



Méthodes de Monte Carlo appliquées à la Finance :

Echantillonnage d'importance et Stratification
pour les processus de diffusion

NICOLAS BAUD VINCENT PORTE

Groupe de Travail

Dirigé par PAUL DEMEY et GAËL RIBOULET,
Groupe de Recherche Opérationnelle, Crédit Lyonnais

mai 2001

Table des matières

Introduction	1
1 Evaluation de produits dérivés par simulation	3
1.1 Principe des méthodes de Monte Carlo	3
1.2 Contrôle de la convergence	4
1.3 Accélération de la convergence	6
2 Changement de dérive déterministe	13
2.1 Discrétisation du processus de diffusion	14
2.2 Construction de la dérive suivant un principe de grandes déviations . . .	15
2.3 Accélération de la convergence par stratification	18
2.4 Applications : options sur un sous-jacent à volatilité stochastique	19
3 Changement de dérive stochastique	22
3.1 Processus optimal ϕ^*	23
3.2 Utilisation des grandes déviations pour l'accélération de la convergence .	24
3.3 Dérive de BLACK et SCHOLES	34
Conclusion	38
Annexes	40
A Principe de Grandes Déviations	40
B Changement de dérive déterministe et stratification : résultats numériques	42
C Changement de dérive stochastique : résultats numériques	43
C.1 Evaluation d'options d'achat européennes : modèle 1 (HESTON)	44
C.2 Evaluation d'options d'achat européennes : modèle 2	45
C.3 Evaluation d'options DIC : modèle 1 (HESTON)	46
C.4 Evaluation d'options DIC : modèle 2	47
D Variable de contrôle : le cas d'une option sur panier	48

Introduction

Le cadre général de notre problème est l'évaluation du prix d'un produit financier dont le payoff terminal dépend des valeurs futures d'un actif sous-jacent. Dans un premier temps, il convient de modéliser correctement l'évolution de la valeur du sous-jacent sur lequel est construit le produit financier.

Le modèle de BLACK-SCHOLES [2] est un des modèles les plus simples. La valeur du sous-jacent suit un mouvement brownien géométrique :

$$\frac{dS_t}{S_t} = bdt + \sigma dW_t$$

où b et σ sont deux constantes, (W_t) est un mouvement brownien standard sous la probabilité historique. Cette modélisation est très intéressante pour l'évaluation d'options. En effet, le marché muni d'un actif sans risque de rendement r est alors complet, ce qui permet de déterminer de façon unique le prix de tout produit financier. Un deuxième avantage est l'existence de formules fermées donnant la valeur de beaucoup de produits financiers, comme les options européennes de vente et d'achat. Cependant, ce modèle ne rend pas compte de plusieurs phénomènes. L'effet leptokurtique (queues épaisses des rendements) n'est pas pris en compte par ce modèle lognormal. De plus, le smile observé sur les volatilités implicites, calculées en inversant la formule de BLACK-SCHOLES pour les options, contredit ce modèle.

HULL et WHITE [15] se sont intéressés à un modèle plus général, dans lequel la valeur du sous-jacent est gouvernée par les équations suivantes :

$$\begin{aligned}\frac{dS_t}{S_t} &= b_t dt + \sigma(Y_t) dW_{1,t} \\ dY_t &= \bar{\eta}_t dt + \gamma_t dW_{2,t} \\ d\langle W_1, W_2 \rangle_t &= \rho_t dt\end{aligned}$$

Ce modèle ne présente pas les inconvénients du modèle de BLACK-SCHOLES. Néanmoins, le marché n'est plus complet et il n'est pas possible de déterminer de façon unique le prix des options. Cependant, l'existence d'une option européenne cotée permet de compléter le marché (ROMANO et TOUZI[18]). Ainsi, en supposant qu'il y a absence d'opportunité d'arbitrage sur les marchés, il existe une unique probabilité \mathbb{Q} telle que le prix I_g d'un produit financier de flux futurs actualisés $g(S_t, 0 \leq t \leq T)$ se mette sous la forme :

$$I_g = \mathbb{E}^{\mathbb{Q}} [g(S_t, 0 \leq t \leq T)].$$

Il n'existe plus en général de formule analytique simple qui permette d'obtenir I_g de façon directe, mais cette représentation probabiliste du prix, issue de la théorie développée par HARRISON et KREPS [11], suggère d'utiliser les méthodes de Monte Carlo pour en estimer la valeur.

La simulation par Monte Carlo est fondée sur la loi des grandes nombres. Le principe est de simuler, de façon indépendante, un grand nombre de variables aléatoires $(S_t)_{0 \leq t \leq T}$,

puis de prendre comme estimateur du paramètre d'intérêt I_g la moyenne empirique des $g(S_t, 0 \leq t \leq T)$ réalisés.

Si l'erreur commise est nulle de façon asymptotique, ce n'est pas le cas à distance finie. Il convient donc de mettre en place des outils de contrôle de la convergence. Ceux-ci reposent essentiellement sur l'utilisation du théorème Central Limite. Il apparaît que la variance de la loi simulée joue un rôle fondamental pour la construction des intervalles de confiance : plus la variance est élevée, plus la taille de l'intervalle est importante. Afin d'améliorer la précision de l'estimateur, on cherchera donc à simuler de nouvelles variables aléatoires d'espérance le paramètre d'intérêt I_g et de variance inférieure à celle de $g(S_t, 0 \leq t \leq T)$.

Dans la première partie de ce mémoire, nous présentons, dans un cadre général, les quatre méthodes de réduction de la variance les plus classiques. Nous nous intéresserons plus particulièrement à l'échantillonnage d'importance et à la stratification, méthodes que nous exploiterons, dans les deux parties suivantes, dans le cas particulier des processus de diffusion.

Appliquée au calcul du paramètre $I_g = \mathbb{E}^{\mathbb{Q}}[g(S_t, 0 \leq t \leq T)]$, l'échantillonnage d'importance est une méthode qui consiste à modifier la loi de la variable $(S_t)_{0 \leq t \leq T}$ simulée de façon à prendre en compte les spécificités de la fonction g à intégrer, les réalisations $g(S_t, 0 \leq t \leq T)$ étant alors pondérées afin de ne pas biaiser l'estimation de I_g . g représentant le payoff actualisé d'un produit dérivé, il s'agit en général d'une fonction positive ou nulle, strictement positive sur le domaine d'exercice. La simulation de trajectoires ne permettant pas l'exercice de l'option n'apportent que peu d'information pour estimer I_g . En revanche, ces trajectoires contribuent de façon significative à la variance de l'estimateur. Il peut donc être intéressant de modifier la loi du processus $(S_t)_{0 \leq t \leq T}$ afin d'augmenter le nombre de trajectoires simulées permettant l'exercice de l'option.

L'approche que nous développons dans les deuxième et troisième parties de ce mémoire consiste à modifier la dérive du processus $(S_t)_{0 \leq t \leq T}$ afin d'accroître la probabilité d'exercice. De façon plus fine, il s'agit de mettre en place des changements de dérive prenant en compte la distribution du payoff sur le domaine d'exercice : on cherche à construire un estimateur intégrant le maximum d'information pertinente sur la fonction à intégrer.

Les différentes méthodes présentées diffèrent par le mode de construction du changement de dérive. Dans la deuxième partie, on se limite aux seuls changements de dérive déterministes. Cette technique présente l'avantage de pouvoir être couplée à une deuxième méthode d'accélération : la stratification. Dans la troisième partie, nous nous intéressons au cas des changements de dérive stochastiques. Dans tous les cas, les changements de dérive proposés s'avèrent optimaux de façon asymptotique.

1. Evaluation de produits dérivés par simulation

1 Evaluation de produits dérivés par simulation

La représentation sous forme probabiliste des produits dérivés, issue de la théorie développée par HARRISON et KREPS [11], et HARRISON et PLISKA [12][13] conduit naturellement à considérer les méthodes de Monte Carlo pour l'évaluation de leurs prix : dans la mesure où un marché financier peut être doté d'une probabilité risque-neutre \mathbb{Q} , le prix d'un actif apparaît comme étant égal à l'espérance, sous cette probabilité, des flux futurs actualisés qu'il génère. Formellement, le problème de l'évaluation d'une option s'écrit

$$I_g = \mathbb{E}^{\mathbb{Q}} [g(S_t, 0 \leq t \leq T)]. \quad (1.1)$$

S_t représentant le prix du sous-jacent à la date t . Le prix d'un actif peut donc être calculé en simulant N trajectoires indépendantes $(S_t^1)_{0 \leq t \leq T}, \dots, (S_t^N)_{0 \leq t \leq T}$. Pour chacune de ces trajectoires, on calcule alors le montant $g(S_t^i, 0 \leq t \leq T)$ des flux futurs actualisés correspondants. La loi forte des grands nombres assure la convergence de la moyenne arithmétique des valeurs simulées vers le prix.

Dans un certain nombre de situations particulières, l'équation (1.1) se résout de façon analytique. C'est par exemple le cas des options d'achat européennes dans le modèle de BLACK et SHOLES. Cependant, pour des produits plus complexes, ou pour une modélisation plus fine du sous-jacent intégrant un aléa supplémentaire dans l'évolution de la fonction de volatilité σ , il est rare que l'on puisse obtenir des formules fermées d'évaluation. Il est alors nécessaire d'appliquer des méthodes numériques pour résoudre le problème d'évaluation de l'option. Les simulations par méthodes de Monte Carlo sont alors une des techniques couramment utilisées.

1.1 Principe des méthodes de Monte Carlo

Les techniques de simulation par méthode de Monte Carlo utilisent la loi forte des grands nombres pour estimer la valeur recherchée I_g . Si les courbes $S^i = \{S_t^i, 0 \leq t \leq T\}$, $1 \leq i \leq N$, représentent N trajectoires indépendantes distribuées suivant une même loi, alors

$$c_N = \frac{1}{N} \sum_{i=1}^N g(S^i) \xrightarrow{p.s.} I_g.$$

Le principe de base des méthodes de Monte Carlo consiste à générer N trajectoires $S^i = \{S_t^i, 0 \leq t \leq T\}$ et à considérer c_N comme estimateur du prix, celui-ci convergeant vers le paramètre d'intérêt I_g .

Le théorème central limite permet d'avoir une idée de l'approximation réalisée. Si le moment d'ordre 2 de $g(S_t, 0 \leq t \leq T)$ est fini, alors :

$$\sqrt{N} \left(\frac{1}{N} \sum_{i=1}^N g(S^i) - I_g \right) \xrightarrow{loi} \mathcal{N}(0, var[g(S)]).$$

1. Evaluation de produits dérivés par simulation

La détermination du nombre N de simulations nécessaires pour obtenir un estimateur satisfaisant du prix constitue l'un des problèmes auquel se trouve confronté le praticien. D'après le théorème central limite que l'on vient de rappeler, plus la variance V de $g(S^i)$ est élevée, plus le nombre de simulations nécessaires à l'obtention d'un estimateur précis est important. Si la réalisation d'une simulation est très coûteuse en temps de calcul, il peut être intéressant de développer une approche alternative jouant non pas sur N mais sur V afin de réduire le ratio $\frac{V}{N}$.

Nous présentons dans les paragraphes suivants les méthodes classiques de réduction de la variance utilisées dans le cadre des schémas de Monte Carlo. Dans un premier temps, nous adoptons un point de vue relativement général, indépendamment des considérations spécifiques à la modélisation financière. Le formalisme propre aux processus utilisés dans les applications financières est développé dans un second temps.

Parmi les différentes méthodes présentées, nous étudierons plus particulièrement la méthode d'échantillonnage d'importance (ou d'échantillonnage pondéré), sur laquelle se fondent les différents estimateurs de prix que nous présentons dans les deux parties suivantes.

Pour commencer, nous nous intéressons au problème du contrôle de la convergence. En particulier, nous rappelons succinctement la façon dont on construit les intervalles de confiance apparaissant dans les diverses applications de ce mémoire.

1.2 Contrôle de la convergence

Le cadre d'analyse est posé de la façon suivante. On considère un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{Q})$ et on se donne un variable aléatoire X à valeurs dans un espace \mathcal{S} , de loi $\mu(dx)$. On souhaite estimer numériquement, à partir d'une méthode de simulation de Monte Carlo, le paramètre d'intérêt

$$I_g = \int_{\mathcal{S}} g(x) \mu(dx) = \mathbb{E}^{\mathbb{Q}}[g(X)] \quad (1.2)$$

où g représente une fonction μ -intégrable. La méthode de Monte Carlo standard consiste à réaliser N tirages aléatoires indépendants (x_1, \dots, x_N) suivant la loi μ et à considérer

$$c_g = \frac{1}{N} \sum_{k=1}^N g(x_k) \quad (1.3)$$

comme estimateur du paramètre I_g . Il s'agit ici d'une application directe de la loi forte des grands nombres qui, compte tenu des hypothèses que l'on vient de formuler, assure que

$$\lim_{N \rightarrow \infty} c_g = I_g \text{ p.s.}$$

1. Evaluation de produits dérivés par simulation

Comme dans le cas des méthodes d'intégration déterministes, l'erreur d'approximation liée à c_g décroît lorsque N augmente. Néanmoins, l'erreur ne s'exprime pas ici en terme d'approximation mathématique mais suivant des considérations statistiques liées à l'aléa de l'échantillonnage. En particulier, comme nous l'avons déjà noté, la convergence de c_g vers le paramètre d'intérêt I_g est une convergence presque sûre, ce qui signifie que

$$\mathbb{Q} \left(\lim_{N \rightarrow \infty} c_g = I_g \right) = 1.$$

Si cette propriété est bien évidemment souhaitable, elle ne permet pas d'évaluer l'erreur commise lorsque N est fini.

Nous présentons ici la méthode d'évaluation de l'erreur communément utilisée en pratique pour résoudre ce problème. Elle repose sur la construction d'un intervalle de confiance. De plus, nous explicitons les approximations réalisées afin de mettre en évidence les limites d'un tel intervalle.

L'intervalle $IC(\lambda)$ que l'on souhaite expliciter est centré sur l'estimateur c_g , et tel que le paramètre d'intérêt I_g appartienne à $IC(\lambda)$ avec probabilité $1 - \lambda$:

$$\mathbb{Q}(I_g \in IC(\lambda)) = 1 - \lambda$$

Lorsque le moment d'ordre 2 de $g(X)$ est fini, le théorème central limite assure que :

$$\sqrt{N} \left(\frac{1}{N} \sum_{i=1}^N g(x_i) - I_g \right) \xrightarrow{loi} \mathcal{N}(0, var[g(X)]).$$

Ainsi, pour N suffisamment grand, on peut supposer que l'estimateur c_g est distribué suivant une loi $\mathcal{N}\left(I_g, \frac{var[g(X)]}{N}\right)$. Cette première approximation permet d'écrire $IC(\lambda)$ sous la forme suivante :

$$IC(\lambda) = \left[c_g - \sqrt{\frac{var[g(X)]}{N}} \Phi^{-1} \left(1 - \frac{\lambda}{2} \right), c_g + \sqrt{\frac{var[g(X)]}{N}} \Phi^{-1} \left(1 - \frac{\lambda}{2} \right) \right] \quad (1.4)$$

où Φ représente la fonction de répartition de la loi normale centrée réduite. Il est important de remarquer que l'approximation réalisée induit ici une première erreur sur la définition de l'intervalle de confiance, erreur qui ne peut être contrôlée : si les queues de distribution de la véritable loi de c_g sont plus épaisses que pour la loi normale, alors l'utilisation de l'intervalle $IC(\lambda)$ défini par la relation (1.4) sur-estime l'efficacité de l'estimateur c_g et, inversement, l'efficacité de c_g est sous-estimée si les queues de distribution sont plus fines.

Par ailleurs, le paramètre $var[g(X)]$ est *a priori* inconnu. Néanmoins, on peut l'estimer à partir de l'échantillon des variables x_i simulées pour construire c_g . Si l'on note v_g l'estimateur de la variance de $g(X)$ construit à partir des réalisations x_i , on a :

$$v_g = \frac{1}{N} \sum_{i=1}^N g(x_i)^2 - \left(\frac{1}{N} \sum_{i=1}^N g(x_i) \right)^2.$$

1. Evaluation de produits dérivés par simulation

L'intervalle de confiance s'écrit alors

$$IC(\lambda) = \left[c_g - \sqrt{\frac{v_g}{N}} \Phi^{-1} \left(1 - \frac{\lambda}{2} \right), c_g + \sqrt{\frac{v_g}{N}} \Phi^{-1} \left(1 - \frac{\lambda}{2} \right) \right] \quad (1.5)$$

Si v_g correspond bien à un estimateur convergent, sans biais, de $\text{var}[g(X)]$, le fait de remplacer $\text{var}[g(X)]$ par v_g dans la formule (1.4) introduit une nouvelle source d'erreur liée à l'aléa de l'échantillonnage.

Notons en outre que l'erreur induite par l'utilisation de la statistique v_g est d'autant plus importante que c_g et v_g ne sont pas des statistiques indépendantes. La corrélation entre ces deux statistiques conduit à sur-estimer, ou sous-estimer, la précision de l'estimateur c_g . Par exemple, dans le cas particulier où $g(X)$ est une variable de Bernoulli de paramètre λ , FISHMAN [4] montre que

$$\lim_{N \rightarrow \infty} \text{corr}(c_g, v_g) = \begin{cases} 1 & \text{si } 0 \leq \lambda < \frac{1}{2} \\ 0 & \lambda = \frac{1}{2} \\ -1 & \frac{1}{2} < \lambda \leq 1 \end{cases} .$$

En général, la corrélation entre c_g et v_g ne tend donc pas vers 0 lorsque le nombre de simulations utilisées tend vers l'infini.

Les intervalles de confiance que nous présentons dans les divers applications de ce mémoire reposent sur la formule (1.5). Les quelques remarques que nous venons de formuler montrent qu'il s'agit d'un outil de contrôle imparfait. Il permet néanmoins d'avoir une première idée de la vitesse de convergence des estimateurs que nous allons construire.

1.3 Accélération de la convergence

Dans les paragraphes suivants, nous présentons quatre méthodes usuelles qui, pour un nombre de simulations N fixé, permettent de construire des estimateurs de I_g de variance inférieure à $\text{var}(c_g)$: utilisation d'une variable de contrôle, utilisation de variables antithétiques, stratification constituent trois méthodes générales de réduction de la variance. La quatrième méthode nous intéresse plus particulièrement dans le cadre de ce mémoire : il s'agit de la méthode d'échantillonnage pondéré (ou d'importance).

