

Mémoire

pour obtenir les diplômes

“Master of Science & Diplôme d’ingénieur de l’ENSIIE “

Année 2020/2021

Modèle de Cheyette et pricing de produits de taux

—

Awalee Consulting

Rédigé par:

Alexis Houssard

E-Mail: alexis.houssard@hotmail.com

Date de début: 19 Avril 2021

Date de fin: 18 octobre 2021



Abstract

This document deals with a study of the so-called "Cheyette" interest rate model, a specification of the HJM framework which provides a continuous forward curve with a markovian short rate. The main idea of this work is to understand the construction, the parameterization and the capacity to reproduce market's prices of this model with a practical but also theoretical point of view. In order to highlight the properties of this forward curve characterization, one will carry out all the applications in one of the most common contexts of interest rate modeling: swaption pricing. Our aim here is to introduce briefly the theory of the generic model and to implement its uni-dimensional version with a parametric local volatility of the short rate. Afterwards, some numerical tests will be provided to find out the best way to calibrate the Cheyette model on implied volatility and to price swaption. Some mathematical arguments will follow the observations made to justify our results. To finish with, one will pay attention to an extension of the model in a multi-dimensional framework and achieve the same kind of tests as described previously. This study represents a base which can be extended with a non-parametric local or stochastic volatility function of the short rate.

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 1 |
| 2 | Le modèle de Cheyette | 3 |
| 2.1 | Rappels sur la modélisation de la courbe des taux | 3 |
| 2.2 | Le cadre HJM | 4 |
| 2.3 | La spécification de Cheyette: un modèle Markovien | 5 |
| 2.4 | Volatilité locale linéaire | 8 |
| 3 | Valorisation d'une swaption européenne | 9 |
| 3.1 | Valorisation sous la probabilité risque neutre | 10 |
| 3.2 | Valorisation sous la mesure terminale | 11 |
| 3.3 | Discretisation des EDS | 12 |
| 3.3.1 | Le schéma d'Euler | 12 |
| 3.3.2 | Le schéma log-Euler | 14 |
| 3.4 | Courbe des prix et choix de simulation | 16 |
| 3.5 | Approximation par un modèle "displaced log-normal" | 21 |
| 3.5.1 | La swaption comme Call sur taux swap | 21 |
| 3.5.2 | La projection Markovienne pour la volatilité locale | 22 |
| 3.5.3 | La formule de Pricing | 27 |
| 4 | Courbe de volatilité implicite | 30 |
| 4.1 | Volatilité implicite log-normale | 30 |
| 4.2 | Volatilité implicite normale | 37 |
| 4.3 | Observations et propriétés sur le skew | 40 |
| 5 | Calibration de volatilité locale en dimension 1 | 45 |
| 5.1 | Paramétrisation du problème | 45 |
| 5.2 | Algorithme de calibration | 47 |
| 5.3 | Tests et résultats | 49 |
| 6 | Conclusion | 54 |
| | Références | 55 |
| | Annexe | 56 |

1 Introduction

La modélisation de la courbe des taux d'intérêts représente un enjeu central dans le monde des mathématiques financières en ce sens qu'elle intervient dans la valorisation de la majorité des produits financiers. En effet, tous les produits dont le retour sur investissement (nous emploierons le terme *payoff* dans la suite et cela désignera d'un point de vue mathématique la quantité perçue par le détenteur du produit à l'instant terminal) dépend de la courbe de taux d'intérêt en un ou plusieurs instants nécessitent, pour être correctement valorisés, de se questionner sur le modèle de taux à considérer. Cette considération mène à une certaine forme de la courbe des taux forward avec ses propres caractéristiques mathématiques. L'objectif est alors de faciliter la mise en place d'une méthode de valorisation, d'un point de vue théorique comme pratique, sans altérer la capacité du modèle à reproduire les phénomènes observables sur le marché tels que le *skew* et le *smile* de volatilité implicite d'instruments de référence.

Au début des années 1990 apparaît la méthodologie de "Heath-Jarrow-Morton" (HJM) dans plusieurs articles des auteurs du même nom. Celle-ci s'imposera comme le cadre théorique le plus fréquemment employé pour modéliser l'entière de la courbe forward, de manière continue et dans le contexte d'un marché complet dépourvu de tout arbitrage. Quelques années plus tard, les travaux de Oren Cheyette (O.Cheyette 1995) notamment ont conduit à une spécification du cadre HJM, conférant au taux court un caractère aléatoire markovien intéressant.

