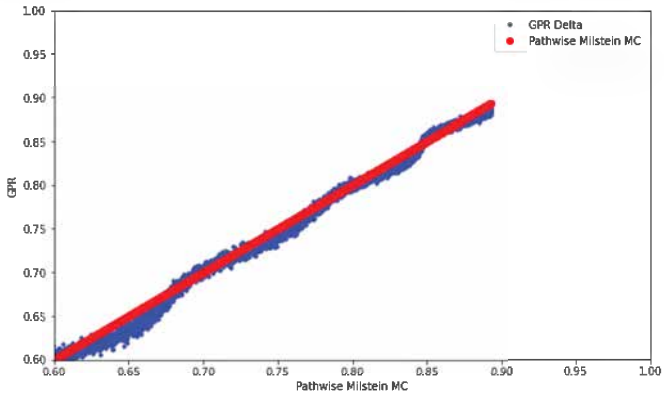


Fig 4.11. Delta pathwise MC method with the drift implicit Milstein scheme vs Delta GPR prediction.

• **Drift implicit Euler** Similarly to the above schemes, we plot the Delta GPR prediction using the drift implicit Euler scheme compared to Delta pathwise MC method.

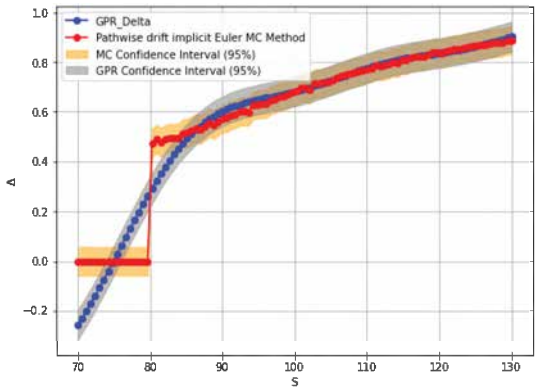


Fig 4.12. Delta: Pathwise MC method using the drift implicit Euler scheme and the GPR Method.

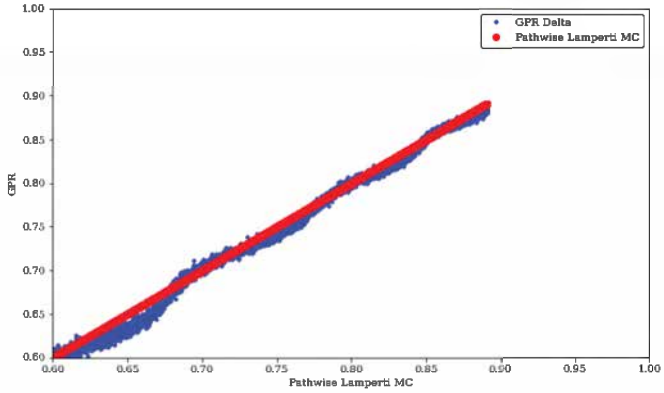


Fig 4.13. Delta pathwise MC method with the drift implicit Euler scheme vs Delta GPR prediction.

4.4.3 Vega Sensitivity

In this section, we examine the computation of Vega using the GPR methodology. Similarly to Delta, we will use formula (4.29). We consider the couple (X, Y) where the input $X = (X_i)_{i=1,\dots,n} = (V_0^{(i)})$ is an $n \times 1$ vector containing volatility initial values sampled uniformly between 0.01 and 0.16, and the output Y is an $n \times 1$ vectors representing the price for a Down and Out barrier option under Heston model simulated by the pathwise Monte Carlo methods developed in Section 4.3. In our case, we use a training set of size 5000. In order to compare the Vega obtained by the GPR, a set of size 10000 has been used. The parameter ranges used for training are giving in Table 4.7.