Variation de contrôle

Si l'on connaît une fonction h pour laquelle on sait calculer $I_h = \mathbb{E}^{\mathbb{Q}}[h(X)]$, on peut construire un nouvel estimateur sans biais de I_g utilisant cette information supplémentaire. On considère pour cela l'estimateur pondéré

$$c_{ctl}(\beta) = c_g - \beta(c_h - I_h) \quad (1.6)$$

1. Evaluation de produits dérivés par simulation

où c_h est l'estimateur de Monte Carlo standard de I_h . La variance de c_Z est

$$\text{var}(c_{ctl}) = \text{var}(c_g) + \beta^2 \text{var}(c_h) - 2\beta \text{cov}(c_g, c_h).$$

Le choix optimal pour β correspond à

$$\beta^* = \frac{\text{cov}(c_g, c_h)}{\text{var}(c_h)},$$

la variance de $c_{ctl}(\beta^*)$ étant alors égale à $(1 - \rho_{g,h}^2) \text{var}(c_g)$. Notons que si la variance de $c_{ctl}(\beta^*)$ est nécessairement inférieure à la variance de c_g , une forte réduction de la variance implique cependant une corrélation très élevée entre la variable $g(X)$ et la variable de contrôle $h(X)$, ce qui limite la portée de cette première méthode.

D'un point de vue pratique, β^* étant inconnu (on ne connaît pas *a priori* $\text{cov}(c_g, c_h)$), on construit dans un premier temps un estimateur de ce paramètre à partir d'un nombre relativement réduit de simulations. La valeur obtenue est alors utilisée pour construire un estimateur c_{ctl} de I_g . Notons que l'estimation préliminaire de β^* ne biaise pas l'estimation de I_g par c_{ctl} , que l'on réalise dans un second temps seulement, de façon indépendante.

Nous présentons en annexe de ce mémoire, page 48, un exemple d'application pour lequel l'utilisation d'une variable de contrôle s'avère particulièrement efficace. Le problème que l'on se pose est l'évaluation d'une option d'achat européenne définie à partir de la moyenne arithmétique d'un panier d'actions. Dans un cadre d'analyse BLACK-SCHOLES, le prix d'un tel produit ne s'exprime pas de façon analytique. L'estimation du prix de ce type d'option par simulation est très rapide si on utilise comme variable de contrôle le même produit, mais défini cette fois-ci à partir de la moyenne géométrique des prix des actions. Dans l'application numérique présentée en annexe, on obtient une réduction de variance d'environ 100 dans le cas d'une option à la monnaie.

Variabes antithétiques

Jusqu'à présent, nous avons présenté les méthodes de Monte Carlo comme des méthodes nécessitant la production d'échantillons de variables aléatoires indépendantes, identiquement distribuées. Néanmoins, il peut s'avérer préférable de construire des échantillons de variables corrélées dans la mesure où cela peut conduire à une réduction de la variance des estimateurs correspondants.

Plus précisément, considérons un premier échantillon de taille $2N$, (x_1, \dots, x_{2N}) , de variables aléatoires indépendantes distribuées suivant la loi μ . Cet échantillon permet d'estimer le paramètre d'intérêt $I_g = \int_{\mathcal{S}} g(x) \mu(dx)$ à partir de la méthode de Monte Carlo standard, en utilisant la statistique c_g . Si l'on dispose par ailleurs d'un second échantillon de taille N , (y_1, \dots, y_N) , de variables aléatoires iid suivant la loi μ , on peut

1. Evaluation de produits dérivés par simulation

construire un nouvel estimateur c_{ant} de I_g en posant :

$$c_{ant} = \frac{1}{2N} \sum_{i=1}^N g(x_i) + g(y_i). \quad (1.7)$$

Si les variables $g(X_i)$ et $g(Y_i)$ sont corrélées négativement, cet estimateur est plus efficace que l'estimateur de Monte Carlo standard fondé sur un échantillon iid de taille $2N$.

RUBINSTEIN [19] propose la méthode suivante pour la simulation des échantillons $(x_i)_{1 \leq i \leq N}$ et $(y_i)_{1 \leq i \leq N}$ dans le cas où $g \circ \mu^-$ est monotone (μ^- représente l'inverse de la fonction de répartition de la loi μ) : x_i est généré à partir de la réalisation d'une loi uniforme u_i en posant

$$x_i = \mu^-(u_i),$$

y_i étant généré de la même façon à partir de $1 - u_i$. Cette méthode est relativement restrictive dans la mesure où d'une part, il faut que la condition de monotonie portant sur $g \circ \mu^-$ soit vérifiée, et d'autre part, il faut disposer de μ^- de façon analytique, ce qui est rarement le cas.

Une méthode plus aisément implémentable a été proposée par GEWEKE [8] dans le cas où la loi μ est symétrique autour d'une valeur m . Cette méthode consiste à simuler un échantillon de N réalisations (x_1, \dots, x_N) de façon indépendante suivant la loi μ et à poser $y_i = 2m - x_i, i \in \{1, \dots, N\}$. Une telle méthode est facilement applicable en finance dans la mesure où les variables aléatoires simulées sont la plupart du temps construites à partir de lois normales.

Stratification

Supposons maintenant que l'on puisse découper l'espace \mathcal{S} des valeurs de X en k ensembles mesurables Str_i , appelés strates, pour lesquels on sait calculer

$$\mu(Str_i) = \mathbb{Q}(X \in Str_i) = \mathbb{Q}(Str(X) = Str_i)$$

où $Str(X)$ représente la variable aléatoire indicatrice de la strate dans laquelle se situe X . Si de plus, on est capable de générer les variables aléatoires conditionnelles $X | (Str(X) = Str_i)$, de lois

$$\mu_i(dx) = \frac{\mathbf{1}_{Str_i}(x)}{\mu(Str_i)} \mu(dx), i \in \{1, \dots, k\},$$

il peut être intéressant de récrire le paramètre d'intérêt $I_g = \mathbb{E}^{\mathbb{Q}}[g(X)]$ sous la forme équivalente suivante :

$$I_g = \sum_{i=1}^k \mu(Str_i) \mathbb{E}^{\mathbb{Q}}[g(X) | X \in Str_i]$$

1. Evaluation de produits dérivés par simulation

où

$$\mathbb{E}^{\mathbb{Q}}[g(X) | Str_i] = \frac{1}{\mu(Str_i)} \int_{Str_i} g(x) \mu(dx).$$

La méthode de Monte Carlo par stratification consiste alors à simuler, au sein de chacune des strates Str_i considérée, un échantillon de N_i réalisations indépendantes $(x_1^i, \dots, x_{N_i}^i)$ issues de variables aléatoires distribuées suivant μ_i . On construit alors un estimateur sans biais de I_g en posant

$$c_{str} = \sum_{i=1}^k \frac{\mu(Str_i)}{N_i} \sum_{j=1}^{N_i} g(x_j^i) \quad (1.8)$$

Posons $N = \sum_{i=1}^k N_i$ et supposons que $N_i = \mu(Str_i) N$. Dans ce cas, la variance de l'estimateur c_{str} s'écrit :

$$var(c_{str}) = \frac{1}{N} \mathbb{E}^{\mathbb{Q}}[var(g(X) | Str(X))].$$

La formule de décomposition de la variance permet d'écrire

$$var(g(X)) = \mathbb{E}^{\mathbb{Q}}[var(g(X) | Str(X))] + var\left(\mathbb{E}^{\mathbb{Q}}[g(X) | Str(X)]\right),$$

si bien que

$$var(c_{str}) = var(c_g) - \frac{1}{N} var\left(\mathbb{E}^{\mathbb{Q}}[g(X) | Str(X)]\right).$$

La stratification permet donc de réduire la variance, par rapport à une méthode de Monte Carlo standard, du terme $\frac{1}{N} var\left(\mathbb{E}^{\mathbb{Q}}[g(X) | Str(X)]\right)$. Notons que cette méthode sera d'autant plus efficace que la variance de la variable $g(X)$ au sein de chacune des strates sera faible, et que la variance inter-strate sera élevée.

Le cas d'un vecteur gaussien. On s'intéresse ici au cas particulier d'un vecteur gaussien, les résultats obtenus dans ce court paragraphe étant utilisés par la suite.

Considérons une variable aléatoire X à valeur dans $\mathcal{S} = \mathbb{R}^d$, de loi $\mathcal{N}(0, I_d)$ et un vecteur unitaire $s \in \mathbb{R}^d$. Le vecteur aléatoire X peut se récrire sous la forme

$$X = (s'X) s + Y$$

où $Y = X - (s'X) s$ est un vecteur gaussien centrée de matrice de variance-covariance $I_d - P_s$, P_s étant la matrice de la projection orthogonale sur le vecteur s . De plus, $s'X$ est une variable aléatoire normale centrée réduite.

On construit alors k strates $Str_i, i \in \{1, \dots, k\}$, suivant la valeur prise par la variable $s'X$. Plus précisément, on définit Str_i par

$$\forall i \in \{1, \dots, k\}, X \in Str_i \iff \frac{i-1}{k} \leq \Phi(s'X) < \frac{i}{k}$$

1. Evaluation de produits dérivés par simulation

où Φ représente la fonction de répartition de la loi normale centrée réduite. On tire alors le vecteur aléatoire $X \mid X \in Str_i$ à partir de la décomposition :

$$X \mid X \in Str_i \stackrel{loi}{=} \Phi^{-1}(u_i) s + Y$$

où u_i suit une loi uniforme sur $[\frac{i-1}{k}; \frac{i}{k}]$ indépendante de Y .

Le problème qui se pose alors est de déterminer une direction s permettant de réduire significativement la variance. Nous traiterons cette question dans la deuxième partie de ce rapport.

Echantillonnage pondéré

L'échantillonnage pondéré constitue une quatrième méthode permettant de construire des estimateurs du paramètre d'intérêt I_g (équation 1.2 page 4), dont la variance, pour un nombre de simulations donné, est inférieure à la variance de l'estimateur de Monte Carlo standard. L'idée de l'échantillonnage d'importance est d'utiliser non plus la loi μ pour la simulation des variables aléatoires, mais une loi instrumentale qui tient compte de la répartition, au sein de l'espace \mathcal{S} , du produit $g(x) \mu(dx)$. Les simulations ainsi obtenues sont alors pondérées pour supprimer le biais induit par le changement de loi lors de l'étape de simulation.

Supposons que l'espace \mathcal{S} puisse être identifié à \mathbb{R}^d , $d \geq 1$, et que la mesure de probabilité μ admette une densité f par rapport à la mesure de Lebesgue sur \mathbb{R}^d :

$$\mu(dx) = f(x) dx.$$

De plus, en limitant le champ de l'étude aux fonctions g positives, on peut réécrire I_g sous la forme

$$I_g = \int_{\mathbb{R}^d} g(x) \mathbf{1}_D(x) f(x) dx \quad (1.9)$$

où D représente le domaine au sein duquel la fonction g prend des valeurs strictement positives.

Considérons à présent une fonction de densité sur \mathbb{R}^d , notée h , telle que

$$\forall x \in \mathbb{R}^d, g(x) > 0 \implies h(x) > 0.$$

Le relation (1.9) peut se réécrire

$$I_g = \int_{\mathbb{R}^d} g(x) \frac{f(x)}{h(x)} \mathbf{1}_D(x) h(x) dx. \quad (1.10)$$

1. Evaluation de produits dérivés par simulation

Il apparaît alors clairement que l'on peut construire un nouvel estimateur c_{is} du paramètre I_g en simulant N variables aléatoires indépendantes y_1, \dots, y_N suivant la loi h et en posant

$$c_{is} = \frac{1}{N} \sum_{i=1}^N \frac{f(y_i)}{h(y_i)} g(y_i) \quad (1.11)$$

La variance de cet estimateur vaut

$$\text{var}(c_{is}) = \int_D \left(g(x) \frac{f(x)}{h(x)} - I_g \right)^2 h(x) dx.$$

Ainsi la densité h^* qui minimise la variance de l'estimateur pondéré s'écrit :

$$h^*(x) = \frac{g(x) f(x)}{I_g},$$

la variance de l'estimateur c_{is}^* étant alors nulle. Notons que ce résultat est sans grand intérêt pratique puisque la construction de h^* fait intervenir le paramètre d'intérêt I_g ! Il montre néanmoins que la densité h à utiliser pour la construction de c_{is} doit être choisie de telle sorte que $\frac{fg}{h}$ soit quasi constant et de variance finie. Le changement de loi conduit donc à un tirage fréquent des valeurs de x pour lesquelles le produit gf prend des valeurs élevées, les valeurs de x telles que $f(x)g(x)$ est faible étant plus rarement sélectionnées.

Le cas des processus de diffusion. Le choix du changement de probabilité pour la simulation des variables aléatoires constitue la difficulté principale des méthodes d'échantillonnage d'importance. La famille de densités possibles, au sein de laquelle on recherchera la loi instrumentale, dépend bien entendu du cadre d'analyse dans lequel on se place.

Dans le cas particulier des modèles financiers, les variables aléatoires utilisées correspondent la plupart du temps à des processus de diffusion. Ainsi, I_g s'écrit

$$I_g = I_g(t, x) = \mathbb{E}^{\mathbb{Q}} [g(X_s, t \leq s \leq T) \mid X_t = x],$$

$X = (X_t)_{0 \leq t \leq T}$ étant un processus à valeur dans \mathbb{R}^d , solution d'une équation différentielle stochastique de la forme

$$dX_t = b(t, X_t) dt + \sigma(t, X_t) dW_t, \quad (1.12)$$

où $(W_t)_{t \geq 0}$ est un mouvement brownien standard de dimension d défini sur l'espace $(\Omega, \mathcal{F}, \mathbb{Q})$. Le théorème de Girsanov permet ici d'introduire un nouveau formalisme afin d'orienter le choix de la loi instrumentale.

1. Evaluation de produits dérivés par simulation

Théorème 1.1 – Girsanov –

Soit $T > 0$, $(\Omega, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathcal{F}_T, \mathbb{Q})$ un espace filtré et (W_t) un \mathcal{F}_t – mouvement brownien d – dimensionnel issu de 0.

Soit $(\phi_t)_{0 \leq t \leq T}$ un processus à valeurs dans \mathbb{R}^d , mesurable, adapté. Si le processus $M^{(\phi)}$ défini par

$$M_t^{(\phi)} = \exp \left[- \int_0^t \phi_s dW_s - \frac{1}{2} \int_0^t |\phi_s|^2 ds \right],$$

est une martingale sous \mathbb{Q} alors on définit une nouvelle probabilité $\mathbb{Q}^{(\phi)}$ sur (Ω, \mathcal{F}_T) en posant¹ :

$$\frac{d\mathbb{Q}^{(\phi)}}{d\mathbb{Q}} = M_T^{(\phi)},$$

et le processus $B_t = W_t + \int_0^t \phi_s ds, t \in [0, T]$ est un \mathcal{F}_t – mouvement brownien pour la probabilité $\mathbb{Q}^{(\phi)}$.

Compte tenu de ce théorème, il apparaît que le paramètre d'intérêt $I_g(t, x)$ peut s'écrire :

$$I_g(t, x) = \mathbb{E}^{\mathbb{Q}^{(\phi)}} \left[g(X_s, t \leq s \leq T) L_{t,T}^{(\phi)} | X_t = x \right], \quad (1.13)$$

où

$$\begin{aligned} L_{t,T}^{(\phi)} &= \frac{M_t^{(\phi)}}{M_T^{(\phi)}} \\ &= \exp \left[\int_t^T \phi_s dB_s - \frac{1}{2} \int_t^T |\phi_s|^2 ds \right] \end{aligned}$$

Un estimateur $c_{is}^{(\phi)}(t, x)$ de $I_g(t, x)$ peut donc être construit en simulant N trajectoires indépendantes suivant la loi du processus $(X_s)_{t \leq s \leq T}$ sous la probabilité $\mathbb{Q}^{(\phi)}$, et en posant

$$c_{is}^{(\phi)}(t, x) = \frac{1}{N} \sum_{i=1}^N g(x_s(i), t \leq s \leq T) L_{t,T}^{(\phi)}(i). \quad (1.14)$$

La simulation de trajectoires suivant la loi du processus $(X_s)_{t \leq s \leq T}$ sous la probabilité $\mathbb{Q}^{(\phi)}$ peut être réalisée en remarquant que $(X_s)_{t \leq s \leq T}$, solution de l'équation différentielle stochastique (1.12) sous \mathbb{Q} , est la solution, sous $\mathbb{Q}^{(\phi)}$, de l'équation différentielle stochastique suivante :

$$dX_s = [b(s, X_s) - \sigma(s, X_s) \phi_s] dt + \sigma(s, X_s) dB_s. \quad (1.15)$$

¹Une condition suffisante portant sur le processus ϕ correspond au critère de NOVIKOV : si $\mathbb{E}^{\mathbb{Q}} \left[\exp \left(\int_0^T |\phi_s|^2 ds \right) \right] < \infty$ alors $\mathbb{E}^{\mathbb{Q}} \left[M_T^{(\phi)} \right] = 1$ et $M^{(\phi)}$ est une \mathbb{Q} – martingale.

2. Changement de dérive déterministe

Le changement de probabilité pour la construction d'un nouvel estimateur $c_{is}^{(\phi)}(t, x)$ de $I_g(t, x)$ conduit donc à la modification du processus de dérive intervenant dans la caractérisation du processus (X_t) . L'idée est ici de modifier l'allure des trajectoires par le choix d'un processus ϕ judicieux, afin de prendre en compte les spécificités de la fonction g à intégrer.

Les deux sections suivantes présentent deux approches différentes conduisant à la construction d'un processus ϕ permettant de réduire la variance de l'estimateur de I_g . Dans la première approche, le processus ϕ utilisé est une fonction en escalier déterministe. Dans la seconde approche, ϕ est un processus stochastique : ϕ_t dépend, de façon déterministe, de la valeur X_t du processus sous-jacent à la date t . Que ϕ soit stochastique ou déterministe, la construction de ce processus repose sur l'utilisation de résultats asymptotiques issus de la théorie des grandes déviations. Appliquées à l'évaluation de produits dérivés, ces méthodes s'avèrent particulièrement efficaces dans le cas d'options très en dehors de la monnaie.