Ce mémoire fait ainsi l'objet de l'étude d'un modèle de Cheyette traité dans un premier temps dans sa version unidimensionnelle et dans un second temps en multi-dimension. Afin de tester ce modèle et d'en comprendre les fondements, nous concentrerons notre étude sur sa capacité à valoriser les swaptions européennes du fait de leur importante liquidité sur le marché. De plus, les payoffs de ces options font intervenir plusieurs instruments avec des maturités différentes, ce qui leur confère une utilité particulière dans le processus de calibration d'un modèle dont certains paramètres dépendent du temps.

Les objectifs du projet développé à travers ce document résident dans trois étapes principales que sont la valorisation des swaptions européennes dans notre cas particulier, le calcul de la courbe de volatilité implicite ainsi que la calibration de la volatilité locale du modèle de Cheyette considéré. Ces trois étapes sont intrinsèquement liées dans le processus permettant de 'pricer' en temps réels des instruments non disponibles sur le marché tout en étant cohérent avec la tendance de ce dernier. En d'autres termes, il s'agira de calibrer la volatilité locale paramétrique du modèle de Cheyette sur le cube de volatilité des swaptions du marché afin de reproduire au mieux la courbe de volatilité implicite et les prix observés.

Enfin, nous invitons le lecteur à explorer différentes extensions du modèle de Cheyette telles que celle à volatilité non-paramétrique ou celle à volatilité stochastique.

Notations

- Probabilité historique \mathbb{P} / probabilité risque neutre \mathbb{Q}
- Ω l'univers dans la probabilité risque neutre. Nous noterons souvent w pour désigner un évènement de l'univers.
- Φ est la fonction de distribution de la loi normale centrée réduite et ϕ sa densité de probabilité.
- $\mathbb{E}_{\mathbb{Q}}$: espérance sous la probabilité \mathbb{Q}
- $W^{\mathbb{Q}}$: un mouvement Brownien sous la mesure \mathbb{Q}
- \mathcal{F} : Filtration propre du mouvement Brownien $W^{\mathbb{Q}}$
- $\langle X, Y \rangle_t$: covariation quadratique des processus X, Y évaluée en t .
- Facteur d'actualisation à l'instant t : $B(t) = e^{-\int_0^t r(s) ds}$
- Le processus actualisé $X(t)/B(t)$ sera noté $\tilde{B}(t)$
- Partie positive d'un réel: $x^+ = \max(x, 0)$
- (\mathcal{H}) Hypothèse de coefficients de diffusions Lipschitz et à croissance au plus linéaire: On considère un drift $\mu(t, x)$ et une volatilité $\sigma(t, x)$. Ils vérifient (\mathcal{H}) si

$$\forall t \geq 0, \forall (x, y) \in \mathbb{R}^d \times \mathbb{R}^d, \exists K > 0, \begin{cases} \|\mu(t, x) - \mu(t, y)\| + \|\sigma(t, x) - \sigma(t, y)\| \leq K\|x - y\| \\ \|\mu(t, x)\|^2 + \|\sigma(t, x)\|^2 \leq K^2(1 + \|x\|^2) \end{cases}$$

- Le prix du zéro coupon $P(t, T) \stackrel{\text{def}}{=} B(t)\mathbb{E}_{\mathbb{Q}}[B(T)^{-1}|\mathcal{F}_t]$

2 Le modèle de Cheyette

2.1 Rappels sur la modélisation de la courbe des taux

Commençons par introduire les principes fondamentaux de la courbe des taux avec la définition du **taux forward**. Le taux forward est le taux qui annule la valeur initiale d'un contrat du même nom, le *Forward Rate Agreement*. Il s'agit d'un contrat entre deux contreparties assurant un prêt (ou un emprunt selon la contrepartie) à un taux connu en l'instant t où le contrat prend effet, entre deux dates futures T et S .

Ce contrat peut être répliqué à l'aide des obligations zéros coupons si on fait l'hypothèse que leurs valeurs sont connues en l'instant initial t . La stratégie de réplication en considérant un taux en convention exponentielle est la suivante:

On note N le nominal du contrat, t la date à laquelle le contrat prend effet et $T < S$ les dates entre lesquelles le taux est calculé.