Parameter	Value
S_0	100
K	100
κ	2
σ	0.5
T	$1Y$
r	0.02
θ	0.1
ρ	-0.5
B	80

Table 4.7: Parameters values for Down and Out barrier options under Heston model ν Computation

Table 4.8 shows that the Vega does not fit very well the pathwise MC Vega computed with the different approximation schemes and this can be shown by the larger average error.

	Semi-exact	Drift implicit Milstein	Drift implicit Euler
MAE	4.4398	2.1241	2.2637
AAE	0.6055	0.5480	0.3615
RMSE	0.7994	0.6454	0.4633

Table 4.8: Difference between the GPR Vega estimate and pathwise sensitivity Greek for Heston model

• **Semi-exact Heston model** In this paragraph, we illustrate the quality of Vega computation obtained with the GPR method by comparing it with the pathwise MC method using the Semi-exact approximation scheme.

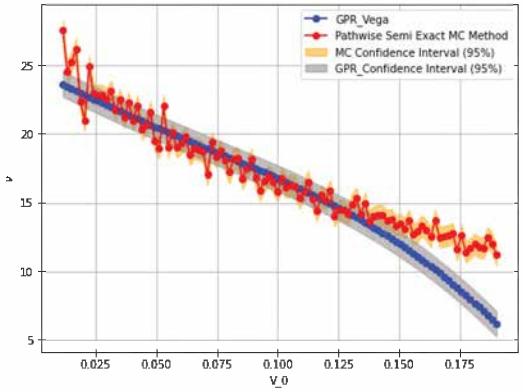


Fig 4.14. Vega: Pathwise MC method using the Semi-exact scheme and the GPR Method.

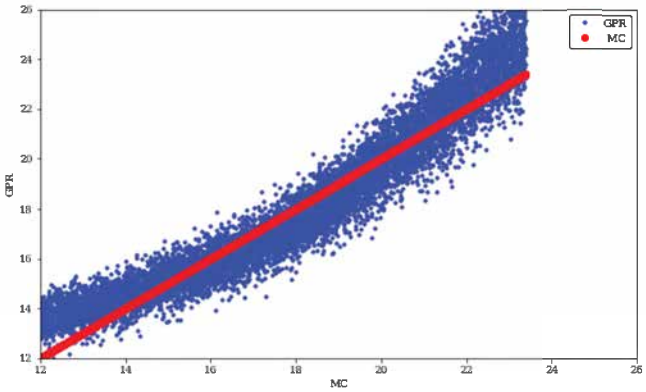


Fig 4.15. MC Vega using Semi-exact scheme vs GPR predicted Vega for Down and Out barrier option.

• **Drift implicit Milstein** We provide similar plot except that the Vega GPR is trained using the drift implicit Milstein scheme.

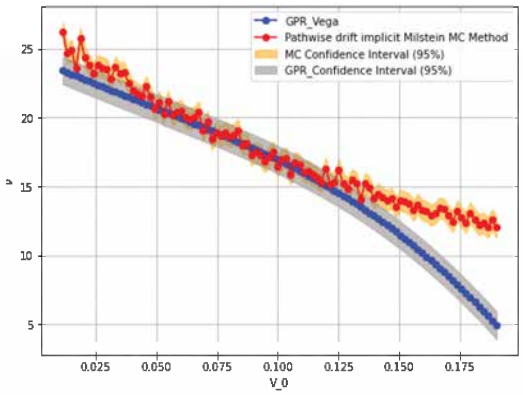


Fig 4.16. Vega: Pathwise MC method using the drift implicit Milstein scheme and the GPR Method.