2 Changement de dérive déterministe

Cette partie est consacrée à la mise en place d'une méthode d'échantillonnage pondéré pour le calcul du paramètre d'intérêt

$$I_g(t, x) = \mathbb{E}^{\mathbb{Q}} [g(X_s, t \leq s \leq T) \mid X_t = x]$$

où (X_s) est un processus défini sur un espace probabilisé $(\Omega, (\mathcal{F}_t), \mathcal{F}, \mathbb{Q})$, à valeurs dans \mathbb{R}^d , solution d'une équation différentielle stochastique de la forme

$$dX_t = b(X_t) dt + \sigma(X_t) dW_t. \quad (2.1)$$

où (W_t) est un (\mathcal{F}_t) -mouvement brownien sous la probabilité \mathbb{Q} . Notons que l'on suppose désormais que les fonctions b et σ sont indépendantes du temps.

Dans la section précédente, nous avons explicité une méthode permettant de définir, à partir d'un processus ϕ , l'estimateur $c_{is}^{(\phi)}(t, x)$ (relation 1.14 page précédente) de $I_g(t, x)$. On dispose *a priori* d'une très grande liberté de choix pour le processus ϕ , dans la mesure où la seule contrainte qui lui est pour l'instant imposée est de vérifier les hypothèses du théorème de Girsanov. En particulier, ϕ peut être déterministe ou stochastique. Dans cette section, nous limitons le domaine des processus ϕ admissibles aux seuls processus déterministes.

La détermination d'une fonction ϕ conduisant à un estimateur $c_{is}^{(\phi)}(t, x)$ de faible variance est menée après discrétisation de l'équation différentielle stochastique vérifiée par (X_t) . Une telle approche restreint le domaine des fonctions possibles aux fonctions en escalier. Le choix de la meilleure fonction possible repose ensuite sur un raisonnement heuristique, raisonnement qui conduit à un choix optimal dans un cadre asymptotique.

2. Changement de dérive déterministe

2.1 Discrétisation du processus de diffusion

Il existe de nombreuses méthodes pour simuler la solution d'une équation différentielle stochastique (cf PARDOUX et TALAY [17]) entre les dates t et T . La méthode que nous utiliserons correspond à la méthode la plus simple. Il s'agit de la méthode d'Euler aléatoire.

Méthode d'Euler aléatoire

Etant donnée l'équation différentielle stochastique (2.1), on fixe un pas de temps $\Delta t = \frac{T-t}{l}$, $l \in \mathbb{N}$, et on construit un processus en temps discret (Y_i) approximant aux différentes dates $i\Delta t$, $0 \leq i \leq l$, le processus X solution de (2.1) en posant :

$$\begin{aligned} Y_0 &= X_t = x \\ Y_{i+1} &= Y_i + b(Y_i) \Delta t + \sigma(Y_i) (W_{t+(i+1)\Delta t} - W_{t+i\Delta t}) \end{aligned} \quad (2.2)$$

Le processus $X_s^{(l)} = Y_{t+\lceil \frac{s-t}{\Delta t} \rceil}$, $t \leq s \leq T$ approxime alors $(X_t)_{t \leq s \leq T}$ dans le sens suivant :

Théorème 2.1 *Pour tout $T > 0$,*

$$\mathbb{E}^{\mathbb{Q}} \left[\sup_{t \leq s \leq T} |X_s^{(l)} - X_s|^2 \right] \leq C_{T-t} \Delta t$$

où C_{T-t} est une constante ne dépendant que de $T - t$

Pour un horizon donné T , il est donc possible de rendre l'écart entre le processus $(X_s)_{t \leq s \leq T}$ et le processus $(X_s^{(l)})_{t \leq s \leq T}$ aussi petit qu'on le souhaite. Il suffit pour cela de choisir un pas de discrétisation Δt suffisamment fin.

L'estimation de $I_g(t, x)$ est réalisée à partir de la simulation de trajectoires du processus X . Une première approximation est donc faite puisque le calcul réellement mené est celui de

$$I_g(t, x, l) = \mathbb{E}^{\mathbb{Q}} \left[g \left(X_s^{(l)}, t \leq s \leq T \right) \mid X_t^{(l)} = x \right] \quad (2.3)$$

L'étude de la convergence de $I_g(t, x, l)$ vers $I_g(t, x)$ lorsque Δt tend vers 0 est traité dans GOBET [10].

Changement de probabilité après discrétisation

Dans la mesure où le schéma de discrétisation (2.2) est utilisé pour construire le processus $X^{(l)}$, le paramètre d'intérêt que l'on cherche à calculer apparaît comme l'espérance

2. Changement de dérive déterministe

d'une variable aléatoire dépendant d'un nombre l de lois normales centrées réduites indépendantes, à valeurs dans \mathbb{R}^d .

GLASSERMAN et al. [9] proposent alors d'étudier la méthode de réduction de variance fondée sur un changement de probabilité en limitant le domaine de processus ϕ possibles aux fonctions en escalier et déterministes.

Si l'on note ϕ_i , $i \in \{0, \dots, l-1\}$, la valeur prise par la fonction ϕ sur l'intervalle de temps $[t + i\Delta t, t + (i+1)\Delta t] = [t_i, t_{i+1}]$, le processus discrétisé $X^{(l)}$ vérifie sous la nouvelle probabilité $\mathbb{Q}^{(\phi)}$ induite par ϕ :

$$\begin{aligned} X_{t_0}^{(l)} &= x \\ X_{t_{i+1}}^{(l)} &= X_{t_i}^{(l)} + \left[b(X_{t_i}^{(l)}) - \sigma(X_{t_i}^{(l)}) \phi_i \right] \Delta t + \sqrt{\Delta t} \sigma(X_{t_i}^{(l)}) Z_{i+1} \\ &= X_{t_i}^{(l)} + b(X_{t_i}^{(l)}) \Delta t + \sqrt{\Delta t} \sigma(X_{t_i}^{(l)}) \left[Z_{i+1} - \sqrt{\Delta t} \phi_i \right] \end{aligned}$$

où Z_1, \dots, Z_l représentent l loi normales centrées réduites indépendantes sous $\mathbb{Q}^{(\phi)}$, à valeur dans \mathbb{R}^d . Dès lors, si l'on note $Z = (Z'_1, \dots, Z'_l)'$, le paramètre d'intérêt s'écrit, sous la probabilité $\mathbb{Q}^{(\phi)}$:

$$I_g(t, x, l) = \mathbb{E}^{\mathbb{Q}^{(\phi)}} \left[g \left(X^{(l)} \left(Z - \sqrt{\Delta t} \phi \right) \right) L_l^{(\phi)} \left(Z - \sqrt{\Delta t} \phi \right) \mid X_t^{(\Delta t)} = x \right]$$

où

$$\begin{aligned} L_l^{(\phi)} \left(Z - \sqrt{\Delta t} \phi \right) &= \exp \left(\sqrt{\Delta t} \phi' \left(Z - \sqrt{\Delta t} \phi \right) + \frac{\Delta t}{2} |\phi|^2 \right) \\ &= \exp \left(\sqrt{\Delta t} \phi' Z - \frac{\Delta t}{2} |\phi|^2 \right) \end{aligned}$$

On est donc ramené au calcul de l'espérance d'une fonction dépendant d'une loi normale centrée réduite de dimension ld .

2.2 Construction de la dérive suivant un principe de grandes déviations

Pour simplifier les notations, $g(X^{(l)}(Z))$ est identifié dans cette section à $g(Z)$ et φ représente le vecteur $-\sqrt{\Delta t} \phi$. De plus, on suppose que g est une fonction positive. Notons $D = \{z \in \mathbb{R}^{ld}, g > 0\}$. On définit alors sur D la fonction F par

$$F(z) = \ln g(z), \forall z \in D.$$

2. Changement de dérive déterministe

Minimisation du moment d'ordre 2

Compte tenu de ce que l'on vient d'écrire, le problème que l'on se pose est le suivant : on souhaite déterminer un vecteur $\varphi \in \mathbb{R}^{ld}$, tel que la variable aléatoire

$$g(Z + \varphi) \exp\left(-\varphi'Z - \frac{1}{2}\varphi'\varphi\right) \mathbf{1}_D(Z + \varphi)$$

ait une variance la plus faible possible, Z étant un vecteur aléatoire gaussien centré réduit sur \mathbb{R}^{ld} , sous la probabilité $\mathbb{Q}^{(\varphi)}$. La minimisation de la variance de cette variable aléatoire revient à minimiser son moment d'ordre 2 dans la mesure où son espérance est indépendante du vecteur φ considéré.

Le problème de minimisation de la variance s'écrit donc :

$$\min_{\varphi} \mathbb{E}^{\mathbb{Q}^{(\varphi)}} \left[g(Z + \varphi)^2 \exp(-2\varphi'Z + \varphi'\varphi) \mathbf{1}_D(Z + \varphi) \right]$$

ce qui s'écrit, en repassant sous la probabilité initiale :

$$\min_{\varphi} \mathbb{E}^{\mathbb{Q}} \left[\exp\left(2F(Z) + \varphi'Z + \frac{1}{2}\varphi'\varphi\right) \mathbf{1}_D \right], \text{ où } Z \sim \mathcal{N}(0, I) \text{ sous } \mathbb{Q}. \quad (2.4)$$

Ce problème est complexe car il s'agit de résoudre un calcul d'espérance similaire à ceux intervenant dans les calculs de prix d'options.

Approximation du problème : la méthode de Laplace abstraite

Une résolution approchée peut néanmoins être menée en utilisant une approximation suggérée par la théorie des grandes déviations, issue du théorème de la méthode de Laplace abstraite (pour plus de précisions, nous renvoyons à l'annexe A page 40). En effet, d'après ce théorème, si F est continue et s'il existe $A \in \mathbb{R}$ tel que $F(z) \leq A + cz'z$ où $c < \frac{1}{4}$, alors :

$$t \ln \mathbb{E}^{\mathbb{Q}} \left[e^{\frac{1}{t}(2F(\sqrt{t}Z) - \sqrt{t}\varphi'Z + \frac{1}{2}\varphi'\varphi)} \mathbf{1}_D \right] \xrightarrow{t \rightarrow 0} \max_{z \in D} \left\{ 2F(z) - \varphi'z + \frac{1}{2}\varphi'\varphi - \frac{1}{2}z'z \right\}$$

Cela amène donc à chercher un vecteur φ^* tel que :

$$\varphi^* \in \arg \min_{\varphi} \left\{ \max_{z \in D} \left\{ 2F(z) - \varphi'z + \frac{1}{2}\varphi'\varphi - \frac{1}{2}z'z \right\} \right\} \quad (2.5)$$

Résoudre ce problème est encore relativement complexe (résolution d'un problème min-max). Nous présentons alors une approche qui conduit à un problème plus simple.

2. Changement de dérive déterministe

Simplification du programme d'optimisation

Une autre idée, afin de trouver le vecteur de dérive φ est de lisser la variable aléatoire dont on cherche l'espérance I_g . Revenons donc au problème de la recherche de la dérive φ minimisant la variance de l'estimateur de la quantité inconnue $I_g(t, x, l)$. Si nous notons f_0 (respectivement f_φ) la densité de la loi gaussienne centrée (respectivement d'espérance φ) de matrice de variance-covariance l'identité, le problème posé peut alors se réécrire comme la minimisation de la variance de la variable aléatoire :

$$g(Z) \frac{f_0(Z)}{f_\varphi(Z)} \mathbf{1}_D(Z)$$

où $Z \sim \mathcal{N}(\varphi, I)$. Sous forme intégrale, cette variance s'écrit :

$$\int_D \left(g(z) \frac{f_0(z)}{f_\varphi(z)} - I_g(t, x, l) \right)^2 f_\varphi(z) dz.$$

D'après ce que nous avons remarqué dans la présentation de l'échantillonnage pondéré (sous-section 1.3), une manière de chercher φ est de faire en sorte que la densité f_φ soit le plus proportionnel possible à la fonction gf_0 . Pour cela, plaçons les maxima de f_φ et de la fonction gf_0 au même endroit. Cela revient alors à chercher le vecteur φ comme solution du programme :

$$\varphi^* \in \arg \max_{z \in D} \left\{ F(z) - \frac{1}{2} z'z \right\} \quad (2.6)$$

Ce programme est bien plus facile à résoudre que le programme trouvé dans la première approche. La proposition suivante démontrée dans GLASSERMAN et al.[9] prouve alors, sous certaines conditions, qu'il est possible de se limiter à la résolution de (2.6).

Proposition 2.1 *Supposons qu'il existe une fonction h telle que $D = \{z : h(z) \leq 0\}$ et qu'il existe $z \in D$ tel que $h(z) < 0$. Supposons que $F(x) - \frac{1}{4}x'x$ soit concave sur D . Si φ^* est solution de (2.6), alors φ^* résout (2.5).*

Pour l'ensemble des applications que nous étudierons par la suite, les conditions de la proposition sur le domaine D et sur F seront vérifiées. Cependant, l'hypothèse la plus contraignante sur le plan pratique est que D doit être convexe, ce qui n'est par exemple pas le cas lorsque l'on étudie les options à barrière.

La résolution approchée du programme (2.4) que nous venons d'effectuer en considérant le programme simplifié (2.6) peut être interprétée de la façon suivante : la définition du changement de probabilité à partir de $\varphi \in \arg \max_{z \in D} \{F(z) - \frac{1}{2}z'z\}$ élimine la part de la variance due à la composante linéaire de F . En effet, φ vérifie la condition du premier ordre :

$$\varphi = \nabla F(\varphi).$$

2. Changement de dérive déterministe

En réalisant un développement de Taylor au premier ordre de F au point φ , on obtient :

$$e^{F(\varphi) + \nabla F(\varphi)'Z + O(Z'Z) - \frac{1}{2}\varphi'Z - \frac{1}{2}\varphi'\varphi} \mathbf{1}_D = e^{F(\varphi) - \frac{1}{2}\varphi'\varphi + O(Z'Z)} \mathbf{1}_D$$

L'idée naturelle prolongeant cette approche, pour réduire encore la variance, est alors de mettre en place une méthode éliminant la part de la variance due à la composante quadratique de F . La méthode de stratification présentée dans la section suivante permet de résoudre un tel problème.

2.3 Accélération de la convergence par stratification

Dans cette partie, nous allons chercher une direction de stratification du vecteur gaussien Z , suivant le formalisme présenté dans le paragraphe 1.3 page 8. Effectuons pour cela un développement de Taylor de F au second ordre. Il vient, en notant H le Hessien de F :

$$e^{F(Z+\varphi) - \varphi'Z - \frac{1}{2}\varphi'\varphi} \mathbf{1}_D = e^{F(\varphi) - \frac{1}{2}\varphi'\varphi + \frac{1}{2}Z'H(\varphi)Z + o(Z'Z)} \mathbf{1}_D$$

Le choix essentiel est donc celui d'une direction s qui permet de réduire significativement la variance. Cela revient donc à chercher un vecteur s qui minimise la variance suivante :

$$\sigma_{s,k}^2 \stackrel{\text{not.}}{=} \text{var} \left[\mathbb{E}^{\mathbb{Q}} \left[\frac{1}{k} \sum_{i=1}^k e^{\frac{1}{2}Z'H(\varphi)Z} \mathbf{1}_D \mid \text{Str}^{(s)}(Z) \right] \right]$$

où $\text{Str}^{(s)}(Z)$ représente la variable indicatrice de la strate dans laquelle se situe Z . Plus précisément,

$$\text{Str}^{(s)}(Z) = \text{Str}_i \iff \left\{ s'Z \in \text{Str}_i = \left[\Phi^{-1} \left(\frac{i-1}{k} \right), \Phi^{-1} \left(\frac{i}{k} \right) \right] \right\},$$

k correspondant au nombre de strates considérées. Ce problème est complexe car la variable aléatoire $s'Z \mid s'Z \in \text{Str}_i$ (variable gaussienne tronquée) a une loi non standard, et ainsi le vecteur aléatoire $Z \mid s'Z \in \text{Str}_i$ ne possède plus une loi standard. Pour simplifier le problème et se ramener au cas gaussien, on utilise alors le lemme suivant :

Lemme 2.1 *si k est une puissance de 2,*

$$\sigma_{s,k}^2 = \frac{1}{k} \mathbb{E}^{\mathbb{Q}} \left[\text{var} \left(e^{\frac{1}{2}Z'H(\varphi)Z} \mid s'Z \right) \right] + o \left(\frac{1}{k} \right)$$

La variable aléatoire $Z \mid s'Z$ est cette fois bien une loi gaussienne. Cela amène donc à chercher la direction optimale de stratification en résolvant le problème :

$$\min_{s: s's=1} \mathbb{E}^{\mathbb{Q}} \left[\text{Var} \left[e^{\frac{1}{2}Z'H(\varphi)Z} \mid s'Z \right] \right] \quad (2.7)$$

Ce problème est alors résolu grâce au théorème suivant :

2. Changement de dérive déterministe

Théorème 2.2 Soient e_1, \dots, e_{ld} les vecteurs propres (normalisés) de $H(\varphi)$ de valeurs propres respectives $\lambda_1, \dots, \lambda_{ld}$. Soit j^* l'indice tel que :

$$\left(\frac{\lambda_{j^*}}{1 - \lambda_{j^*}}\right)^2 = \max_{i=1, \dots, ld} \left(\frac{\lambda_i}{1 - \lambda_i}\right)^2$$

alors e_{j^*} résout le problème (2.7).

En pratique, il s'agit donc, une fois trouvé le changement de dérive φ optimal, de calculer la matrice hessienne de F en ce point. On calcule ensuite les valeurs propres de la matrice obtenue, et l'on choisit comme direction de stratification un vecteur propre unitaire e_{j^*} associée à la valeur propre λ_{j^*} vérifiant :

$$\left(\frac{\lambda_{j^*}}{1 - \lambda_{j^*}}\right)^2 = \max_{i=1, \dots, l} \left(\frac{\lambda_i}{1 - \lambda_i}\right)^2$$

Nous allons voir dans la suite plusieurs applications de cette méthode.