- En t : La contrepartie prêteuse paie $NP(t, T) - Ne^{(S-T)F(t, T, S)}P(t, S)$
- T : Le prêt est effectif en T , on livre $-N$
- $S > T$: La contrepartie prêteuse récupère les intérêts $+Ne^{(S-T)F(t, T, S)}$

En l'absence d'arbitrage, deux stratégies fournissant les mêmes *cash-flows* ont la même valeur initiale. On en déduit que la valeur du *Forward Rate Agreement* en t est de $NP(t, T) - Ne^{(S-T)F(t, T, S)}P(t, S)$.

Le taux forward est le taux qui annule cette quantité, on peut donc exprimer le taux forward $F(t, T, S)$:

$$F(t, T, S) = \frac{1}{S - T} \ln \frac{P(t, T)}{P(t, S)}$$

Le taux forward instantané peut être défini comme le taux forward pour lequel la durée entre les dates T et S est infiniment courte, à l'instar d'un taux *overnight* pour citer un équivalent sur le marché des taux. De manière quantitative, ce taux f vaut

$$f(t, T) = \lim_{S \rightarrow T} F(t, T, S) = -\frac{\partial \ln P(t, T)}{\partial T}$$

La courbe forward $T \rightarrow f(t, T)$ permet donc directement de déduire une expression du prix de l'obligation zéro coupon pour toutes les maturités puisque

$$P(t, T) = e^{-\int_t^T f(t, u) du}$$

Enfin, le **taux court** en l'instant t est défini de la manière suivante

$$r(t) = f(t, t).$$

Il ne correspond à aucun actif échangeable sur le marché, mais c'est un taux calculable en chaque instant et qui intervient dans la valorisation des instruments financiers sous la probabilité risque neutre.

Modélisation de la structure par terme

On peut distinguer deux classes de modèles de taux d'intérêts, les modèles appelés "**Short rate models**" et les modèles dits "**Term structure models**" (respectivement modèles de taux court et modèles à structure par terme en français). Les short rate models ont pour objectif de décrire la dynamique du taux court, la plupart du temps sous la forme d'une diffusion satisfaisant une équation différentielle stochastique générique telle que: $dr(t) = (\theta(t) - \alpha(t)r(t)) dt + \sigma(t)r(t)^\beta dW_t$. Certes ces modèles permettent une reconstruction de la courbe zéro coupon grâce à la relation $P(t, T) = \mathbb{E}_{\mathbb{Q}}[e^{-\int_t^T r_s ds} | \mathcal{F}_t]$, mais introduire un modèle de ce type qui reproduit une fonction de volatilité réaliste pour les taux forwards s'avère complexe.

La modélisation de la structure par terme, quant à elle, propose de décrire directement la courbe forward et se voit beaucoup plus avantageuse d'un point de vue mathématique. En d'autres termes, connaissant la courbe forward initiale $T \rightarrow P(0, T)$, on cherche à modéliser la courbe forward $t \rightarrow f(t, T)$ de manière continue pour tous les horizons T . On construira aussi la fonction P de sorte qu'elle soit dérivable par rapport à T . Il s'agit en fait de décrire la dynamique du taux forward $f(t, T)$ sur le domaine désiré. Il paraît en effet plus logique de déduire un taux court instantané de la courbe forward qui, elle, a une dépendance à la maturité, plutôt que de procéder à la démarche inverse.

2.2 Le cadre HJM

La majeure partie des modèles de structure par terme est issue de ce que l'on appelle le cadre HJM que nous présenterons succinctement ici.

On se propose d'écrire la dynamique du taux forward sous la forme suivante:

$$df(t, T) = \mu_f(t, T) dt + \sigma_f(t, T)^\top dW_t, \quad t \leq T$$

où W est un mouvement Brownien sous la probabilité historique \mathbb{P} .