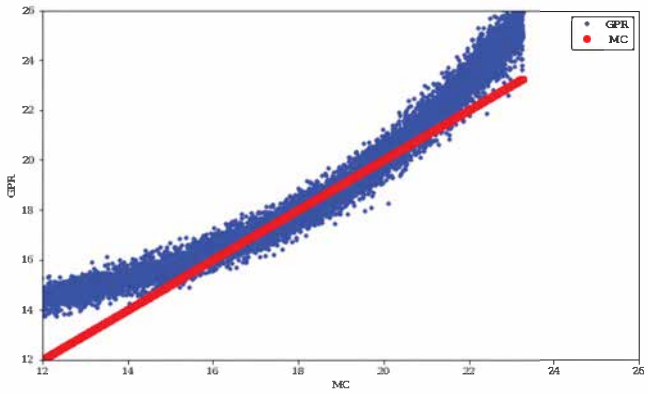


Fig 4.17. Vega pathwise MC method with the drift implicit Milstein scheme vs Vega GPR prediction.

Drift implicit Euler Similarly to the above schemes, we plot the Vega GPR pricing prediction using the drift implicit Euler scheme compared to the Vega pathwise MC method.

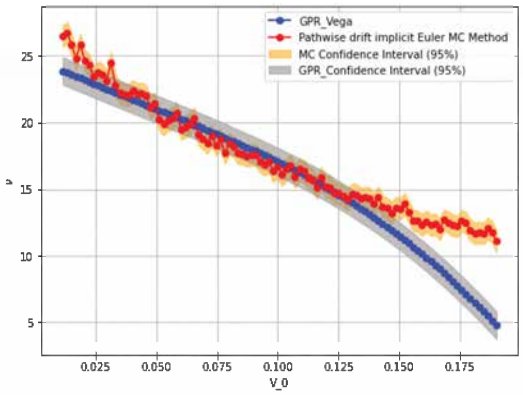


Fig 4.18. Vega: Pathwise MC method using the drift implicit Euler scheme and the GPR Method.

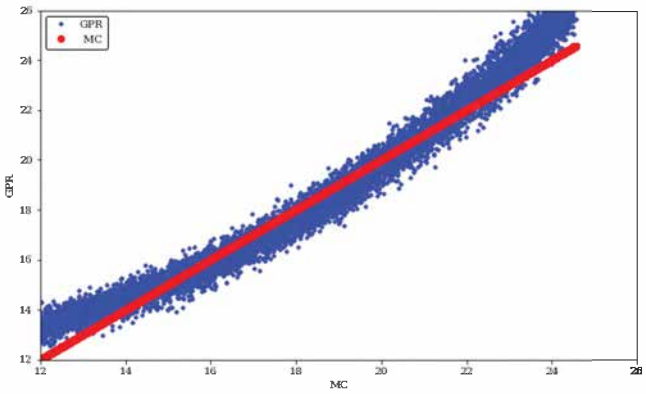


Fig 4.19. Vega pathwise MC method with the drift implicit Euler scheme vs Vega GPR prediction.

4.5 Conclusions and future work

In this paper, we have tested the efficiency of the GPR method. We illustrated the use of this technique for pricing Down and Out barrier option under the Heston model. We proved that when using the Brownian bridge interpolation in the training phase the GPR performs better than using the indicator. For the Greeks computation, we developed new formulas to compute the Delta and Vega Down and Out barrier options under the log-Heston model then compare it to GPR. However, we found that the Delta is not accurate when the Spot price is close to the barrier. The GPR method seems also not very accurate for the Vega computation. Therefore, the GPR method can be reliable when pricing barrier options but we need to be more careful when using it for computing sensitivities.
The challenge for the future works is to try to improve the accuracy of GPR by looking for an appropriate kernel so that it captures for example values close to the barrier, and a fast matrix inversion of the corresponding covariance matrix.