2.4 Applications : options sur un sous-jacent à volatilité stochastique

On se place ici dans le cadre d'un modèle à volatilité stochastique. La spécification que nous avons choisie est celle de HULL et WHITE [15]. Dans ce modèle, la valeur du sous-jacent est gouvernée, sous une mesure martingale \mathbb{Q} , par l'équation différentielle stochastique :

$$\begin{aligned} \frac{dS_t}{S_t} &= rdt + \sqrt{V_t}dW_{1,t} \\ dV_t &= \nu V_t dt + \xi V_t dW_{2,t} \end{aligned}$$

où $(W_{1,t}), (W_{2,t})$ sont deux mouvements browniens standards tels que $\langle dW_{1,t}, dW_{2,t} \rangle = \rho dt$, où r représente le rendement de l'actif sans risque, ν et ξ sont 2 paramètres du modèle. Comme cas particulier de ce modèle, on retrouve le modèle de BLACK-SCHOLES ($\nu = \xi = 0$).

La théorie d'accélération des techniques de Monte Carlo présentée dans ce paragraphe s'effectue sur le processus discrétisé :

$$\begin{aligned} S_{i+1} &= S_i \left(1 + r\Delta t + \sqrt{V_i\Delta t}Z_i\right) \\ V_{i+1} &= V_i e^{(v - \frac{1}{2}\xi^2)\Delta t + \xi\sqrt{\Delta t}(\rho Z_i + \sqrt{1 - \rho^2}Z_{i+1})} \end{aligned}$$

où $Z = (Z_1, \dots, Z_{2l})$ est un vecteur gaussien centré et de variance l'identité.

Application numérique 2.1 Pour les options que nous étudierons dans la suite de cette section, nous avons choisi les valeurs suivantes :

2. Changement de dérive déterministe

- La valeur S_0 du sous-jacent est choisie à 100, le prix d'exercice K à 110, la volatilité initiale $\sqrt{V_0}$ à 0,2, la corrélation ρ entre les browniens vaut $-0,5$, le rendement r vaut 0, la tendance ν de la variance est nulle, et le nombre l de pas de discrétisation est choisi à 32.
- Nous allons ensuite jouer sur la maturité T des options considérées afin de rendre les options plus ou moins hors la monnaie. En effet, plus la maturité sera faible plus la probabilité que l'option ait un payoff nul sera grande. Enfin cette étude s'effectuera suivant différentes valeurs de ξ , calibrant la perturbation de la volatilité du sous-jacent.

Premier exemple : option asiatique

L'option asiatique est caractérisée par un payoff terminal fonction d'une moyenne arithmétique de valeurs du sous-jacent à différentes dates. Ainsi, pour des dates à intervalles réguliers, l'option asiatique de prix d'exercice K que l'on étudie aura un payoff terminal de la forme $(\bar{S}(Z) - K)^+$. On veut donc trouver l'espérance de

$$e^{-rT} (\bar{S}(Z) - K)^+ = \left(\frac{1}{l} \sum_{i=1}^l S_i(Z) - K \right)^+$$

où Z est un vecteur gaussien de dimension $2l$. En appliquant ce qui précède, la dérive optimale est obtenue en maximisant (de façon numérique), sur le domaine $D = \{z \in \mathbb{R}^{2l} \mid (\bar{S}(z) - K) > 0\}$, la fonction :

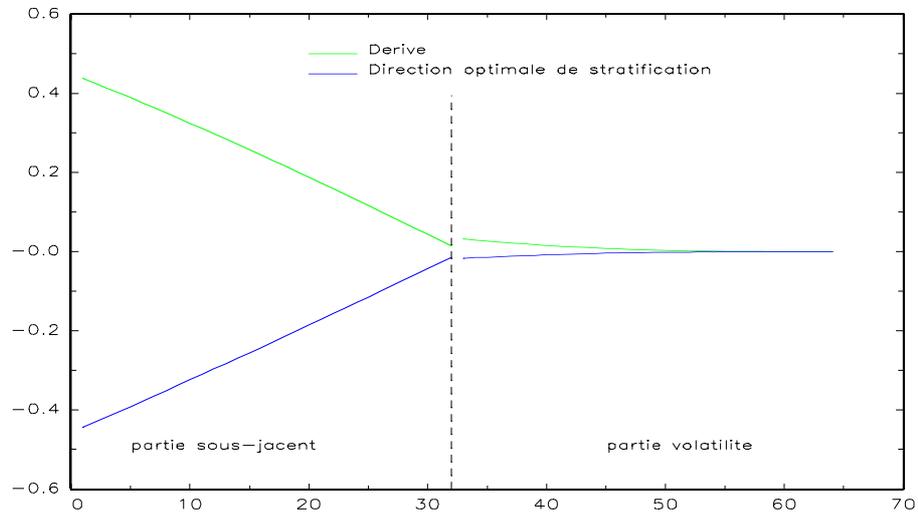
$$\ln(\bar{S}(z) - K) - \frac{1}{2} z'z.$$

Puis, une fois la dérive optimale φ estimée, on calcule la matrice hessienne de la fonction $\ln(\bar{S}(\cdot) - K)$ au point φ . On choisit alors la direction optimale de stratification à partir d'une diagonalisation numérique de la matrice obtenue, en utilisant le résultat du théorème 2.2. Le graphique 2.1 page suivante présente les vecteurs obtenus dans le cas d'une option de maturité 1 an.

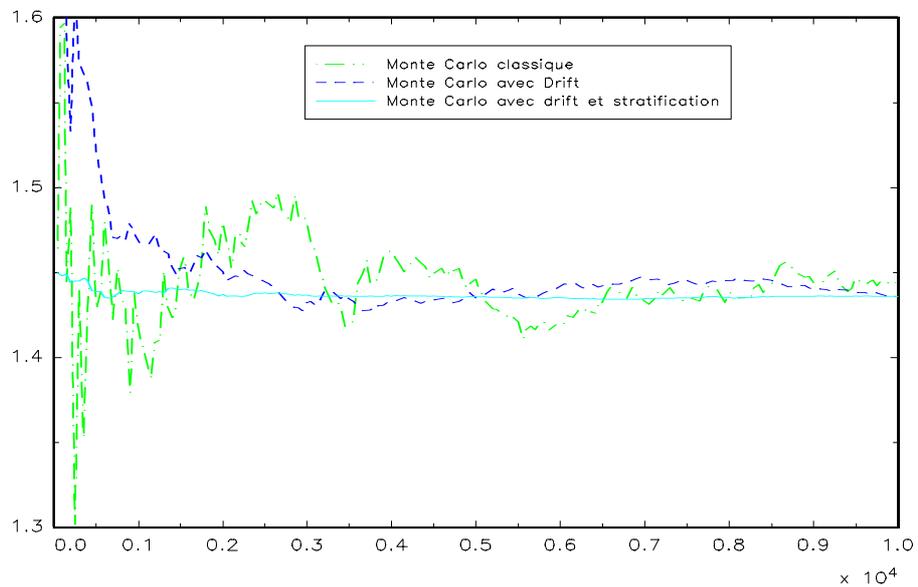
Nous présentons dans le tableau B.1 page 42 l'évolution des réductions de variance lorsque la maturité diminue, pour $\xi = 0.2$ et $\xi = 0$.

Les réductions de variance dépendent bien évidemment des paramètres du modèle, mais sont significatives dans de nombreux cas. Pour une maturité d'un an, les paramètres choisis sont raisonnables, et l'option asiatique aura souvent un payoff non nul sous la probabilité initiale. Pourtant les accélérations de convergence sont déjà non négligeables. Ainsi, on trouve pour $\xi = 0.2$ une réduction de variance d'un facteur 15 avec changement de dérive, de 864 avec changement de dérive et stratification, par rapport à une méthode de Monte Carlo classique. Dans le graphique 2.2 page suivante, nous présentons les courbes de convergence des différentes méthodes correspondant à ce jeu de paramètres.

2. Changement de dérive déterministe



Graphique 2.1. Dérive et direction optimale de stratification pour une option asiatique, modèle de HULL et WHITE. $K = 110$, $S_0 = 100$, $V_0 = 0.04$, $\xi = 0.2$, $\rho = -0.5$, $\nu = 0$, et $l = 32$



Graphique 2.2. Estimation du prix d'une option asiatique de maturité 1 an et de prix d'exercice 110, modèle de HULL et WHITE. $S_0 = 100$, $V_0 = 0.04$. $\xi = 0.2$, $\nu = 0$, $\rho = -0.5$ et $l = 32$.

3. Changement de dérive stochastique

De plus, lorsque l'exercice de l'option devient rare, les réductions de variance peuvent être très importantes. Ainsi, si la volatilité est choisie constante (i.e. $\nu = \xi = 0$), ce qui correspond au cas particulier BLACK-SCHOLES, et que l'on choisit une maturité plus petite (en l'occurrence $T = 0.25$), le facteur de réduction de variance est de 41 avec changement de dérive, et de 7847 avec changement de dérive et stratification.

Sur ces premiers résultats, on voit conformément à l'intuition que le facteur de réduction de variance est d'autant plus important que l'option est en dehors de la monnaie.

Deuxième exemple : option européenne d'achat

Dans le cadre du modèle à volatilité stochastique étudié, et de la discrétisation utilisée, on s'intéresse maintenant à un payoff terminal de la forme : $(S_T(Z) - K)^+$, et l'on applique le changement de dérive déterministe et la stratification suivant la méthode précédemment présentée. Le domaine $D = \{z \in \mathbb{R}^{2l} \mid S_T(z) > K\}$ est bien convexe et on obtient la dérive optimale φ en maximisant sur D de façon numérique l'équation :

$$\ln(S_T(z) - K) - \frac{1}{2}z'z.$$

La stratification optimale s est obtenue ensuite en calculant la matrice hessienne de la fonction $\ln(S_T(\cdot) - K)$ au point φ . Les réductions de variance sont alors significatives lorsque l'option est très en dehors de la monnaie. Le tableau B.2 page 42 met en lumière la croissance du facteur de réduction de la variance lorsque la maturité diminue.

Conclusion

La méthode présentée ci-dessus fondée sur un changement de dérive et sur une stratification permet une accélération de la convergence, surtout lorsque les options sont très en dehors de la monnaie. Cependant, une des hypothèses les plus contraignantes de cette méthode est que le domaine D où le payoff de l'option est non nul doit être convexe. Ce n'est par exemple pas le cas lorsque l'on s'intéresse à une option à barrière. De plus, en se limitant à un changement de dérive déterministe, on n'utilise pas l'information accumulée grâce à l'évolution de la valeur du sous-jacent. Nous présentons dans la suite de ce mémoire une deuxième méthode qui permet de traiter ce cas par un changement de dérive stochastique.

3 Changement de dérive stochastique

Dans la partie précédente, la recherche d'un nouveau processus de dérive permettant de construire un estimateur de $I_g(t, x)$ de variance inférieure à celle de l'estimateur de Monte Carlo standard est effectuée après discrétisation du processus de diffusion

3. Changement de dérive stochastique

(X_t) . La méthode présentée dans cette section n'intègre le problème de la discrétisation qu'après détermination de la nouvelle dérive. De plus, on ne restreint plus le domaine des processus ϕ possibles aux seules fonctions déterministes : ϕ peut être stochastique dans le sens où il peut s'agir d'une fonction du processus (X_t) .

Comme nous l'avons déjà noté dans la première partie de ce mémoire, lors de l'étude générale des méthodes d'échantillonnage pondéré, il est possible d'exprimer le changement de probabilité qui conduit à un estimateur de variance nulle en fonction de $I_g(t, x)$. Un tel résultat semble *a priori* inutile dans la mesure où le paramètre d'intérêt, qui est inconnu, intervient dans la construction de l'estimateur optimal. Il peut néanmoins nous aider à choisir un changement de probabilité pertinent, qui induira une réduction significative de la variance.

Nous présentons dans les paragraphes suivants une méthode d'échantillonnage d'importance dans laquelle le changement de dérive est réalisé à partir d'une approximation du changement de dérive optimal. Nous étudierons ensuite l'efficacité de la méthode en l'appliquant au calcul du prix des options d'achat européennes et des options à barrière de type DIC (Down and In Call).

3.1 Processus optimal ϕ^*

Jusqu'ici, nous avons supposé que g pouvait dépendre de l'ensemble des valeurs prises par le processus X entre les dates t et T . Dans ce paragraphe seulement, nous ne considérons que les fonctions g dépendant uniquement de la valeur terminale X_T . Dans ce cas, il est en effet possible d'exhiber un changement de dérive optimal, conduisant à une variance nulle.

Le paramètre d'intérêt $I_g(t, x)$ se réécrit, sous la probabilité $\mathbb{Q}^{(\phi)}$ induite par le processus ϕ :

$$I_g(t, x) = \mathbb{E}^{\mathbb{Q}^{(\phi)}} \left[g(X_T) L_{t,T}^{(\phi)} \mid X_t = x \right].$$

On peut alors montrer, en appliquant le calcul d'Itô, que la variable $g(X_T) L_{t,T}^{(\phi)}$ s'écrit :

$$g(X_T) L_{t,T}^{(\phi)} = I_g(t, x) + \int_t^T L_{t,s}^{(\phi)} \left[\sigma(X_s)' \nabla_x I_g(s, X_s) + \phi_s I_g(s, X_s) \right] dB_s$$

où $\nabla_x I_g$ représente le gradient de la fonction I_g par rapport aux coordonnées de X . Il apparaît donc que la variance de $g(X_T) L_{t,T}^{(\phi)}$ sera nulle dès que

$$\phi_s = \phi_s^* = -\frac{1}{I_g(s, X_s)} \sigma(X_s)' \nabla_x I_g(s, X_s). \quad (3.1)$$

3. Changement de dérive stochastique

La fonction I_g étant inconnue, il est impossible d'implémenter un changement de dérive utilisant ϕ^* . Néanmoins, l'utilisation d'un processus ϕ construit à partir de la formule (3.1), en remplaçant I_g par une fonction qui l'approxime, peut permettre une réduction significative de la variance. Notons de plus que si la formule (3.1) a été établie pour des fonctions g ne dépendant que de X_T , on peut supposer que son utilisation, dans le cadre plus général de fonctions dépendant de tout ou partie de la trajectoire, induira également une réduction de la variance.

La recherche de processus ϕ pertinents peut être menée suivant deux directions différentes.

1. La première approche consiste à utiliser une approximation asymptotique de la fonction I_g dans la formule (3.1) dans le cas où g correspond au payoff de l'option *Up and In Bond*² (UIB). On utilise pour cela un résultat issu de la théorie des grandes déviations. Le changement de dérive qui en résulte permet de simuler un plus grand nombre de trajectoires appartenant au domaine d'exercice de l'option UIB. Dès lors, il peut être appliqué à tout type d'option caractérisée par un domaine d'exercice similaire : c'est par exemple le cas de l'option d'achat européenne. Nous montrerons comment la méthode peut-être étendue au cas des options à barrière en considérant le cas particulier de l'option *DIC*.
2. Dans la deuxième approche, on remplace la fonction $I_g(t, x)$ dans la formule (3.1) par la fonction $I_g(t, x)$ que l'on obtient lorsque l'on spécifie un nouveau modèle sous-jacent pour le processus (X_s) "approximant" le modèle initial, et dans lequel $I_g(t, x)$ s'écrit sous la forme d'une formule fermée. Le modèle de BLACK et SCHOLLES constitue un candidat naturel pour appliquer ce type de méthode.

3.2 Utilisation des grandes déviations pour l'accélération de la convergence

La méthode que l'on cherche à mettre en place ici repose sur l'idée suivante : si l'exercice d'une option, définie par sa fonction de payoff g , correspond à un évènement rare, une proportion élevée de trajectoires simulées conduit à un payoff nul. Il est alors difficile d'évaluer avec précision la répartition du payoff au sein du domaine d'exercice. De façon plus problématique, il se peut même que l'on ne parvienne à simuler aucune trajectoire induisant un payoff strictement positif. On s'attend donc à ce que la variance de l'estimateur du prix $I_g(t, x)$ soit élevée. Afin de réduire cette variance, il peut être intéressant de réaliser un changement de dérive rendant le domaine d'exercice plus probable.

²L'option *Up and in Bond* permet à son détenteur de toucher 1 unité de numéraire à la maturité si le sous-jacent passe au-dessus d'une barrière au cours de la vie du contrat.

3. Changement de dérive stochastique

Le cas de l'option *Up and In Bond*

Dans cette optique, on commence par se restreindre au calcul de $I_g(t, x)$ dans le cas où ce paramètre correspond au prix $UIB(t, x)$ d'un *Up and In Bond*, c'est à dire d'une option versant 1 unité de numéraire si l'actif sous-jacent passe au-dessus d'une barrière K prédéfinie. Notons Γ le domaine de \mathbb{R}^d à atteindre pour que l'option soit exercée :

$$\Gamma = \left\{ x \in \mathbb{R}^d, x_1 \geq K \right\}.$$

Introduisons le temps d'atteinte τ de ce domaine, défini sur $]t, \infty]$:

$$\tau = \inf \{ s \geq t, X_s \in \Gamma \}.$$

Compte tenu de ces nouvelles notations, $UIB(t, x)$ se réécrit :

$$UIB(t, x) = \mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{\tau \leq T} | X_t = x].$$

Une première approche : le modèle de Black et Scholes. Dans le cas particulier où le processus de diffusion considéré est à valeur dans \mathbb{R} , solution de

$$dX_t = bX_t dt + \sigma X_t dW_t,$$

on dispose d'une formule analytique de $UIB(t, x)$:

$$UIB(t, x) = \mathbb{Q}(\tau < t | X_0 = x) = \Phi(d_1) + \left(\frac{x}{K}\right)^\gamma \Phi(d_2) \quad (3.2)$$

où Φ représente la fonction de répartition de la loi normale centrée réduite, $\gamma = 1 - \frac{2b}{\sigma^2}$, $d_1 = \frac{1}{\sigma\sqrt{t}} \left[\ln \frac{x}{K} - \frac{\gamma}{2} \sigma^2 t \right]$, et $d_2 = d_1 + \gamma \sigma \sqrt{t}$. Il est bien évident que le changement de dérive que l'on peut obtenir ici ne sera d'aucune utilité pour le calcul de $UIB(t, x)$. Néanmoins, ce changement pourra servir pour estimer le prix d'options qui, même dans un cadre BLACK-SCHOLES, ne possèdent pas de formules d'évaluation fermées. Compte tenu de la formule (3.2), il est relativement aisé de démontrer que

$$\lim_{t \rightarrow 0} -2t \ln \mathbb{Q}(\tau < t | X_0 = x) = \left[\frac{1}{\sigma} \ln \left(\frac{x}{K} \right) \right]^2$$

On peut exploiter ce résultat pour construire un processus de dérive à partir de la formule (3.1) : on approxime $UIB(t, x)$ par $\exp \left[-\frac{1}{2(T-t)\sigma^2} \ln \left(\frac{x}{K} \right)^2 \right]$, ce qui conduit à poser

$$\phi(t, x) = \frac{1}{\sigma(T-t)} \ln \left(\frac{x}{K} \right) \mathbf{1}_{x < K}.$$

Cette dernière formule montre que trois paramètres influencent l'amplitude de la correction apportée à la trajectoire :

3. Changement de dérive stochastique

- le temps restant pour atteindre la barrière K .
- le niveau de la volatilité : toutes choses égales par ailleurs, plus la volatilité du processus est élevée, plus il est probable que le processus atteigne la barrière K avant la date limite T .
- la distance de x à K .