Pour assurer l'absence d'arbitrage dans ce cadre, il est suffisant d'imposer l'existence d'une probabilité risque neutre sous laquelle le prix du zéro coupon actualisé est une martingale. Mathématiquement parlant, cela équivaut à l'existence de σ_P telle que

$$\frac{dP(t, T)}{P(t, T)} = r(t) dt - \sigma_P(t, T)^\top dW_t^{\mathbb{Q}}.$$

On rappelle l'expression liant le prix du zéro coupon au taux forward $f(t, T) = -\frac{\partial \ln P(t, T)}{\partial T}$.

Pour des raisons évidentes de non-arbitrage, le prix du zéro coupon est borné inférieurement par 0. Puisque la fonction $x \rightarrow \ln(x) \in \mathcal{C}^2(\mathbb{R}^{+*}, \mathbb{R})$, à l'aide des hypothèses relaxées présentées en Annexe 1, on peut appliquer le lemme d'Itô au processus $(P(t, T))_t$.

$$d \ln (P(t, T)) = \left(r(t) - \frac{1}{2} \sigma_P(t, T)^\top \sigma_P(t, T) \right) dt - \sigma_P(t, T)^\top dW_t^\mathbb{Q}.$$

Par le raisonnement détaillé en Annexe 2, on aboutit à la dynamique du taux forward suivante:

$$\begin{aligned} df(t, T) &= d \left(-\frac{\partial \ln P(t, T)}{\partial T} \right) \\ &= -\frac{\partial}{\partial T} d \ln (P(t, T)) \\ \iff df(t, T) &= \frac{\partial \sigma_P}{\partial T}(t, T)^\top \sigma_P(t, T) dt + \frac{\partial \sigma_P}{\partial T}(t, T)^\top dW_t^\mathbb{Q} \end{aligned}$$

Si on considère un changement de probabilité tel que l'on ait $W_t^\mathbb{Q} = W_t + \int_0^t \lambda_u du$ (λ est la prime de risque définie telle que le ZC actualisé soit une martingale sous \mathbb{Q}) alors on en déduit par identification

$$\sigma_f(t, T) = \frac{\partial \sigma_P}{\partial T}(t, T),$$

ce qui conduit à

$$\sigma_P(t, T) = \int_t^T \sigma_f(t, u) du$$

avec la condition initiale $\sigma_P(t, t) = 0$.

De même, il vient directement l'expression sur les drift, connue sous le nom de **HJM drift condition**:

$$\sigma_f(t, T)^\top \lambda_t = \mu_f(t, T) - \sigma_f(t, T)^\top \int_t^T \sigma_f(t, u) du$$

.

Il est donc convenable de réécrire la dynamique du taux forward sous la forme

$$\boxed{df(t, T) = \sigma_f(t, T)^\top \left(\int_t^T \sigma_f(t, u) du \right) dt + \sigma_f(t, T)^\top dW_t^\mathbb{Q}} \quad (1)$$

Enfin,

$$r(t) = f(t, t) = f(0, t) + \int_0^t \sigma_f(u, t)^\top \left(\int_u^t \sigma_f(u, s) ds \right) du + \int_0^t \sigma_f(u, t)^\top dW_u^\mathbb{Q} \quad (2)$$

2.3 La spécification de Cheyette: un modèle Markovien

Comme il est explicité dans un des papiers originaux d'O. Cheyette sur sa reformulation du modèle HJM, ce dernier présente de nombreux avantages théoriques, toutefois il impose

certaines contraintes sur l'aspect pratique de la modélisation des taux. En effet, si le modèle HJM est si générique, c'est grâce à sa capacité à générer une dynamique de taux forward à n'importe quel instant et pour une maturité quelconque à condition de connaître la structure de volatilité forward en 0. Cheyette évoque un espace des dynamiques de dimension infinie. Il n'en reste pas moins que cette généralité impose dans certains cas, par exemple pour des instruments financiers sur de longues maturités avec des échéances régulières, de simuler une dynamique pour le taux forward à chaque échéance. Cette nécessité exclut donc toute possibilité d'employer des méthodes de discrétisation et d'estimation Monte Carlo pour des raisons de complexité. La simple valorisation d'une courbe zero coupon implique encore une fois de générer une infinité de taux forward. De même, on remarque que la modélisation du taux forward précédemment induite (1) est "*path dependent*", ce qui peut compliquer les simulations.