4.6 Appendix

Theorem 4.6.1. *Interchanging differentiation and expectation (See Pagès [69, Theorem 2.2]) Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, let I be a nontrivial interval of $\mathbb{R}$. Let $\varphi : I \times \Omega \to \mathbb{R}$ be a $\mathcal{B}or(I) \otimes \mathcal{A}$-measurable function. If the function φ satisfies:*

1. *for every $x \in I$, the random variable $\varphi(x, .) \in L^1_{\mathbb{R}}(\Omega, \mathcal{A}, \mathbb{P})$,*
2. *$\mathbb{P}(d\omega) - a.s.$, $\frac{\partial \varphi}{\partial x}(x, \omega)$ exists at every $x \in I$*
3. *There exists a $Y \in L^1_{\mathbb{R}}(\Omega, \mathcal{A}, \mathbb{P})$ such that, for every $x \in I$,*

$$\mathbb{P}(d\omega) - a.s. \quad \Big|\frac{\partial \varphi(x, \omega)}{\partial x}\Big| \leq Y(\omega),$$

then, the function $\Phi(x) := \mathbb{E}\varphi(x, .)$ is defined and differentiable at every $x \in I$, with derivative

$$\Phi(x) = \mathbb{E}\Big(\frac{\partial \varphi}{\partial x}(x, .)\Big).$$

Bibliography

[1] Milton Abramowitz and Irene A. Stegun, editors. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York; John Wiley & Sons, Inc., New York, 1984. Reprint of the 1972 edition, Selected Government Publications.

[2] Mohamed Ben Alaya, Ahmed Kebaier, and Djibril Sarr. Deep calibration of interest rates model. *arXiv preprint arXiv:2110.15133*, 2021.

[3] Aurélien Alfonsi. On the discretization schemes for the cir (and bessel squared) processes. *Monte Carlo Methods and Applications*, 11(4):355–384, 2005.

[4] Aurélien Alfonsi. High order discretization schemes for the cir process: application to affine term structure and heston models. *Mathematics of Computation*, 79(269):209–237, 2010.

[5] Aurélien Alfonsi. Strong order one convergence of a drift implicit euler scheme: Application to the cir process. *Statistics & Probability Letters*, 83(2):602–607, 2013.

[6] Martin Altmayer and Andreas Neuenkirch. Discretising the heston model: an analysis of the weak convergence rate. *IMA Journal of Numerical Analysis*, 37(4):1930–1960, 2017.

[7] Nachman Aronszajn. Theory of reproducing kernels. *Transactions of the American mathematical society*, 68(3):337–404, 1950.

[8] Paolo Baldi. Exact asymptotics for the probability of exit from a domain and applications to simulation. *Ann. Probab.*, 23(4):1644–1670, 1995.

[9] Subhasish Basak, Sébastien Petit, Julien Bect, and Emmanuel Vazquez. Numerical issues in maximum likelihood parameter estimation for gaussian process interpolation. In *International Conference on Machine Learning, Optimization, and Data Science*, pages 116–131. Springer, 2021.

[10] Mohamed Ben Alaya, Kaouther Hajji, and Ahmed Kebaier. Importance sampling and statistical Romberg method. *Bernoulli*, 21(4):1947–1983, 2015.

[11] Mohamed Ben Alaya, Kaouther Hajji, and Ahmed Kebaier. Adaptive importance sampling for multilevel monte carlo euler method. *Stochastics*, 0(0):1–25, 2022.

[12] Mohamed Ben Alaya and Ahmed Kebaier. Asymptotic behavior of the maximum likelihood estimator for ergodic and nonergodic square-root diffusions. *Stoch. Anal. Appl.*, 31(4):552–573, 2013.

[13] Abdel Berkaoui, Mireille Bossy, and Awa Diop. Euler scheme for sdes with non-lipschitz diffusion coefficient: strong convergence. *ESAIM: Probability and Statistics*, 12:1–11, 2008.

[14] Christopher M Bishop and Nasser M Nasrabadi. *Pattern recognition and machine learning*, volume 4. Springer, 2006.

[15] Mark Broadie and Özgür Kaya. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations research*, 54(2):217–231, 2006.

[16] Sylvestre Burgos. The computation of greeks with multilevel monte carlo. *PhD thesis, University of Oxford*, 2014.