Généralisation aux modèles à volatilité stochastique. Le résultat obtenu dans le cas particulier de modèle de BLACK et SCHOLES se généralise à l'ensemble des processus de diffusion solutions d'une équation différentielle stochastique de la forme

$$dX_t = b(X_t) dt + \sigma(X_t) dW_t.$$

On dispose en effet du résultat suivant, portant sur le temps d'atteinte τ défini précédemment :

$$\lim_{t \rightarrow 0} -2t \ln \mathbb{Q}(\tau < t \mid X_0 = x) = \mathfrak{D}^2(x) \quad (3.3)$$

où \mathfrak{D} est l'unique solution de viscosité de l'équation différentielle non linéaire

$$\begin{aligned} \nabla \mathfrak{D}(x)' \sigma(x) \sigma(x)' \nabla \mathfrak{D}(x) &= 1 & \text{si } x \in \mathbb{R}^d \setminus \Gamma \\ &= 0 & \text{sinon} \end{aligned} \quad (3.4)$$

Pour plus de détails sur ce résultat, le lecteur pourra se référer à FLEMING et SONER [5], FOURNIÉ, LASRY et LIONS [6], et FOURNIÉ, LASRY et TOUZI [7].

Comme dans le cadre BLACK-SCHOLES, ce résultat permet d'exhiber une fonction approximant la paramètre d'intérêt $UIB(t, x)$ puisque asymptotiquement,

$$UIB(t, x) \approx \exp -\frac{\mathfrak{D}^2(x)}{2(T-t)}.$$

Le changement de probabilité utilisé pour réduire la variance de l'estimateur de $UIB(t, x)$ est alors construit à partir du processus ϕ suivant :

$$\phi(t, x) = \frac{\mathfrak{D}(x)}{T-t} \sigma(x)' \nabla \mathfrak{D}(x). \quad (3.5)$$

L'implémentation de ce changement de dérive requiert au préalable la recherche de la solution de viscosité à l'équation différentielle (3.4). Nous présentons ici une approche suggérée par FOURNIÉ, LASRY et TOUZI [7], qui est fondée sur l'interprétation de \mathfrak{D} en terme de distance au domaine Γ .

3. Changement de dérive stochastique

Interprétation de la fonction \mathfrak{D} . Si l'on munit l'espace \mathbb{R}^d de la métrique locale $A(x) = [\sigma(x)\sigma(x)']^{-1}$, $x \in \mathbb{R}^d$, associée à la fonction de volatilité σ du processus de diffusion, alors $\mathfrak{D}(x)$ s'interprète comme la longueur du plus court chemin permettant de rejoindre le domaine Γ partant de x . En effet la solution \mathfrak{D} du programme (3.4) vérifie également

$$\mathfrak{D}(x) = \inf_{y \in \Gamma} \mathfrak{D}(x, y) \quad (3.6)$$

où

$$\mathfrak{D}(x, y) = \inf_{w \in H(x, y)} \int_0^1 \left\langle A(w(t)) \dot{w}(t), \dot{w}(t) \right\rangle^{\frac{1}{2}} dt$$

avec $H(x, y) = \{w \in C^1([0, 1], \mathbb{R}^d) : w(0) = x, w(1) = y\}$, \dot{w} représentant la dérivée w de par rapport à t .

La présentation de \mathfrak{D} sous cette nouvelle forme permet d'analyser de façon qualitative l'impact du changement de dérive (3.5) sur la distribution des trajectoires du processus (X_t) .

Notons w^* la solution du problème (3.6). Asymptotiquement, l'évènement $\{\tau < T\}$ correspond à l'ensemble des trajectoires du processus (X_t) oscillant de façon très rapprochée autour du chemin critique w^* , celui-ci étant le plus court chemin reliant la position x occupée en t au domaine cible Γ . Sous la probabilité initiale, cet ensemble de trajectoires constitue un évènement rare dans la mesure où $\mathbb{Q}(\tau < T)$ tend vers 0. Le résultat (3.3) explicite la vitesse de convergence de cette probabilité vers 0 et montre qu'elle est directement reliée à la distance \mathfrak{D} du point x à Γ .

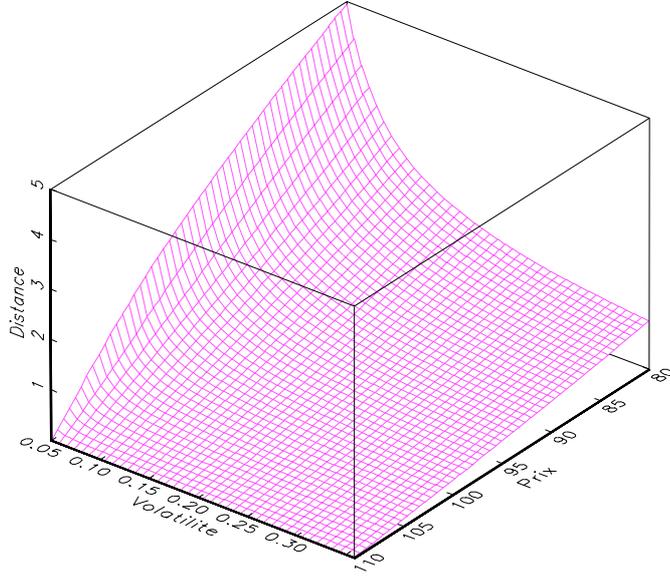
En appliquant le changement de dérive (3.5), on modifie la distribution de (X_t) de sorte que la probabilité d'occurrence de l'évènement $\{\tau < T\}$ reste suffisamment importante à mesure que l'on se rapproche de la maturité T . Le changement de dérive modifie donc l'allure des trajectoires de sorte qu'elles se propagent autour du chemin critique w^* .

Calcul numérique de \mathfrak{D} . La représentation de \mathfrak{D} sous la forme (3.6) permet de résoudre numériquement le calcul de \mathfrak{D} . On peut en effet appliquer une méthode de programmation dynamique en utilisant l'algorithme suivant (cf FOURNIÉ, LASRY et LIONS [6]) :

Algorithme 3.1 – *calcul numérique de \mathfrak{D}* –

1. *Discrétiser un domaine de l'espace d'état \mathbb{R}^d du processus (X_t) en une grille de points.*
2. *Initialiser les points de Γ appartenant à la grille à 0.*

3. Changement de dérive stochastique



Graphique 3.1. Distance \mathfrak{D} à $\Gamma = \{(s, \sigma) : s \geq 110\}$, Modèle de HESTON. $\rho = -0.6$, $\gamma = 0.1$.

3. Sélectionner un point x sur la grille et calculer, pour l'ensemble des points y dont on a déjà calculé les distances $\mathfrak{D}(y)$ à Γ , la distance $\mathfrak{D}(x, y) = \sqrt{(y-x)' A(x) (y-x)}$ associée à la métrique locale $A(x)$. Poser alors

$$\mathfrak{D}(x) = \min_y \mathfrak{D}(y) + \mathfrak{D}(x, y).$$

Notons que cette algorithmme permet de construire les chemins critiques correspondant à la métrique A . En effet, le point x étant sélectionné dans l'étape 3 de l'algorithme, le point y minimisant $\mathfrak{D}(y) + \mathfrak{D}(x, y)$ correspond au premier point succédant x sur le chemin critique allant de x à Γ .

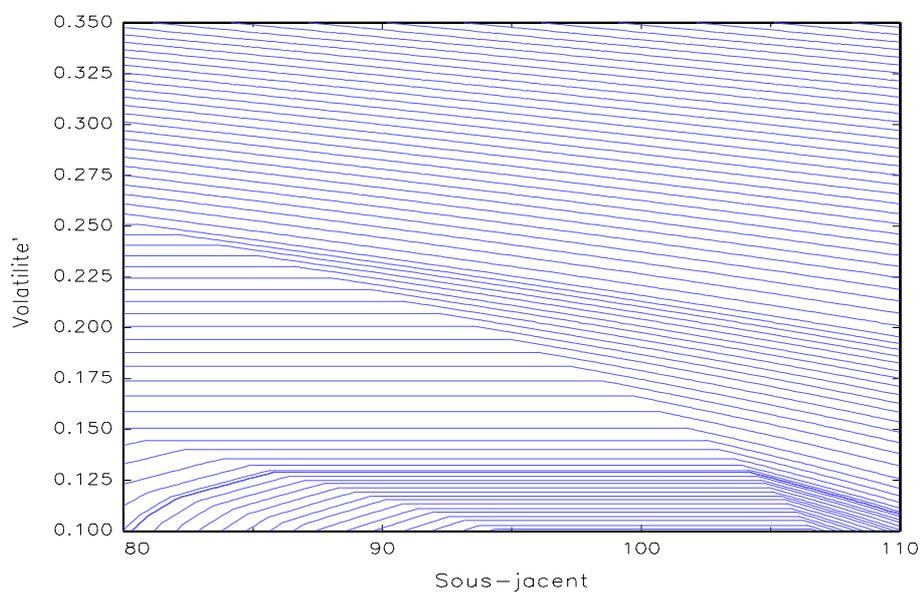
Exemple 3.1 Nous présentons dans cet exemple les résultats que l'on obtient dans le cas du modèle à volatilité stochastique de Heston [14]. $(X_t) = (S_t, V_t)$ est à valeurs dans \mathbb{R}^2 , solution de l'équation différentielle stochastique

$$\begin{aligned} \frac{dS_t}{S_t} &= rdt + \sqrt{V_t} dW_{1,t} \\ dV_t &= \kappa(\theta - V_t) dt + \gamma \sqrt{V_t} dW_{2,t} \end{aligned}$$

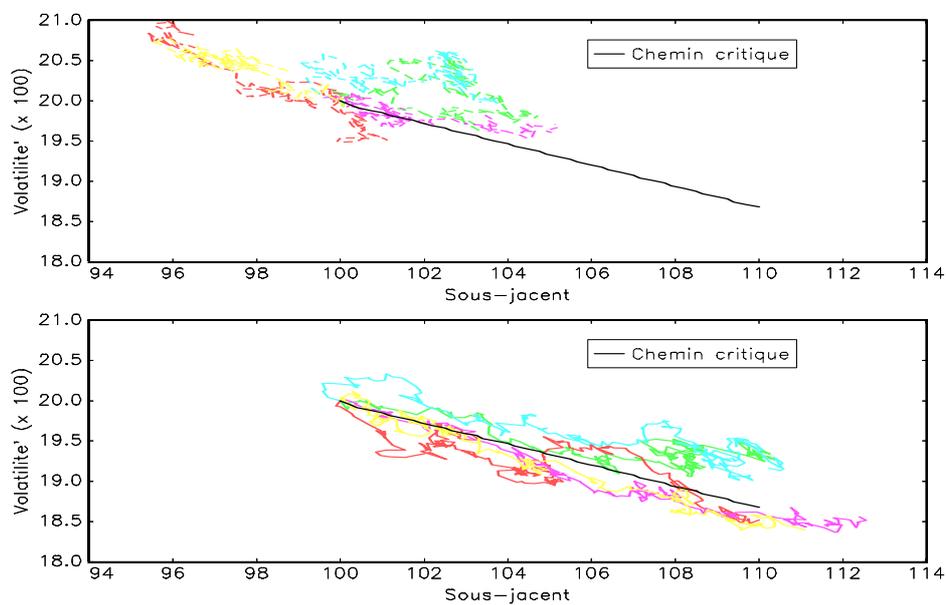
où $(W_{1,t})$ et $(W_{2,t})$ sont deux mouvements browniens standards unidimensionnels tels que

$$\langle dW_{1,t}, dW_{2,t} \rangle = \rho dt$$

3. Changement de dérive stochastique



Graphique 3.2. Chemins critiques associés à $\Gamma = \{(s, \sigma) : s \geq 110\}$, Modèle de HESTON. $\rho = -0.6, \gamma = 0.1$.



Graphique 3.3. Trajectoires simulées avant (pointillés) et après (traits pleins) changement de probabilité pour l'évaluation d'une option *UIB* de maturité de 7 jours et de barrière 110, Modèle de HESTON. $\rho = -0.6, \gamma = 0.1, r = 0, \kappa = 0.5, \theta = 0.2^2$. Initialement, $S_0 = 100$ et $\sigma_0 = 0.2$

3. Changement de dérive stochastique

Le graphique 3.1 page 28 représente la surface \mathfrak{D} que l'on obtient dans le cas où :

$$\begin{aligned}\gamma &= 0.1 \text{ et } \rho = -0.6 \\ \Gamma &= \{(s, v), s \geq 110\}\end{aligned}$$

Le graphique 3.2 page précédente correspond aux chemins critiques associés à ce même jeu de paramètres. La corrélation entre les deux mouvements browniens $(W_{1,t})$ et $(W_{2,t})$ étant négative, une augmentation de S_t sera en général accompagnée d'une diminution de la variance V_t . Ceci explique l'allure des chemins critiques : partant du point (s_0, σ_0) avec $s_0 < 110$, le domaine Γ est atteint plus rapidement en réduisant le niveau de la volatilité dans la mesure où cela induit généralement une hausse du sous-jacent.

Le graphique 3.3 page précédente illustre la façon dont les trajectoires sont modifiées en utilisant le changement de dérive (3.5). Nous avons représenté 5 trajectoires différentes pour t compris entre 0 et 7 jours, simulées pour évaluer une option *Up and In* de maturité restante égale à 7 jours. A la date initiale, $S_0 = 100$ et $\sigma_0 = 0.2$, et le domaine d'exercice correspond à $\Gamma = \{(s, \sigma), s \geq 110\}$. Le temps restant avant la maturité étant très court, la probabilité initiale de l'évènement Γ est très faible. Il sera donc très difficile de simuler des trajectoires conduisant à l'exercice de l'option. Sur le graphique, on voit que les 5 trajectoires simulées sous la probabilité initiale (représentées en pointillés) restent concentrées autour de la valeur initiale (S_0, σ_0) . Sous la nouvelle probabilité, on obtient, à partir de la même graine, 5 trajectoires (en traits pleins) qui se propagent autour du chemin critique (en noir).

Changement de dérive pour les calls européens et les options DIC

L'intérêt du changement de dérive que l'on vient de construire dans le cas de l'option *Up and In Bond* est de modifier la distribution des trajectoires afin d'accroître la probabilité de l'exercice de l'option. Dès lors un tel changement de dérive peut s'avérer intéressant dans un cadre plus général que celui de l'évaluation de $UIB(t, x)$: on peut l'utiliser pour évaluer toute option possédant un domaine d'exercice semblable à celui de l'option *Up an In*. C'est par exemple le cas de l'option d'achat européenne ou des options à barrière.

Si l'utilisation de ce changement de dérive ne prend pas en compte les spécificités de la distribution du payoff sur le domaine d'exercice, on peut néanmoins s'attendre à ce qu'elle induise une réduction de la variance dans la mesure où l'on simulera davantage de trajectoires appartenant au domaine d'exercice. Par suite, cette réduction devrait être d'autant plus importante que la probabilité d'exercer est faible, c'est-à-dire dans le cas des options très hors la monnaie.

Evaluation d'options d'achat européennes. Dans le cas du call européen, le domaine d'exercice s'exprime quasiment de la même façon que pour l'option binaire : le détenteur d'une telle option touche la différence, lorsqu'elle est positive, entre le prix S_T

3. Changement de dérive stochastique

du sous-jacent à la maturité T et le prix d'exercice K de l'option. Formulé dans le cadre d'un modèle à volatilité stochastique, le domaine d'exercice Γ s'écrit

$$\Gamma = \{(s, \sigma) \in \mathbb{R}_+^2, s \geq K\}.$$

Dans le cas des calls européens, on utilisera donc directement la formule (3.5) pour construire un estimateur de moindre variance que l'estimateur de Monte Carlo standard.

Application numérique 3.1 *Nous avons appliqué la technique de réduction de variance fondée sur l'utilisation directe du changement de dérive (3.5) à un ensemble de calls européens pour deux modèles à volatilité stochastique différents³ :*

- *Le premier modèle correspond au modèle de HESTON que l'on a déjà présenté dans l'exemple 3.1 page 28 . Nous reprenons ici le même jeu de paramètres : $r = 0$, $\kappa = 0.5$, $\theta = 0.2^2$, $\gamma = 0.1$ et $\rho = -0.6$.*
- *Dans le second modèle, on change l'équation différentielle caractérisant la diffusion de la volatilité au cours du temps. On suppose que $Y_t = \ln \sigma_t, t \geq 0$ est solution de*

$$dY_t = (c - kY_t) dt + \gamma dW_{2,t}$$

Le jeu de paramètres retenus en application est : $r = 0$, $k = 0$, $\gamma = 0.25$, $c = -\frac{\gamma^2}{2}$ et $\rho = -0.5$.

L'ensemble des options considérées ont un prix d'exercice de 110, le paramètre variant étant la maturité. Le sous-jacent vaut initialement 100 et la volatilité initiale est de 0.2.