Pour pallier ces problèmes, O. Cheyette propose un remaniement du cadre HJM mettant en exergue le caractère markovien du taux court. De manière générale, travailler avec un taux court qui s'exprime comme une fonction déterministe d'un processus de Markov permet de faciliter le calcul d'espérance conditionnelle et de contourner les problèmes liés à la "path dependence".

La spécification de Cheyette réside dans l'utilisation d'une **structure de volatilité forward à variables séparables**. C'est-à-dire,

$$\boxed{\begin{aligned} \sigma_f(t, T, \omega) &= g(t, \omega) h(T) & g : \mathbb{R}^+ \times \Omega &\rightarrow \mathbb{R} \\ & & h : \mathbb{R}^+ &\rightarrow \mathbb{R}^+ \end{aligned}} \quad (3)$$

On étudie ici la version mono-facteur et unidimensionnelle et on omettra l'argument ω dans les longues expressions. Dans le cas multi-facteurs, si on considère un mouvement Brownien en dimension d alors $g : \mathbb{R}^+ \times \Omega \rightarrow \mathbb{R}^d$.

L'équation (2) devient alors

$$r(t) = f(0, t) + h(t) \int_0^t g(u)^\top g(u) \left(\int_u^t h(s) ds \right) du + h(t) \int_0^t g(u)^\top dW_u^\mathbb{Q}$$

Le taux court peut alors être exprimé (Voir section 4.5.2 de Andersen 2010a) comme solution d'une équation différentielle stochastique de la forme

$$dr(t) = \left[a(t) - \chi(t)r(t) \right] dt + h(t)g(t)^\top dW_t^\mathbb{Q}$$

où

$$\begin{cases} a(t) = \frac{\partial f(0, t)}{\partial t} + \chi(t)f(0, t) + \int_0^t e^{-2 \int_u^t \chi(s) ds} \sigma_r(u)^\top \sigma_r(u) du & , \sigma_r(t) = \sigma_f(t, t) = g(t)h(t) \\ \chi(t) = -\frac{h'(t)}{h(t)} \end{cases}$$

Au vu de la dynamique de r , le taux court est bien un processus de Markov puisque le drift et la volatilité satisfont les conditions du théorème d'existence et d'unicité d'une solution markovienne.

Remarque: Ce type de modèle est dit **Quasi-Gaussien à un facteur**. Quasi Gaussien en raison du caractère stochastique de g . Si cette fonction avait été déterministe, le taux court serait Gaussien car somme d'un terme déterministe et de l'intégrale stochastique d'une fonction déterministe. Le modèle est dit mono-facteur car il n'admet qu'une seule source d'aléa W .

Toute la puissance de cette spécification réside dans la capacité à développer les calculs pour obtenir une forme simplifiée du prix de l'obligation zero coupon que l'on pourra contrôler avec les paramètres du modèle. Nous nous proposons ici de retranscrire le raisonnement développé dans (Andersen 2010b, Chap.13) afin d'obtenir ce résultat clé.

On a donc, avec ce qui précède, $h(t) = e^{-\int_0^t \chi(s) ds}$.

D'après (2),

$$f(t, T) = f(0, T) + h(T) \int_0^t g(u)^2 \int_u^T h(s) ds du + h(T) \int_0^t g(u) dW_u^{\mathbb{Q}}. \quad (4)$$

Si on pose

$$x(t) = h(t) \int_0^t g(u)^2 \int_u^t h(s) ds du + h(t) \int_0^t g(u) dW_u^{\mathbb{Q}},$$

on remarque d'abord que $f(t, t) = f(0, t) + x(t) = r(t)$.

Ensuite, il s'agit de différencier l'expression de x pour en déduire une dynamique. En appliquant la formule de Leibniz, il vient:

$$\begin{aligned} dx(t) &= \frac{h'(t)}{h(t)} x(t) dt + y(t) dt + \sigma_r(t) dW_t^{\mathbb{Q}} \\ &= (y(t) - \chi(t)x(t)) dt + \sigma_r(t) dW_t^{\mathbb{Q}} \end{aligned}$$

où on a posé $y(t) = h(t)^2 \int_0^t g(u)^2 du$.