[17] Peter Carr, Andrey Itkin, and Dmitry Muravey. Semi-analytical pricing of barrier options in the time-dependent heston model. *The Journal of Derivatives*, 30(2):141–171, 2022.

[18] Ching-Sung Chou and Hsien-Jen Lin. Some properties of CIR processes. *Stochastic Anal. Appl.*, 24(4):901–912, 2006.

[19] John MC Clark and RJ Cameron. The maximum rate of convergence of discrete approximations for stochastic differential equations. In *Stochastic Differential Systems Filtering and Control: Proceedings of the IFIP-WG 7/1 Working Conference Vilnius, Lithuania, USSR, Aug. 28–Sept. 2, 1978*, pages 162–171. Springer, 2005.

[20] John Cox. Notes on option pricing i: Constant elasticity of variance diffusions. *Unpublished note, Stanford University, Graduate School of Business*, 1975.

[21] John C Cox, Jonathan E Ingersoll Jr, and Stephen A Ross. A theory of the term structure of interest rates. In *Theory of valuation*, pages 129–164. World Scientific, 2005.

[22] John C Cox and Stephen A Ross. The valuation of options for alternative stochastic processes. *Journal of financial economics*, 3(1-2):145–166, 1976.

[23] Andrei Cozma, Matthieu Mariapragassam, and Christoph Reisinger. Convergence of an euler scheme for a hybrid stochastic-local volatility model with stochastic rates in foreign exchange markets. *SIAM Journal on Financial Mathematics*, 9(1):127–170, 2018.

[24] Stéphane Crépey and Matthew Dixon. Gaussian process regression for derivative portfolio modeling and application to cva computations. *arXiv preprint arXiv:1901.11081*, 2019.

[25] Dmitry Davydov and Vadim Linetsky. Pricing and hedging path-dependent options under the cev process. *Management Science*, 47(7):949–965, 2001.

[26] Jan De Spiegeleer, Dilip B Madan, Sofie Reyners, and Wim Schoutens. Machine learning for quantitative finance: fast derivative pricing, hedging and fitting. *Quantitative Finance*, 18(10):1635–1643, 2018.

[27] Freddy Delbaen and Hiroshi Shirakawa. A note on option pricing for the constant elasticity of variance model. *Asia-Pacific Financial Markets*, 9:85–99, 2002.

[28] Steffen Dereich, Andreas Neuenkirch, and Lukasz Szpruch. An euler-type method for the strong approximation of the cox–ingersoll–ross process. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 468(2140):1105–1115, 2012.

[29] Matthew F Dixon, Igor Halperin, and Paul Bilokon. *Machine learning in finance*, volume 1170. Springer, 2020.

[30] David C Emanuel and James D MacBeth. Further results on the constant elasticity of variance call option pricing model. *Journal of Financial and Quantitative Analysis*, 17(4):533–554, 1982.

[31] Sophie Emerson, Ruairí Kennedy, Luke O'Shea, and John O'Brien. Trends and applications of machine learning in quantitative finance. In *8th international conference on economics and finance research (ICEFR 2019)*, 2019.

[32] W Gautsch and WF Cahill. Handbook of mathematical functions with formula, graphs, and mathematical tables ed m abramowitz and ia stegun, 1964.

[33] Stefan Gerhold, Friedrich Hubalek, and Richard B Paris. The running maximum of the cox-ingersoll-ross process with some properties of the kummer function. *arXiv preprint arXiv:2004.10697*, 2020.

[34] Thomas Gerstner, Bastian Harrach, and Daniel Roth. Convergence of milstein brownian bridge monte carlo methods and stable greeks calculation. *arXiv preprint arXiv:1906.11002*, 2019.

[35] M. B. Giles. Multilevel Monte Carlo methods. *Acta Numerica*, 24:259–328, 5 2015.

[36] Michael B. Giles. Multilevel Monte Carlo path simulation. *Oper. Res.*, 56(3):607–617, 2008.