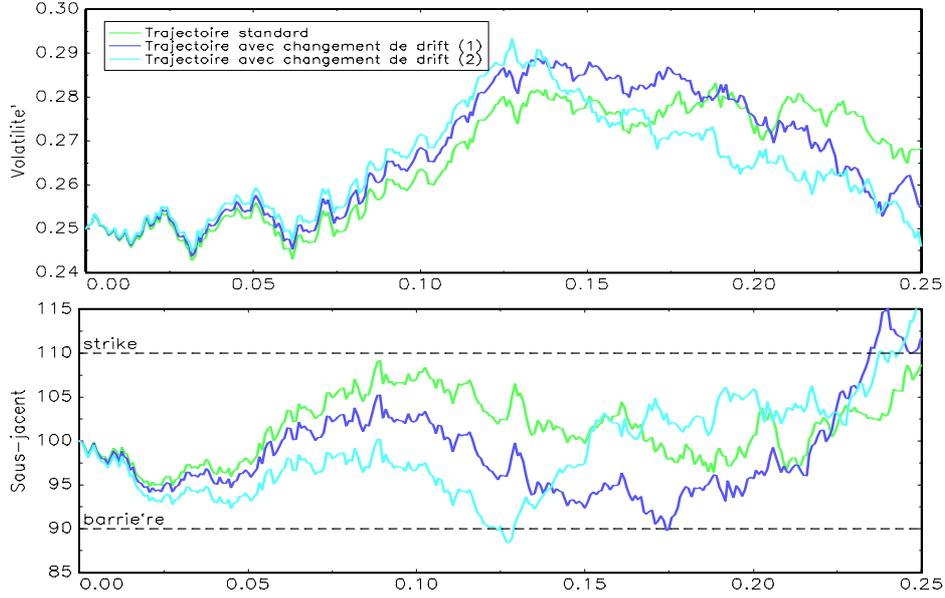
Les tableaux C.3 page 44 et C.5 page 45 situés en annexe de ce mémoire présentent les résultats obtenus.

Conformément à ce que l'on pouvait penser *a priori*, l'utilisation du changement de dérive (3.5) induit une réduction de variance, celle-ci étant d'autant plus importante que l'on se rapproche de la maturité. Néanmoins, le fait d'accroître le nombre de simulations appartenant au domaine d'exercice sans se préoccuper de la répartition du payoff sur ce domaine ne permet pas d'obtenir des réductions de variance très importantes. Par exemple, dans le cas du modèle de HESTON la réduction de variance est de 11.6 pour l'option de maturité 1 mois.

Evaluation d'options à barrière DIC. L'option à barrière DIC est une option d'achat européenne que le détenteur peut exercer si le sous-jacent passe sous une certaine barrière H inférieure au prix d'exercice K , avant la maturité du contrat. Le domaine d'exercice de l'option est donc plus complexe que dans le cas de l'option *Up and In* ou du call européen. Néanmoins, il est possible de construire un changement de dérive fondé sur la formule (3.5) en décomposant les trajectoires en deux portions :

³La spécification complète des deux modèles est rappelée à l'annexe B de ce mémoire.

3. Changement de dérive stochastique



Graphique 3.4. Changement de dérive appliqué à une option DIC, modèle de HESTON. $r = 0$, $\kappa = 0.5$, $\theta = 0.25^2$, $\gamma = 0.1$, $\rho = -0.6$.

- tant que la barrière H n'est pas atteinte, on utilise le changement de dérive (3.5) associé au domaine $\Gamma_H = \{(s, \sigma) \in \mathbb{R}_+^2, s \leq H\}$;
- une fois la barrière atteinte, on se retrouve dans le cas du call européen et on applique le changement de dérive (3.5) associé à $\Gamma_K = \{(s, \sigma) \in \mathbb{R}_+^2, s \geq K\}$.

Cette première approche accroît le nombre de trajectoires simulées permettant l'exercice de l'option. Néanmoins, lors du premier changement de dérive, on ne prend pas en compte la distance de la barrière au prix d'exercice. Avant que la barrière ne soit atteinte, on sous-estime donc la distance au domaine d'exercice, ce qui réduit la portée de la méthode.

Une deuxième approche consiste à utiliser un changement de dérive fondé sur la formule (3.5) pour atteindre la barrière H en intégrant la distance de la barrière au strike dans le calcul de la \mathcal{D} . Pour cela, on modifie la première étape de l'algorithme 3.1 page 27 en initialisant le domaine Γ_H à sa distance au domaine Γ_K . On prend ainsi en compte le fait qu'une fois la barrière atteinte, le processus devra remonter vers le prix d'exercice. En fait, une telle méthode s'avère assez naturelle compte tenu de l'interprétation des chemins critiques présentée précédemment : en procédant ainsi, on calcule la longueur du chemin critique pour un point n'ayant pas encore atteint la barrière.

Le graphique 3.4 illustre l'impact de la prise en compte de la distance de Γ_H à Γ_K dans la construction du changement de dérive : elle permet d'atteindre plus rapidement la barrière H , ce qui rend la remontée du processus vers le strike plus probable.

3. Changement de dérive stochastique

Application numérique 3.2 *On considère à nouveau les deux modèles à volatilité stochastique utilisés pour les applications numériques 3.1 . Seul le paramètre θ du modèle de Heston est modifié : on le fixe à 0.25.*

L'ensemble des options considérées ont un prix d'exercice de 110 et une barrière de 90, le paramètre variant étant la maturité. Le sous-jacent vaut initialement 100. La volatilité initiale est de 0.25 dans la cas du modèle de HESTON, elle vaut 0.2 dans le cas du second modèle.

Les tableaux C.7 page 46 et C.9 page 47 situés en annexe de ce mémoire présentent les résultats obtenus.

Remarque 3.1 *Notons que dans le cas des options à barrière, la discrétisation des trajectoires lors de la simulation induit un biais sur le prix à estimer.*

Considérons deux valeurs du sous-jacent simulées aux dates t et $t + \Delta t$, S_t et $S_{t+\Delta t}$, et supposons que la barrière H n'ait toujours pas été atteinte à la date t . Si $S_{t+\Delta t} > H$, il se peut néanmoins que la barrière ait été touchée entre t et $t + \Delta t$. Le fait de ne pas prendre en compte cette possibilité lors de la simulation conduit à un biais important sur le prix, même lorsque le pas de discrétisation est très fin.

Pour remédier à ce problème, nous avons utilisé une variable de Bernoulli de paramètre q_t égal à la probabilité que le sous-jacent passe sous la barrière H entre t et $t + \Delta t$ sachant S_t et $S_{t+\Delta t}$, et en supposant la volatilité constante entre t et $t + \Delta t$:

$$q_t = \exp \left(-\frac{1}{2\sigma_t^2 \Delta t} \ln \frac{H}{S_t} \ln \frac{H}{S_{t+\Delta t}} \right).$$

Si le résultat du tirage est 1, alors on considère que la barrière a été touchée.

Les commentaires que l'on peut formuler au vu des résultats obtenus sont similaires à ceux que l'on a fait dans le cas du call. Notons que les domaines d'exercice considérés étant plus rares pour l'option DIC que pour le call, on observe des réductions de variance supérieures : le nombre de trajectoires simulées entrant dans le domaine d'exercice joue un rôle très important sur la variance de la variable aléatoire simulée par la méthode de Monte Carlo standard.

D'ailleurs, il convient de noter que pour les très courtes maturités les réductions de variance affichées n'ont plus de réelle signification : en simulant les trajectoires par la méthode de Monte Carlo standard, le nombre de trajectoires simulées autorisant l'exercice de l'option devient négligeable et on ne parvient à estimer ni le prix ni la variance de l'estimateur de façon correcte. On peut à cet effet souligner le cas extrême de l'option DIC de maturité 1 mois, évaluée dans le cadre du modèle 2 : sur 100000 trajectoires simulées, aucune n'atteint le domaine d'exercice.

3. Changement de dérive stochastique

Conclusion

Il ressort des deux applications numériques que l'on vient de présenter que la méthode de Monte Carlo utilisant la formule (3.5) conduit à une réduction de variance par rapport à la méthode de Monte Carlo standard dès que le domaine d'exercice de l'option considérée est de faible probabilité : la majeure partie de la variance de l'estimateur standard provient alors du fait qu'il est très difficile de réaliser des tirages au sein du domaine d'exercice. Le fait de changer la dérive du processus permet de résoudre ce problème.

Néanmoins, on constate que pour les options peu éloignées de la monnaie, la méthode est peu efficace : dans ce cas la probabilité du domaine d'exercice est relativement importante, si bien que le problème de la simulation de trajectoires au sein du domaine d'exercice n'explique que de façon secondaire la variance de l'estimateur de Monte Carlo standard. Si l'on souhaite réduire la variance de façon plus significative pour les options proches de la monnaie, il apparaît nécessaire de davantage prendre en compte la distribution du payoff au sein du domaine d'exercice. C'est ce que permet dans une certaine mesure la méthode suivante, également fondée sur un changement de dérive stochastique.

3.3 Dérive de Black et Scholes

Dans la section 3.1 page 23, nous avons explicité le changement de dérive optimal (cf formule (3.1)), conduisant à un estimateur de variance nul, dans le cas où la fonction de payoff g ne dépend que de la valeur du processus sous-jacent à la maturité du contrat. Rappelons que l'utilisation directe de la formule (3.1) est impossible dans la mesure où le paramètre d'intérêt $I_g(t, x)$ est inconnu. L'approche exposée dans cette section consiste à approximer $I_g(t, x)$ par le paramètre que l'on obtiendrait en spécifiant un modèle de diffusion pour le processus (X_t) conduisant à une formule fermée : dans de nombreux cas on pourra considérer le modèle de BLACK et SCHOLES.

Justification empirique. Si l'on se réfère à la pratique des salles de marché, on peut penser qu'une telle approche améliorera la convergence des estimateurs. En effet, bien que la volatilité du rendement des actions ne soit pas déterministe, le modèle de BLACK et SCHOLES demeure très utilisé.

La justification d'une telle pratique repose en partie sur le fait que le modèle de BLACK et SCHOLES permet d'évaluer très facilement de nombreux produits dérivés. Bien entendu, cette raison ne peut être invoquée que dans la mesure où, de façon plus fondamentale, ce modèle constitue une bonne approximation du comportement des actifs sous-jacents. L'utilisation de modèles à volatilité stochastique permet de raffiner les résultats obtenus en utilisant le modèle de BLACK et SCHOLES : ils permettent par exemple de capturer l'effet smile observé sur les prix de marchés. De tels modèles cherchent à mieux appréhender le comportement des marchés : ils apparaissent comme une extension plus précise du

3. Changement de dérive stochastique

modèle de BLACK et SCHOLES qui demeure néanmoins le modèle fondamental sur le marché des actions.

Dès lors, on peut *a priori* s'attendre à ce que l'utilisation du modèle de BLACK et SCHOLES pour évaluer $I_g(t, x)$ conduise à une bonne approximation du comportement du paramètre d'intérêt lorsque le processus sous-jacent est un processus à volatilité stochastique.

Justification théorique. D'un point de vue mathématique, on peut se placer dans le cadre d'analyse suivant, proposé par FOURNIÉ, LASRY et TOUZI [7]. Considérons le processus (S_t, Y_t) , solution de l'équation différentielle stochastique :

$$\begin{aligned}\frac{dS_t}{S_t} &= rdt + \sigma(Y_t) dW_{1,t} \\ dY_t &= \eta(Y_t) + \gamma(Y_t) dW_{2,t}\end{aligned}$$

où $(W_{1,t})$ et $(W_{2,t})$ sont deux mouvements browniens standards tels que $\langle dW_{1,t}, dW_{2,t} \rangle = \rho dt$, et où η, σ, γ sont trois fonctions différentiables à tout ordre, dont les dérivées sont bornées. En faisant dépendre la dynamique du processus de volatilité d'un paramètre ϵ , ce processus apparaît comme appartenant à la famille de processus $(S_t^\epsilon, Y_t^\epsilon)$, solutions des équations différentielles stochastiques :

$$\begin{aligned}\frac{dS_t^\epsilon}{S_t^\epsilon} &= rdt + \sigma(Y_t^\epsilon) dW_{1,t} \\ dY_t^\epsilon &= \epsilon\eta(Y_t^\epsilon) + \sqrt{\epsilon}\gamma(Y_t^\epsilon) dW_{2,t}\end{aligned}$$

où $(W_{1,t})$ et $(W_{2,t})$ sont deux mouvements browniens standards tels que $\langle dW_{1,t}, dW_{2,t} \rangle = \epsilon\rho dt$. On dispose alors du résultat suivant : lorsque ϵ tend vers 0, le processus (S_t^ϵ) tend vers (S_t) dans le sens où

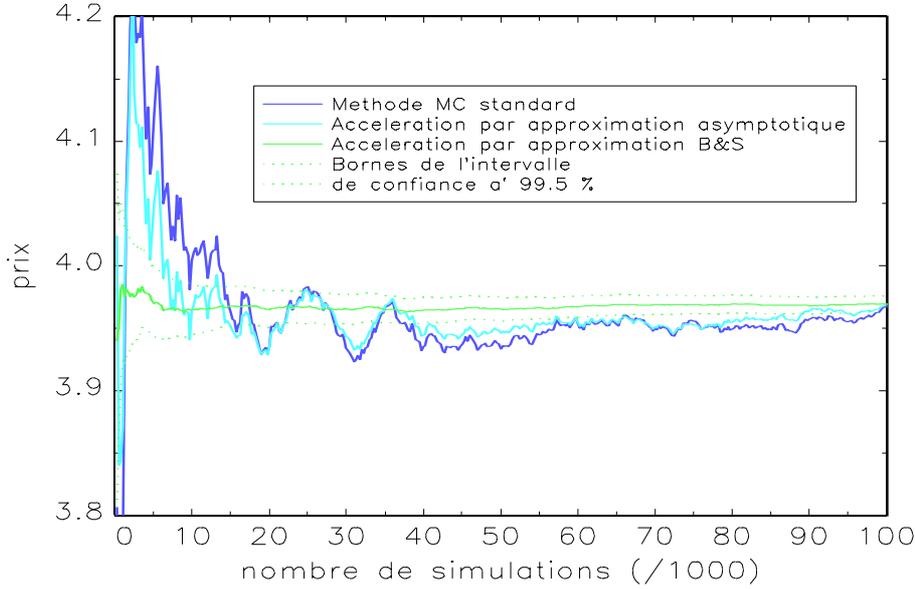
$$\lim_{\epsilon \rightarrow 0} \sup_{0 \leq t \leq T} \mathbb{E}^{\mathbb{Q}} |S_t^\epsilon - S_t^0| = 0.$$

Dès lors, le processus de diffusion BLACK-SCHOLES de paramètres $(r, \sigma(Y_0))$ apparaît comme une approximation asymptotique du processus (S_t, Y_t) initialement considéré. Une approximation naturelle du paramètre d'intérêt

$$I_g^\epsilon(t, s, y) = E^{\mathbb{Q}} [g(S_s^\epsilon, Y_s^\epsilon, t \leq s \leq T) | S_t^\epsilon = s, Y_t^\epsilon = y]$$

est donc obtenue en réalisant un développement asymptotique autour de la valeur $\epsilon = 0$. FOURNIÉ, LEBUCHOUX et TOUZI [7] ont établi un ensemble de résultats concernant les développements asymptotiques à tout ordre. Dans notre approche, nous nous limiterons au développement à l'ordre 0, ce qui revient à utiliser les prix BLACK-SCHOLES comme approximation du paramètre d'intérêt.

3. Changement de dérive stochastique



Graphique 3.5. Estimation du prix d'un call européen de maturité 1 an et de prix d'exercice 110, modèle de HESTON. $\rho = -0.6$, $\gamma = 0.1$, $r = 0$, $\kappa = 0.5$, $\theta = 0.2^2$. Initialement, $S_0 = 100$ et $\sigma_0 = 0.2$

Evaluation d'options d'achat européennes L'accélération de la convergence repose sur l'utilisation de la formule (3.1), le paramètre I_g et son gradient étant évalués en utilisant les formules classiques du prix et des grecs d'un call européen dans le modèle de BLACK-SCHOLES. Si l'on note

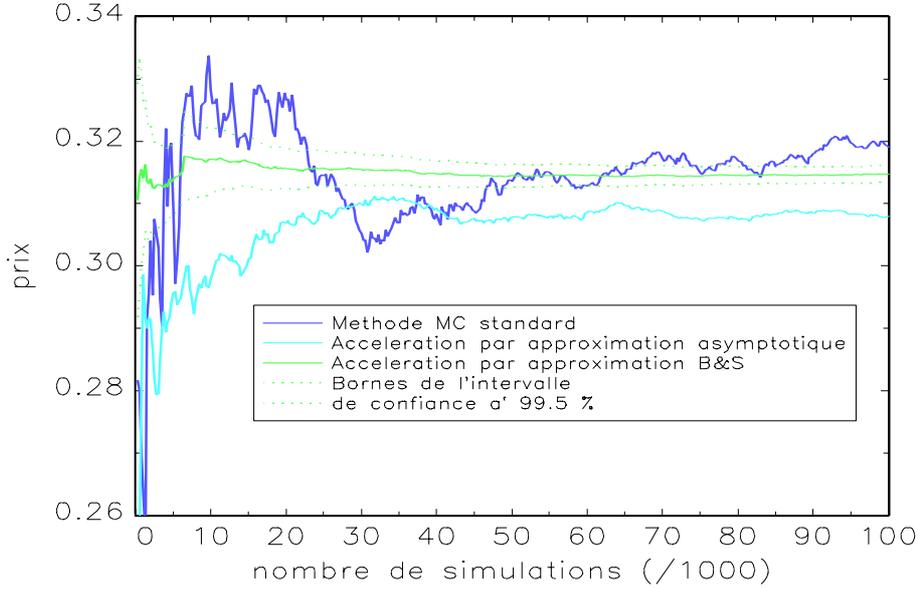
$$d_1 = \frac{\ln\left(\frac{s}{Ke^{-r(T-t)}}\right) + \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}},$$

le prix, le delta et le vega d'un call européen de maturité T et de strike K , valent à la date t , en fonction du prix s et de la volatilité σ :

$$\begin{aligned} CALL(t, s, K, T, \sigma) &= s\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_1 - \sigma\sqrt{T-t}) \\ \Delta_{CALL}(t, s, T, K, \sigma) &= \Phi(d_1) \\ v_{CALL}(t, s, T, K, \sigma) &= s\sqrt{T-t}\Phi'(d_1) \end{aligned}$$

Application numérique 3.3 On considère les deux mêmes modèles que ceux utilisés dans l'application numérique 3.1 page 31, en conservant des jeux de paramètres identiques. Les résultats obtenus sont présentés dans les tableaux C.4 page 44 et C.6 page 45, en annexe de ce mémoire. Le graphique 3.5 permet de comparer les vitesses de convergence associées à la méthode de Monte Carlo standard et aux deux méthodes de Monte Carlo accélérées étudiées dans cette partie, dans le cas d'un call de maturité 1 an.