De plus, on va exprimer le taux forward en fonction de notre nouvelle variable x par une transformation sur (4).

$$\begin{aligned} f(t, T) &= f(0, T) + h(T) \left[\int_0^t g(u)^2 \left(\int_u^t h(s) ds + \int_t^T h(s) ds \right) du + \frac{h(t)}{h(T)} \int_0^t g(u) dW_u^{\mathbb{Q}} \right] \\ &= f(0, T) + \frac{h(T)}{h(t)} x(t) + h(T) \int_0^t g(u)^2 du \int_t^T h(s) ds \\ &= f(0, T) + \frac{h(T)}{h(t)} \left[x(t) + y(t)G(t, T) \right], \quad G(t, T) = \frac{1}{h(t)} \int_t^T h(s) ds \end{aligned}$$

A l'aide de ce qui précède, on obtient facilement:

$$\begin{aligned} P(t, T) &= e^{-\int_t^T f(t, u) du} \\ &= e^{-\int_t^T f(0, u) du - \int_t^T \frac{h(u)}{h(t)} (x(t) + y(t)G(t, u)) du} \\ &= P(0, T) e^{-G(t, T)x(t) - \frac{y(t)}{h(t)} \int_t^T h(u)G(t, u) du} \end{aligned}$$

Or par interversion d'intégrales,

$$\begin{aligned}
\frac{1}{h(t)} \int_t^T h(u) G(t, u) \, du &= \frac{1}{h(t)^2} \int_t^T h(u) \int_t^u h(s) \, ds \, du \\
&= \frac{1}{h(t)^2} \int_t^T h(s) \int_s^T h(u) \, du \, ds \\
&= G(t, T)^2 - \frac{1}{h(t)^2} \int_t^T h(s) \int_t^s h(u) \, du \, ds \\
\iff &\frac{1}{h(t)^2} \int_t^T h(s) \int_t^s h(u) \, du \, ds = \frac{1}{2} G(t, T)^2
\end{aligned}$$

Finalement, on aboutit à l'important résultat qui suit:

Proposition 1

$$\begin{aligned}
&\forall 0 \leq t \leq T, \quad r(t) = f(0, t) + x(t) \\
&\text{Et} \quad P(t, T, x(t), y(t)) = \frac{P(0, T)}{P(0, t)} e^{-G(t, T) x(t) - \frac{1}{2} G(t, T)^2 y(t)} \\
&\text{Où} \quad \begin{cases} dx(t) = (y(t) - \chi(t)x(t)) \, dt + \sigma_r(t, \omega) \, dW_t^{\mathbb{Q}}, \\ dy(t) = (\sigma_r(t, \omega)^2 - 2\chi(t)y(t)) \, dt, \\ x(0) = y(0) = 0 \end{cases}
\end{aligned}$$

Cette proposition nous permet de conclure que, dans le modèle de Cheyette, la courbe forward et à plus forte raison la courbe zéro coupon sont entièrement caractérisées par les deux variables d'état x et y . Ces dernières seront au coeur de notre étude puisqu'elles contrôlent le prix des swaptions à travers celui des zéro coupon, nous nous intéresseront plus particulièrement à l'effet de la structure de la volatilité σ_r sur les prix et la volatilité implicite des swaptions européens.

2.4 Volatilité locale linéaire

Le modèle décrit dans la section précédente est valable pour tous les types de volatilité locale. Nous nous focaliserons ici à l'étude du cas où la volatilité locale est linéaire. En d'autres termes, on choisit une structure de volatilité sous la forme d'une fonction linéaire des variables d'état x et y avec des paramètres dépendant du temps.

On note ainsi

$$\sigma_r(t, x, y) = \lambda(t)(a(t) + b(t) x) \quad (5)$$

Deux raisons justifient en grande partie ce choix de volatilité (voir également Chap.13 Andersen 2010b). D'une part, cette structure laisse paraître de manière quasi transparente l'effet des paramètres et des variables d'état sur le comportement de la volatilité. De plus, la calibration sera facilitée puisqu'elle ne portera que sur deux fonctions, à savoir λa et λb . D'autre part, nous verrons que cette forme simplifie les équations et confère une liberté qui

conduira à un certain nombre d'approximations utiles. On notera que y joue le rôle d'une variable d'ajustement, c'est pourquoi il ne semble pas aberrant de considérer la volatilité comme fonction de x uniquement.