[37] Michael B. Giles, Kristian Debrabant, and Andreas Rössler. Analysis of multilevel Monte Carlo path simulation using the Milstein discretisation. *Discrete Contin. Dyn. Syst. Ser. B*, 24(8):3881–3903, 2019.

[38] Michael B Giles, Desmond J Higham, and Xuerong Mao. Analysing multi-level monte carlo for options with non-globally lipschitz payoff. *Finance and Stochastics*, 13(3):403–413, 2009.

[39] Michael B. Giles, Desmond J. Higham, and Xuerong Mao. Analysing multi-level Monte Carlo for options with non-globally Lipschitz payoff. *Finance Stoch.*, 13(3):403–413, 2009.

[40] Michael B Giles and Lukasz Szpruch. Antithetic multilevel monte carlo estimation for multidimensional sdes. In *Monte Carlo and Quasi-Monte Carlo Methods 2012*, pages 367–384. Springer, 2013.

[41] Michael B Giles and Lukasz Szpruch. Antithetic multilevel monte carlo estimation for multi-dimensional sdes without lévy area simulation. 2014.

[42] Mike Giles. Improved multilevel monte carlo convergence using the milstein scheme. In *Monte Carlo and Quasi-Monte Carlo Methods 2006*, pages 343–358. Springer, 2006.

[43] Paul Glasserman. *Monte Carlo methods in financial engineering*, volume 53. Springer, 2004.

[44] Paul Glasserman and Paul Glasserman. Discretization methods. *Monte Carlo Methods in Financial Engineering*, pages 339–376, 2003.

[45] Paul Glasserman and Kyoung-Kuk Kim. Gamma expansion of the heston stochastic volatility model. *Finance and Stochastics*, 15(2):267–296, 2011.

[46] Paul Glasserman and Jeremy Staum. Conditioning on one-step survival for barrier option simulations. *Operations Research*, 49(6):923–937, 2001.

[47] Emmanuel Gobet. Advanced monte carlo methods for barrier and related exotic options. In Alain Bensoussan and Qiang Zhang, editors, *Special Volume: Mathematical Modeling and Numerical Methods in Finance*, volume 15 of *Handbook of Numerical Analysis*, pages 497–528. Elsevier, 2009.

[48] Ludovic Goudenège, Andrea Molent, and Antonino Zanette. Variance reduction applied to machine learning for pricing bermudan/american options in high dimension. *arXiv preprint arXiv:1903.11275*, 2019.

[49] Steven L Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of financial studies*, 6(2):327–343, 1993.

[50] Desmond J Higham and Xuerong Mao. Convergence of monte carlo simulations involving the mean-reverting square root process. *Journal of Computational Finance*, 8(3):35–61, 2005.

[51] Desmond J. Higham, Xuerong Mao, and Andrew M. Stuart. Strong convergence of Euler-type methods for nonlinear stochastic differential equations. *SIAM J. Numer. Anal.*, 40(3):1041–1063, 2002.

[52] Martin Hutzenthaler, Arnulf Jentzen, and Marco Noll. Strong convergence rates and temporal regularity for cox-ingersoll-ross processes and bessel processes with accessible boundaries. *arXiv preprint arXiv:1403.6385*, 2014.

[53] Monique Jeanblanc, Marc Yor, and Marc Chesney. *Mathematical methods for financial markets.* London: Springer, 2009.

[54] Christian Kahl and Peter Jäckel. Fast strong approximation monte carlo schemes for stochastic volatility models. *Quantitative Finance*, 6(6):513–536, 2006.

[55] Ioannis Karatzas, Ioannis Karatzas, Steven Shreve, and Steven E Shreve. *Brownian motion and stochastic calculus*, volume 113. Springer Science & Business Media, 1991.

[56] Ahmed Kebaier. Statistical Romberg extrapolation: a new variance reduction method and applications to option pricing. *Ann. Appl. Probab.*, 15(4):2681–2705, 2005.