3. Changement de dérive stochastique



Graphique 3.6. Estimation du prix d'une option DIC de maturité 6 mois, de barrière 90 et de prix d'exercice 110, modèle de HESTON. $\rho = -0.6$, $\gamma = 0.1$, $r = 0$, $\kappa = 0.5$, $\theta = 0.25^2$. Initialement, $S_0 = 100$ et $\sigma_0 = 0.25$

Evaluation d'options à barrière DIC Dans le cas de l'option à barrière DIC, les formules analytiques du prix, du delta et du vega dans le cas d'une modélisation BLACK-SCHOLES du sous-jacent sont les suivantes :

$$\begin{aligned}
 DIC(t, s, K, H, T, \sigma) &= \left(\frac{s}{H}\right)^{\gamma-1} \frac{K}{He^{r(T-t)}} PUT\left(s, \frac{(He^{r(T-t)})^2}{K}\right) \\
 \Delta_{DIC} &= \left(\frac{s}{H}\right)^{\gamma-1} \frac{K}{He^{r(T-t)}} \left[\frac{\gamma-1}{s} PUT\left(s, \frac{(He^{r(T-t)})^2}{K}\right) + \Delta_{PUT}\left(s, \frac{(He^{r(T-t)})^2}{K}\right) \right] \\
 v_{DIC} &= \left(\frac{s}{H}\right)^{\gamma-1} \frac{K}{He^{r(T-t)}} \left[v_{PUT}\left(s, \frac{(He^{r(T-t)})^2}{K}\right) - \frac{2}{\sigma} (\gamma-1) \ln \frac{S}{H} PUT\left(s, \frac{(He^{r(T-t)})^2}{K}\right) \right]
 \end{aligned}$$

avec $\gamma = 1 - \frac{2r}{\sigma^2}$, le prix et les grecs du put européen étant obtenus en utilisant le fait que le prix d'un put de strike K , pour un sous-jacent valant s en t , s'écrit

$$PUT(s, K) = e^{-r(T-t)} CALL\left(K, se^{2r(T-t)}\right).$$

Application numérique 3.4 On reprend ici l'application numérique 3.2 page 32. Les résultats obtenus sont présentés dans les tableaux C.8 page 46 et C.10 page 47, en annexe de ce mémoire. Le graphique 3.6 permet de comparer les vitesses de convergence associées à la méthode de Monte Carlo standard et aux deux méthodes de Monte Carlo accélérées étudiées dans cette partie, dans le cas d'une option à barrière DIC de maturité 6 mois.

Conclusion

Dans les deux types d'option considérés, on constate une nette amélioration des résultats par rapport à la méthode précédente, aussi bien pour les options proches de la monnaie que pour les options très hors la monnaie.

Conclusion

Les méthodes de Monte Carlo permettent de résoudre numériquement le calcul d'intégrales multiples suivant une approche probabiliste. Appliquées à la finance, ces méthodes permettent de calculer les prix de produits dérivés complexes, qu'on ne peut pas expliciter de façon analytique : en simulant un grand nombre de fois le prix d'un actif sous-jacent, on parvient à estimer le prix d'un produit dérivé en considérant la moyenne empirique, sur l'échantillon simulé, de ses flux futurs actualisés.

L'utilisation d'une telle méthode requiert un bon contrôle de la convergence de l'estimateur du prix ainsi obtenu. Ce contrôle peut s'effectuer soit en augmentant le nombre de simulations, ce qui peut s'avérer coûteux en temps de calcul, soit en mettant en place des méthodes de réduction de la variance des variables aléatoires simulées.

Dans ce mémoire, nous avons présenté trois méthodes de réduction de la variance appliquées aux processus de diffusion, reposant sur la même idée directrice. La plupart des options sont caractérisées par une fonction de payoff positive ou nulle, la fonction étant strictement positive sur le domaine d'exercice. La simulation de trajectoires en dehors du domaine d'exercice apporte peu d'information et conduit donc à une variance relativement élevée de l'estimateur. L'idée est alors de modifier la fonction de dérive du processus sous-jacent afin d'obtenir davantage de trajectoires autorisant l'exercice. En procédant ainsi, on modifie bien entendu la loi du processus et il convient par conséquent de pondérer les variables aléatoires obtenues afin de ne pas biaiser l'estimateur du paramètre d'intérêt.

Si le théorème de Girsanov permet de poser le cadre théorique du changement de dérive, il n'apporte pas d'information sur la dérive à choisir pour réduire la variance de façon significative. Il convient donc de se référer à d'autres types de résultats. Dans les trois cas que nous abordons, nous avons construit les changements de dérive à partir de résultats asymptotiques, deux d'entre eux étant issus de la théorie des Grandes Déviations.

La première approche que nous avons considérée consiste à utiliser un changement de dérive déterministe, indépendant des trajectoires simulées. La nouvelle dérive est déterminée en exploitant la méthode de Laplace abstraite. Cette technique d'accélération présente l'avantage de pouvoir être améliorée en la couplant à une méthode de stratification. Les résultats numériques que nous avons obtenus dans le cas de modèles à volatilités stochastiques montrent que cette méthode est relativement efficace. Néanmoins, une condition nécessaire à son utilisation est que le domaine d'exercice de l'option considérée soit convexe, ce qui limite la portée de la méthode. On ne peut notamment pas l'appliquer dans le cas des options à barrière.

Conclusion

La deuxième méthode que nous avons mise en place repose sur un changement de dérive stochastique. Ce changement transforme les trajectoires de sorte qu'une fois modifiées, elles se propagent autour du seul chemin qui, asymptotiquement, permet d'atteindre le domaine d'exercice. Les applications numériques que nous avons réalisées montrent que cette méthode est efficace lorsqu'une proportion importante de la variance des variables simulées provient du fait que l'exercice de l'option est un évènement de faible probabilité. Si ce n'est pas le cas, le taux de réduction de la variance est minime, la variance provenant alors essentiellement de la distribution du payoff sur le domaine d'exercice et non de la taille de ce domaine.

La troisième méthode que nous avons implémentée permet, dans certains cas, de résoudre ce problème. Elle présente un intérêt en finance lorsque l'on spécifie la dynamique du sous-jacent sous la forme d'un modèle à volatilité stochastique. La modélisation BLACK-SCHOLES conduit alors à une bonne approximation du comportement du sous-jacent sur le domaine d'exercice. Si l'on dispose d'une formule d'évaluation fermée dans ce modèle simplifié, on peut construire un changement de dérive exploitant cette information supplémentaire. Une telle méthode produit des taux de réduction de la variance significatifs, même pour des options proches de la monnaie.

Annexes

A Principe de Grandes Déviations

Soient (Ω, \mathcal{F}) un espace Polonais (espace métrique complet séparable, muni de sa tribu borélienne) et $(\mu_t)_{t>0}$ une famille de probabilités convergeant vers la masse de Dirac δ_m , où m est un point de Ω . Soit alors $C \in \mathcal{F}$ tel que $m \notin \overline{C}$, on a :

$$\mu_t(C) \xrightarrow[t \rightarrow 0]{} 0$$

Le borélien C est qualifié d'évènement "rare" ; la théorie des grandes déviations s'intéresse à la vitesse de convergence vers 0 de $\mu_t(C)$ lorsque t tend vers 0.

Définition A.1 Principe de Grandes déviations

On dira que $(\mu_t)_{t>0}$ suit un principe de grandes déviations s'il existe $v : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ et $\lambda : \Omega \rightarrow \mathbb{R}_+ \cup \{+\infty\}$ telles que :

1. $\lim_{t \rightarrow 0} v(t) = +\infty$
2. Pour tout $s \in \mathbb{R}_+$, $\{\lambda \leq s\}$ est compact dans Ω .
3. Pour tout $C \in \mathcal{F}$, on a :

$$-\text{Inf}_{\omega \in \overset{\circ}{C}} \lambda(\omega) \leq \underline{\lim}_{t \rightarrow 0} \frac{1}{v(t)} \log(\mu_t(C)) \leq \overline{\lim}_{t \rightarrow 0} \frac{1}{v(t)} \log(\mu_t(C)) \leq -\text{Inf}_{\omega \in \overline{C}} \lambda(\omega)$$

Le couple (v, λ) , ou, plus simplement, λ est appelé "fonctionnelle d'action" pour la famille $(\mu_t)_{t>0}$

Par exemple, dans la section (2), les processus discrétisés sont gouvernés par un vecteur gaussien. Or si Z est un vecteur gaussien dans \mathbb{R}^d , centré, de matrice de variance-covariance l'identité, alors sur \mathbb{R}^d munie de sa tribu borélienne, on définit la famille de probabilités $(\alpha_t)_{t>0}$ par $\alpha_t(\cdot) = P(\sqrt{t}Z \in \cdot)$. $(\alpha_t)_{t>0}$ suit alors un principe de grandes déviations avec pour fonctionnelle d'action $(t, \frac{1}{2}z'z)$. Plus précisément, on a la convergence suivante :

$$\lim_{t \rightarrow 0} t \ln P(\sqrt{t}Z \in \mathcal{B}) = - \min_{z \in \mathcal{B}} \frac{1}{2} z'z$$

où \mathcal{B} est un borélien de \mathbb{R}^d . La théorie des grandes déviations s'intéresse donc à la distribution asymptotique de la famille de probabilités $(\alpha_t)_{t>0}$.

Nous souhaitons en fait utiliser une approximation non pas de la probabilité d'appartenir au borélien \mathcal{B} , mais de l'espérance d'une variable aléatoire. Le théorème suivant permet de réaliser, à partir d'un principe de grandes déviations, cette approximation.

A. Principe de Grandes Déviations

Théorème A.1 (*Méthode de Laplace abstraite*)

Sur un espace polonais (Ω, \mathcal{F}) , soit $(\mu_t)_{t>0}$ une famille de probabilités vérifiant un Principe de Grandes Déviations associé à la fonctionnelle d'action (v, λ) . Soit θ une application mesurable de Ω dans \mathbb{R} telle que :

1. $\exp(v(t)\theta) \in L^1(\mu_t) \quad (t > 0)$
2. θ est continue en tout point de $\{\lambda < +\infty\}$
- 3.

$$\lim_{L \rightarrow +\infty} \overline{\lim}_{t \rightarrow 0} \frac{1}{v(t)} \log \left\{ \int_{(\theta \geq L)} \exp(v(t)\theta) d\mu_t \right\} = -\infty$$

alors :

$$\lim_{t \rightarrow 0} \frac{1}{v(t)} \ln \int \exp(v(t)\theta) d\mu_t = \sup_w \{\theta(w) - \lambda(w)\}$$

De plus, on montre alors que la condition :

$$(\exists q > 1) \left(\limsup_{t \rightarrow 0} \frac{1}{v(t)} \ln \left(\int \exp(qv(t)\theta) d\mu_t \right) < +\infty \right) \quad (\text{A.1})$$

entraîne la condition (3) du théorème. L'application de ce théorème permet de déduire alors une approximation de la quantité (en reprenant les notations de la section 2) :

$$\mathbb{E}^{\mathbb{Q}} \left[\exp \left(2F(Z) + \varphi'Z + \frac{1}{2}\varphi'\varphi \right) \mathbf{1}_D \right], \quad Z \sim \mathcal{N}(0, I) \text{ sous } \mathbb{Q}.$$

Puisque la famille de probabilité $(\alpha_t)_{t>0}$ défini par $\alpha_t(\cdot) = P(\sqrt{t}Z \in \cdot)$ suit un principe de grandes déviations associé à la fonctionnelle $(t, z \rightarrow \frac{1}{2}z'z)$, on va pouvoir appliquer le théorème précédent pour obtenir une approximation asymptotique de la quantité recherchée. En effet si $F(Z) \leq A + cZ'Z$ avec $c < \frac{1}{4}$, la condition ?? est alors vérifiée, et on en conclut que

$$t \ln \mathbb{E}^{\mathbb{Q}} \left[e^{\frac{1}{t}(2F(\sqrt{t}Z) - \sqrt{t}\varphi'Z + \frac{1}{2}\varphi'\varphi)} \mathbf{1}_D \right] \xrightarrow{t \rightarrow 0} \max_{z \in D} \left\{ 2F(z) - \varphi'z + \frac{1}{2}\varphi'\varphi - \frac{1}{2}z'z \right\}$$

B. Changement de dérive déterministe et stratification : résultats numériques

B Changement de dérive déterministe et stratification : résultats numériques

Les résultats numériques présentés dans cette annexe illustrent les performances des méthodes d'accélération de la convergence présentées dans la deuxième partie de ce mémoire, appliquées au modèle à volatilité stochastique de HULL et WHITE : le processus $(X_t) = (S_t, V_t)$ à valeurs dans \mathbb{R}^2 est solution de l'équation différentielle stochastique

$$\begin{aligned}\frac{dS_t}{S_t} &= rdt + \sqrt{V_t}dW_{1,t} \\ dV_t &= \nu V_t dt + \xi V_t dW_{2,t}\end{aligned}$$

où $(W_{1,t})$ et $(W_{2,t})$ sont deux mouvements browniens standards unidimensionnels tels que

$$\langle dW_{1,t}, dW_{2,t} \rangle = \rho dt.$$

T	ξ	Prix	Réduction de variance	
			Fonction d'importance	Fonction d'importance et stratification
1	0.2	1.44	15	864
0.5		0.538	22	1963
0.25		0.134	41	5177
1	0	1.49	15	1633
0.5		0.564	21	3317
0.25		0.143	41	7847

TAB. B.1. Evaluation d'options asiatiques, modèle de HULL et WHITE. Nombre de simulations : 10^5 ; nombre de strates : 50; Prix d'exercice : 110; Prix initial : 100; Volatilité initiale : 0.2; $r = 0$, $\nu = 0$, $\rho = -0.5$.

T	ξ	Prix	Réduction de variance	
			Fonction d'importance	Fonction d'importance et stratification
1	0.2	4.17	11	569
0.5		2.14	13	1070
0.25		0.92	17	2197
1	0	4.28	12	1358
0.5		2.21	13	2332
0.25		0.95	18	4300

TAB. B.2. Evaluation d'options asiatiques, modèle de HULL et WHITE. Nombre de simulations : 10^5 ; nombre de strates : 50; Prix d'exercice : 110; Prix initial : 100; Volatilité initiale : 0.2; $r = 0$, $\nu = 0$, $\rho = -0.5$.

C Changement de dérive stochastique : résultats numériques

Les résultats numériques présentés dans cette annexe illustrent les performances des méthodes d'accélération de la convergence présentées dans la troisième partie de ce mémoire, appliquées aux deux modèles à volatilité stochastique suivants :

Modèle 1 : modèle de Heston [14]. Le processus $(X_t) = (S_t, V_t)$ à valeurs dans \mathbb{R}^2 est solution de l'équation différentielle stochastique

$$\begin{aligned}\frac{dS_t}{S_t} &= rdt + \sqrt{V_t}dW_{1,t} \\ dV_t &= \kappa(\theta - V_t)dt + \gamma\sqrt{V_t}dW_{2,t}\end{aligned}$$

où $(W_{1,t})$ et $(W_{2,t})$ sont deux mouvements browniens standards unidimensionnels tels que

$$\langle dW_{1,t}, dW_{2,t} \rangle = \rho dt$$

Modèle 2. Le processus $(X_t) = (S_t, Y_t)$ à valeurs dans \mathbb{R}^2 est solution de l'équation différentielle stochastique

$$\begin{aligned}\frac{dS_t}{S_t} &= rdt + e^{Y_t}dW_{1,t} \\ dY_t &= (c - kY_t)dt + \gamma dW_{2,t}\end{aligned}$$

où $(W_{1,t})$ et $(W_{2,t})$ sont deux mouvements browniens standards unidimensionnels tels que

$$\langle dW_{1,t}, dW_{2,t} \rangle = \rho dt$$

On pourra se référer à FOURNIÉ, LASRY et TOUZI [7] pour d'autres résultats numériques portant sur ce type de modèle.

C. Changement de dérive stochastique : résultats numériques

C.1 Evaluation d'options d'achat européennes : modèle 1 (Heston)

Maturité	Prix estimé (*10 ²)	Ecart-type (*10 ²) (% du prix)	Réduction de variance
1	396.8	2.11 (0.53%)	1.7
$\frac{1}{2}$	202.5	1.10 (0.54%)	2.1
$\frac{1}{4}$	85.50	0.48 (0.56%)	3.2
$\frac{1}{6}$	45.05	0.26 (0.57%)	4.5
$\frac{1}{12}$	10.31	0.062 (0.60%)	11.6
$\frac{1}{15}$	5.534	0.034 (0.61%)	17.7
$\frac{1}{18}$	3.093	0.019 (0.62%)	26.5
$\frac{1}{20}$	2.134	0.013 (0.63%)	34.2

TAB. C.3. Evaluation de calls européens, modèle de Heston : accélération de la convergence par approximation asymptotique.

Nombre de simulations : 10^5 ; Prix d'exercice : 110; Prix initial : 100; Volatilité initiale : 0.2; $r = 0$, $\kappa = 0.5$, $\theta = 0.2^2$, $\gamma = .1$, $\rho = -0.6$.

Maturité	Prix estimé (*10 ²)	Ecart-type (*10 ⁴) (% du prix)	Réduction de variance
1	396.9	23.00 (0.058%)	140.2
$\frac{1}{2}$	202.4	11.01 (0.054%)	211.8
$\frac{1}{4}$	85.49	4.78 (0.056%)	320.2
$\frac{1}{6}$	45.07	2.65 (0.059%)	431.3
$\frac{1}{8}$	26.22	1.62 (0.062%)	561.5
$\frac{1}{10}$	16.15	1.04 (0.065%)	719.0
$\frac{1}{12}$	10.33	0.70 (0.067%)	911.8
$\frac{1}{14}$	6.792	0.47 (0.070%)	1149.3
$\frac{1}{16}$	4.555	0.33 (0.073%)	1442.6
$\frac{1}{18}$	3.104	0.23 (0.075%)	1803.7
$\frac{1}{20}$	2.143	0.16 (0.077%)	2248.5
$\frac{1}{22}$	1.149	0.12 (0.079%)	2793.2
$\frac{1}{24}$	1.052	0.08 (0.082%)	3460.0

TAB. C.4. Evaluation de calls européens, modèle de Heston : accélération de la convergence par approximation BLACK-SCHOLES.