Remarque: Les auteurs affirment dans Andersen 2010b, Section 13.1.2 que (x, y) définit clairement un processus Markovien en deux dimensions. Cependant, avec le choix d'une volatilité affine, le drift de y dépend de x^2 ce qui rend les conditions de Lipschitz et de croissance linéaire discutables. On admettra ce caractère Markovien sur une région où x est localement borné ou artificiellement borné dans les simulations. Cela garantit que le drift et la volatilité sont localement Lipschitz et que la croissance linéaire est satisfaite sur un compact de l'espace d'arrivée du processus. On aura donc existence d'un processus Markovien jusqu'au temps de sortie de ce compact (Voir E.Chevalier.et.al. 2015).

3 Valorisation d'une swaption européenne

On se propose de rappeler brièvement la définition d'une swaption européenne. Il s'agit d'une option conférant le droit et non l'obligation d'entrer dans un swap européen à la maturité (du swap qui coïncide ici avec celle de la swaption) que l'on notera T_0 . Le swap est un contrat entre deux contreparties qui assure l'échange d'un taux fixe, le *strike* noté K , contre un taux variable aux différentes échéances $T_1 < T_2 < \dots < T_N$. Nous considérerons le taux forward composé entre T_i et T_{i+1} en tant que taux variable échangé en T_{i+1} , $i = 0, \dots, N-1$. Mathématiquement, on peut écrire le payoff du swap de la manière suivante (avec $\tau_i = T_{i+1} - T_i$):

$$\begin{aligned} Swap(T_0) &= B(t) \mathbb{E}_{\mathbb{Q}} \left[\sum_{i=0}^{N-1} B(T_{i+1})^{-1} \tau_i (F(T_i, T_i, T_{i+1}) - K) | \mathcal{F}_t \right] \\ &= B(t) \sum_{i=0}^{N-1} \mathbb{E}_{\mathbb{Q}} \left[B(T_{i+1})^{-1} \left(\frac{1}{P(T_i, T_{i+1})} - 1 \right) | \mathcal{F}_t \right] - K \sum_{i=0}^{N-1} \tau_i P(t, T_{i+1}) \end{aligned}$$

Or d'après la propriété de la tour et le fait que les zero coupons actualisés soient des martingales sous \mathbb{Q} en tant qu'actifs échangeables, il vient:

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}} \left[B(T_{i+1})^{-1} \left(\frac{1}{P(T_i, T_{i+1})} - 1 \right) | \mathcal{F}_t \right] &= \mathbb{E}_{\mathbb{Q}} \left[\left(\frac{1}{P(T_i, T_{i+1})} - 1 \right) \mathbb{E}_{\mathbb{Q}} \left[B(T_{i+1})^{-1} | \mathcal{F}_{T_i} \right] | \mathcal{F}_t \right] \\ &= \mathbb{E}_{\mathbb{Q}} \left[\left(\frac{1}{P(T_i, T_{i+1})} - 1 \right) B(T_i)^{-1} P(T_i, T_{i+1}) | \mathcal{F}_t \right] \\ &= B(t)^{-1} \left(P(t, T_i) - P(t, T_{i+1}) \right) \end{aligned}$$

Par sommation, on a:

$$Swap(T_0) = 1 - P(T_0, T_N) - K \sum_{i=0}^{N-1} \tau_i P(t, T_{i+1}) \quad (6)$$

Finalement, le **prix de la swaption**, à l'instant présent, V_0 de strike K , de maturité T_0 et de tenor structure (la maturité du swap sous-jacent ainsi que la fréquence de paiement) $T_0 < T_1 < \dots < T_N$ s'exprime de la manière suivante:

$$V_0 = \mathbb{E}_{\mathbb{Q}} \left[e^{-\int_0^{T_0} r(s) ds} Swap(T_0)^+ \right] \quad (7)$$

3.1 Valorisation sous la probabilité risque neutre

Afin de valoriser les swaptions, nous allons exploiter leurs expressions sous la forme d'espérance afin de les approximer grâce à une méthode de Monte Carlo très classique.