[57] Ahmed Kebaier and Jérôme Lelong. Coupling importance sampling and multilevel Monte Carlo using sample average approximation. *Methodol. Comput. Appl. Probab.*, 20(2):611–641, 2018.

[58] Peter E Kloeden, Hannes Keller, and Björn Schmalfuß. Towards a theory of random numerical dynamics. *Stochastic Dynamics*, pages 259–282, 1999.

[59] Peter E Kloeden, Eckhard Platen, Peter E Kloeden, and Eckhard Platen. *Stochastic differential equations*. Springer, 1992.

[60] Vincent Lemaire and Gilles Pagès. Multilevel richardson–romberg extrapolation. 2017.

[61] Vadim Linetsky. Computing hitting time densities for cir and ou diffusions: Applications to mean-reverting models. *Journal of Computational Finance*, 7:1–22, 2004.

[62] Alexander Lipton and Artur Sepp. Toward an efficient hybrid method for pricing barrier options on assets with stochastic volatility. *arXiv preprint arXiv:2202.07849*, 2022.

[63] Roger Lord, Remmert Koekkoek, and Dick Van Dijk. A comparison of biased simulation schemes for stochastic volatility models. *Quantitative Finance*, 10(2):177–194, 2010.

[64] Michael Ludkovski. Kriging metamodels and experimental design for bermudan option pricing. *arXiv preprint arXiv:1509.02179*, 2015.

[65] Mike Ludkovski and Yuri Saporito. Krighedge: Gaussian process surrogates for delta hedging. *Applied Mathematical Finance*, 28(4):330–360, 2021.

[66] Alexander G de G Matthews, Mark Rowland, Jiri Hron, Richard E Turner, and Zoubin Ghahramani. Gaussian process behaviour in wide deep neural networks. *arXiv preprint arXiv:1804.11271*, 2018.

[67] Thomas Müller-Gronbach. Strong approximation of systems of stochastic differential equations. *Habilitationsanschrift, Technische Universität Darmstadt*, 2002.

[68] Andreas Neuenkirch and Lukasz Szpruch. First order strong approximations of scalar sdes defined in a domain. *Numerische Mathematik*, 128(1):103–136, 2014.

[69] Gilles Pagès. Numerical probability. *Universitext, Springer*, 2018.

[70] Allan Pinkus and Samy Zafrany. *Fourier series and integral transforms*. Cambridge University Press, Cambridge, 1997.

[71] Olivier Pironneau. Calibration of heston model with keras. 2019.

[72] Philip E Protter and Philip E Protter. *Stochastic differential equations*. Springer, 2005.

[73] Saburou Saitoh. Theory of reproducing kernels and its applications. *Longman Scientific & Technical*, 1988.

[74] Myron Scholes and Fischer Black. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

[75] Mark Schroder. Computing the constant elasticity of variance option pricing formula. *the Journal of Finance*, 44(1):211–219, 1989.

[76] Steven E Shreve. An introduction to singular stochastic control. In *Stochastic differential systems, stochastic control theory and applications*, pages 513–528. Springer, 1988.

[77] Denis Talay. Efficient numerical schemes for the approximation of expectations of functionals of the solution of a sde, and applications. In *Filtering and Control of Random Processes: Proceedings of the ENST-CNET Colloquium Paris, France, February 23–24, 1983*, pages 294–313. Springer, 1984.

[78] Nico M. Temme. *Asymptotic methods for integrals*. Hackensack, NJ: World Scientific, 2015.

[79] Grace Wahba. *Spline models for observational data*. SIAM, 1990.

[80] Christopher KI Williams and Carl Edward Rasmussen. *Gaussian processes for machine learning*, volume 2. MIT press Cambridge, MA, 2006.

[81] Chao Zheng. Weak convergence rate of a time-discrete scheme for the heston stochastic volatility model. *SIAM Journal on Numerical Analysis*, 55(3):1243–1263, 2017.