Nombre de simulations : 10^5 ; Prix d'exercice : 110; Prix initial : 100; Volatilité initiale : 0.2; $r = 0$, $\kappa = 0.5$, $\theta = 0.2^2$, $\gamma = 0.1$, $\rho = -0.6$.

C. Changement de dérive stochastique : résultats numériques

C.2 Evaluation d'options d'achat européennes : modèle 2

Maturité	Prix estimé (*10)	Ecart-type (*100) (% du prix)	Réduction de variance
1	40.58	2.26 (0.55%)	1.55
$\frac{1}{2}$	20.71	1.19 (0.57%)	1.93
$\frac{1}{4}$	8.773	0.53 (0.60%)	2.78
$\frac{1}{6}$	4.640	0.29 (0.62%)	3.85
$\frac{1}{8}$	2.709	0.17 (0.63%)	5.25
$\frac{1}{10}$	1.674	0.11 (0.65%)	7.03
$\frac{1}{12}$	1.074	0.07 (0.65%)	9.36

TAB. C.5. Evaluation de calls européens, modèle 2 : accélération de la convergence par approximation asymptotique.

Nombre de simulations : 10^5 ; Prix d'exercice : 110 ; Prix initial : 100 ; Volatilité initiale : 0.2 ; $r = 0$, $k = 0$, $\gamma = 0.25$, $c = -\frac{\gamma^2}{2}$, $\rho = -0.5$.

Maturité	Prix estimé (*10 ²)	Ecart-type (*10 ⁴) (% du prix)	Réduction de variance
1	406.4	19.24 (0.047%)	214.1
$\frac{1}{2}$	207.4	9.91 (0.048%)	277.3
$\frac{1}{4}$	84.84	4.46 (0.051%)	387.8
$\frac{1}{6}$	46.47	2.51 (0.054%)	509.0
$\frac{1}{8}$	27.13	1.54 (0.057%)	653.0
$\frac{1}{10}$	16.78	1.00 (0.060%)	828.4
$\frac{1}{12}$	10.78	0.67 (0.062%)	1044.7
$\frac{1}{14}$	7.108	0.46 (0.065%)	1312.3
$\frac{1}{16}$	4.785	0.32 (0.067%)	1642.4
$\frac{1}{18}$	3.273	0.23 (0.069%)	2047.0
$\frac{1}{20}$	2.268	0.16 (0.071%)	2544.8
$\frac{1}{22}$	1.588	0.12 (0.073%)	3156.5
$\frac{1}{24}$	1.122	0.08 (0.075%)	3905.2

TAB. C.6. Evaluation de calls européens, modèle 2 : accélération de la convergence par approximation BLACK-SCHOLES.

Nombre de simulations : 10^5 ; Prix d'exercice : 110 ; Prix initial : 100 ; Volatilité initiale : 0.2 ; $r = 0$, $k = 0$, $\gamma = 0.25$, $c = -\frac{\gamma^2}{2}$, $\rho = -0.5$.

C. Changement de dérive stochastique : résultats numériques

C.3 Evaluation d'options DIC : modèle 1 (Heston)

Maturité	Prix estimé (*10 ²)	Ecart-type (*10 ³) (% du prix)	Réduction de variance
1	133.6	9.54 (0.71%)	3.38
$\frac{1}{2}$	30.79	2.42 (0.78%)	7.60
$\frac{1}{4}$	3.246	0.290 (0.89%)	28.90
$\frac{1}{6}$	0.458	0.0466 (1.02%)	117.2
$\frac{1}{8}$	0.072	0.00815 (1.14%)	741.9
$\frac{1}{10}$	0.013	0.00175 (1.38%)	2016.8
$\frac{1}{12}$	0.002	0.00033 (1.43%)	62.95

TAB. C.7. Evaluation de DIC, modèle de Heston : accélération de la convergence par approximation asymptotique.

Nombre de simulations : 10^5 ; Prix d'exercice : 110 ; Barrière : 90 ; Prix initial : 100 ; Volatilité initiale : 0.25 ; $r = 0$, $\kappa = 0.5$, $\theta = 0.25^2$, $\gamma = .1$, $\rho = -0.6$.

Maturité	Prix estimé (*10 ²)	Ecart-type (*10 ³) (% du prix)	Réduction de variance
2	373.0	6.42 (0.17%)	33.6
1	135.4	4.46 (0.33%)	15.4
$\frac{1}{2}$	31.47	0.497 (0.15%)	187.2
$\frac{1}{4}$	3.262	0.0527 (0.16%)	872.4
$\frac{1}{6}$	0.457	$8.21e - 3$ (0.18%)	3782.6
$\frac{1}{8}$	0.074	$1.60e - 3$ (0.22%)	19278
$\frac{1}{10}$	0.013	$2.80e - 4$ (0.22%)	78983
$\frac{1}{12}$	0.002	$5.51e - 5$ (0.24%)	2228.3

TAB. C.8. Evaluation de DIC, modèle de Heston : accélération de la convergence par approximation BLACK-SCHOLES.

Nombre de simulations : 10^5 ; Prix d'exercice : 110 ; Barrière : 90 ; Prix initial : 100 ; Volatilité initiale : 0.25 ; $r = 0$, $\kappa = 0.5$, $\theta = 0.25^2$, $\gamma = .1$, $\rho = -0.6$.

C. Changement de dérive stochastique : résultats numériques

C.4 Evaluation d'options DIC : modèle 2

Maturité	Prix estimé (*10 ³)	Ecart-type (*10 ³) (% du prix)	Réduction de variance
1	616.7	5.97 (0.96%)	3.0
$\frac{1}{2}$	92.26	1.16 (1.25%)	7.9
$\frac{1}{3}$	18.89	0.28 (1.49%)	17.1
$\frac{1}{4}$	4.414	0.08 (1.82%)	27.8
$\frac{1}{6}$	0.285	6.63e - 3 (2.32%)	185.1
$\frac{1}{9}$	5.970e - 3	2.66e - 4 (4.45%)	2552
$\frac{1}{12}$	1.396e - 4	9.04e - 6 (6.48%)	0

TAB. C.9. Evaluation de DIC, modèle 2 : accélération de la convergence par approximation asymptotique.

Nombre de simulations : 10^5 ; Prix d'exercice : 110; Barrière : 90; Prix initial : 100; Volatilité initiale : 0.2; $r = 0$, $k = 0$, $\gamma = 0.25$, $c = -\frac{\gamma^2}{2}$, $\rho = -0.5$.

Maturité	Prix estimé (*10 ³)	Ecart-type (*10 ³) (% du prix)	Réduction de variance
2	2141.5	9.05 (0.42%)	7.6
1	616.7	3.02 (0.49%)	11.8
$\frac{1}{2}$	93.48	0.33 (0.35%)	98.9
$\frac{1}{3}$	18.88	7.56e - 2 (0.40%)	240.1
$\frac{1}{4}$	4.341	1.75e - 2 (0.40%)	585.8
$\frac{1}{5}$	1.071	4.81e - 3 (0.45%)	2301.6
$\frac{1}{6}$	0.282	1.42e - 3 (0.50%)	4064.6
$\frac{1}{7}$	7.560e - 2	3.98e - 4 (0.53%)	692.8
$\frac{1}{8}$	2.081e - 2	1.11e - 4 (0.53%)	31.1
$\frac{1}{9}$	5.939e - 3	3.66e - 5 (0.62%)	134839
$\frac{1}{10}$	1.691e - 3	1.20e - 5 (0.71%)	0
$\frac{1}{11}$	4.942e - 4	3.36e - 6 (0.68%)	0
$\frac{1}{12}$	1.456e - 4	1.21e - 6 (0.83%)	0

TAB. C.10. Evaluation de DIC, modèle 2 : accélération de la convergence par approximation BLACK-SCHOLES.

Nombre de simulations : 10^5 ; Prix d'exercice : 110; Barrière : 90; Prix initial : 100; Volatilité initiale : 0.2; $r = 0$, $k = 0$, $\gamma = 0.25$, $c = -\frac{\gamma^2}{2}$, $\rho = -0.5$.

D. Variable de contrôle : le cas d'une option sur panier

D Variable de contrôle : le cas d'une option sur panier

Cet exemple illustre la méthode d'accélération de la convergence à partir d'une variable de contrôle, présentée dans la première partie de ce rapport, page 6. On s'intéresse à un produit dérivé défini à partir du cours à la date T de d actions sous-jacentes S_T^1, \dots, S_T^d . Le payoff de ce produit s'écrit

$$X_T = \max(A_T - K, 0)$$

où

$$A_T = \frac{1}{d} \sum_{i=1}^d S_T^i.$$

En supposant que le taux d'intérêt instantané r est constant sur la période considérée, le prix de cette actif s'écrit

$$P_A = \mathbb{E}e^{-rT} X_T,$$

l'espérance étant calculée sous la probabilité risque-neutre du marché.

On suppose que l'ensemble des actions considérées suivent un processus de diffusion BLACK-SCHOLES sous la probabilité risque-neutre. Formellement, on considère un mouvement brownien standard sur \mathbb{R}^d , noté $(B_t)_{t \geq 0}$, $\Sigma = (\sigma_i)_{1 \leq i \leq d}$ un vecteur de coordonnées strictement positives et $\Omega = (\rho_{ij})_{1 \leq i, j \leq d}$ une matrice de corrélation. On note Γ la matrice associée à Ω résultant de la décomposition de Cholesky⁴ de cette dernière. On modélise alors l'évolution des d actions composant le panier sous-jacent par le processus aléatoire sur \mathbb{R}^d , $(S_t^1, \dots, S_t^d)_{t \geq 0}$, caractérisé par l'équation différentielle stochastique suivante :

$$dS_t^i = rdt + \sigma_i dW_t^i, i \in \{1, \dots, d\}$$

où $W_t = \Gamma B_t, t \geq 0$, de telle sorte que

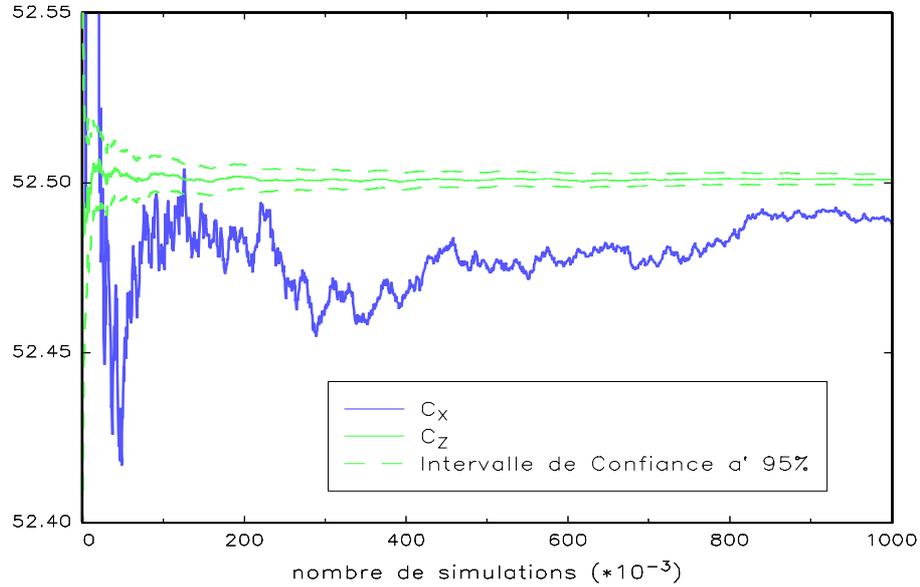
$$\forall i, j \in \{1, \dots, N\}, \langle W^i, W^j \rangle_t = \rho_{ij} t.$$

Un premier estimateur sans biais de ce paramètre, noté c_X , peut être obtenu par une méthode de Monte Carlo standard, en simulant N vecteurs (B_t^1, \dots, B_t^d) et en calculant les N valeurs (x_t^1, \dots, x_t^N) prises par la variable aléatoire X_t . On pose alors :

$$c_X = \frac{e^{-rT}}{N} \sum_{i=1}^N x_T^i$$

⁴ Γ est l'unique matrice triangulaire inférieure telle que $\Gamma\Gamma' = \Omega$.

D. Variable de contrôle : le cas d'une option sur panier



Graphique D.1. Accélération par variable de contrôle, $K = 50$, $\beta = 1.042$

La variance de cet estimateur est

$$\text{var}(c_X) = \frac{e^{-2rt}}{N} \text{var}(X_T)$$

L'utilisation d'une variable de contrôle permet de réduire cette variance pour un nombre de simulations donné. On considère pour cela un nouveau produit dérivé défini non plus à partir de la moyenne arithmétique A_T des sous-jacents, mais à partir de la moyenne géométrique G_T

$$G_T = \sqrt[d]{\prod_{i=1}^d S_T^i}$$

Dans un cadre d'analyse BLACK-SCHOLES, le prix

$$P_G = \mathbb{E}e^{-rT} (G_T - K)_+$$

est connu sous forme analytique. On peut facilement établir que

$$P_G = G_0 e^{(\tilde{r}-r)T} \Phi(d_1) - K e^{-rT} \Phi(d_2)$$

où

$$\begin{aligned} d_1 &= \frac{\ln\left(\frac{G_0}{K}\right) + (\tilde{r} + \frac{1}{2}\tilde{\sigma}^2)T}{\tilde{\sigma}\sqrt{T}} & \tilde{\sigma} &= \frac{1}{d}\sqrt{\Sigma'\Omega\Sigma} \\ d_2 &= d_1 - \tilde{\sigma}\sqrt{T} & \tilde{r} &= r + \frac{1}{2}\tilde{\sigma}^2 - \frac{1}{2d}\sum\sigma_i^2 \end{aligned}$$

D. Variable de contrôle : le cas d'une option sur panier

Le graphique D.1 page précédente représente le comportement de l'estimateur standard c_X et de l'estimateur contrôlé c_Z pour le jeu de paramètres suivant : $T = 1$; $r = 0$; $d = 10$; $S_0 = \{80, 85, 90, 95, 100, 105, 110, 115, 120, 125\}$; $\Sigma = \{0.15, 0.16, 0.17, 0.18, 0.19, 0.20, 0.21, 0.22, 0.23, 0.24\}$; $\rho_{ii} = 1$, $\rho_{ij} = 0.5$ si $i \neq j$; $K = 50$. La variance de l'estimateur contrôlé c_Z est 360 fois plus petite que la variance de l'estimateur standard c_X . Notons que le taux de réduction de la variance dépend du jeu de paramètres. Si l'on considère une option à la monnaie ($K = 102.5$), on obtient une réduction dans un rapport de 100, le rapport s'établissant à 10 dans le cas d'une option très hors la monnaie ($K = 150$).

Références

- [1] AZENCOTT, R., “Formules de Taylor stochastiques et développement asymptotique d’intégrales de Feynman”, *Séminaire de Probabilité XVI, Lecture Notes in Mathematics*, 921, 237–285, 1981.
- [2] BLACK F, AND C. SCHOLES, “The pricing of options and corporate liabilities”, *Journal of Political Economy*, 81, 635–654, 1973.
- [3] CASELLA, G, AND C.P. ROBERT, “Monte Carlo Statistical Methods”.
- [4] FISHMAN, G.S, “Monte Carlo, Concepts, Algorithms and Applications”, Springer Verlag.
- [5] FLEMING W.H., AND H.M. SONER, ”Controlled Markov Processes and Viscosity Solutions”, Springer Verlag, 1992.
- [6] FOURNIÉ E., J.M. LASRY, AND P.L. LIONS, ”Some Nonlinear Methods for Studying Far-from-the-money Contingent Claims”, *Numerical Methods in Finance*, L.C.G. Rogers et D.Talay, eds., Cambridge University Press, 115–145.
- [7] FOURNIÉ E., J.M. LASRY, AND N. TOUZI, ”Monte Carlo Methods for Stochastic Volatility Models”, *Numerical Methods in Finance*, L.C.G. Rogers et D.Talay, eds., Cambridge University Press, 146–164.
- [8] GEWEKE, J., “Antithetic acceleration of Monte Carlo integration in Bayesian inference”, *Journal of Econometrics*, 38, 73–90, 1988.
- [9] GLASSERMAN, P., P. HEIDELBERGER, AND P. SHAHABUDDIN, “Asymptotically Optimal importance Sampling and Stratification for Pricing Path-Dependent Options”, 1998.
- [10] GOBET, E., “Schémas d’Euler pour diffusions tuée. Application aux options à barrière”, *Université Paris VII, Thèse*, 1998.
- [11] HARRISON, M.J., AND D.M. KREPS, “Martingales and arbitrage in multiperiod securities markets”, *Journal of Economic Theory*, 29, 381–408, 1979.
- [12] HARRISON, M.J., AND S.R. PLISKA, “Martingales and stochastic integrals in the theory of continuous trading”, *Stochastic Processes and their Applications*, 11, 215–260, 1981.
- [13] HARRISON, M.J., AND S.R. PLISKA, “A stochastic calculus model of continuous trading : complete markets”, *Stochastic Processes and their Applications*, 15, 313–316, 1983.
- [14] HESTON, S.L., “A closed-form solution for options with stochastic volatility with application to bond and currency options”, *The Review of Financial Studies*, 1993.
- [15] HULL J., AND A. WHITE, “The pricing of options on assets with stochastic volatilities”, *Journal of Finance*, 3, 281–300, 1987.
- [16] NEWTON, N.J., “Variance Reduction for Simulated diffusions”, *SIAM J. Appl. Math.* 54, 1780–1805, 1994.

- [17] PARDOUX, E., AND D. TALAY, "Discretization and simulation of stochastic differential equations", *Acta Applicandae Mathematicae*, 3, 23–47, 1985.
- [18] ROMANO, M., AND N. TOUZI, "Contingent Claims and Market Completeness in a stochastic Volatility Model", *Mathematical Finance*, Vol. 7, No. 4 (October 1997), 399-412.
- [19] RUBINSTEIN, R.Y., "Simulation and the Monte Carlo Method", *J. Wiley, New York, 1981*.
- [20] VARADHAN S.R.S., "Large Deviations and Applications", 1984.