Reprenons la formule du prix de la swaption européenne 7 adaptée à notre modèle, cela donne:

$$\begin{aligned} V_0 &= \mathbb{E}_{\mathbb{Q}} \left[e^{-\int_0^{T_0} r(s) ds} Swap(T_0, x(T_0), y(T_0))^+ \right] \\ &= P(0, T_0) \mathbb{E}_{\mathbb{Q}} \left[e^{-\int_0^{T_0} x(s) ds} \left(1 - P(T_0, T_N, x(T_0), y(T_0)) - K \sum_{n=0}^{N-1} \tau_n P(T_0, T_{n+1}, x(T_0), y(T_0)) \right)^+ \right] \end{aligned} \quad (8)$$

Où on rappelle que

$$\forall n \in [0, N], \quad P(T_0, T_n, x, y) = \frac{P(0, T_n)}{P(0, T_0)} e^{-G(T_0, T_n)x - \frac{1}{2}G(T_0, T_n)^2 y}$$

Nous allons ainsi calculer un estimateur Monte Carlo de l'espérance \hat{V}_0 à partir de M simulations indépendantes (dans la pratique, les simulations sont pseudos indépendantes) $(S_m)_{m=1, M}$ tel que:

$$\begin{cases} \hat{V}_0 = \frac{1}{M} \sum_{m=1}^M S_m \\ \forall m, S_m = e^{I_m} \left(P(0, T_0) - P(0, T_0)P(T_0, T_N, x_m, y_m) - K P(0, T_0) A(T_0, x_m, y_m) \right)^+ \end{cases}$$

Où on note x_m, y_m, I_m la m-ième simulation des variables aléatoire $x(T_0), y(T_0), I(T_0)$ grâce à un schéma d'Euler sur lequel nous donnerons des précisions dans la section 3.3.

Nous allons également donner une notation à la somme qui subsiste dans le payoff de (8), que l'on nomme souvent "*Annuity*".

$$A(t, x, y) = \sum_{n=0}^{N-1} \tau_n \frac{P(0, T_{n+1})}{P(0, t)} e^{-G(t, T_{n+1})x - \frac{1}{2}G(t, T_{n+1})^2 y}$$

Remarque: La forme de volatilité locale (5) est générique car elle autorise des fonctions λ, a, b déterministes quelconques. Toutefois, en gardant à l'esprit l'objectif de calibration de la volatilité qui se fera avec des paramètres constants par morceaux, nous simulerons les prix de swaption dans les tests avec des paramètres constants.

3.2 Valorisation sous la mesure terminale

Dans cette section, notre but est de se débarrasser du terme "path dependent" $e^{-\int_0^{T_0} x(s) ds}$ dans l'espérance en introduisant une mesure de probabilité appropriée. Une solution classique est de choisir le prix zero coupon comme numéraire. Mais il est préférable d'opter pour un numéraire qui reste "en vie" pour toutes les dates de la structure de tenor. On se porte ainsi vers le numéraire $P(t, T_N)$ et la mesure associée est appelée **la mesure terminale**.

$(e^{-\int_0^t r(s) ds} P(t, T_N))_t$ est une martingale strictement positive sous la mesure risque neutre (à condition que $\sigma_P(t, T)$ vérifie la condition de Novikov pour tout T), on pose donc

$$\frac{d\mathbb{Q}^{T_N}}{d\mathbb{Q}} \Big|_{\mathcal{F}_t} = \frac{P(t, T_N)}{P(0, T_N)} e^{-\int_0^t r(s) ds}$$

Cela implique

$$\begin{aligned} V_0 &= P(0, T_N) \mathbb{E}_{\mathbb{Q}^{T_N}} \left[\frac{1}{P(T_0, T_N)} (Swap(T_0))^+ \right] \\ &= P(0, T_N) \mathbb{E}_{\mathbb{Q}^{T_N}} \left[\frac{1}{P(T_0, T_N)} \left(1 - P(T_0, T_N, x(T_0), y(T_0)) - KA(T_0, x(T_0), y(T_0)) \right)^+ \right] \end{aligned}$$

Focalisons nous à présent sur la dynamique de x et y sous la nouvelle mesure. Pour commencer, en l'absence d'arbitrage, l'obligation actualisée de maturité T_N est une martingales sous la mesure risque neutre, donc:

$$\begin{aligned} d\left(\frac{P(t, T_N)}{B(t)}\right) &= -\frac{P(t, T_N)}{B(t)} \sigma_P(t, T_N)^\top dW_t^\mathbb{Q}. \\ \implies \frac{P(t, T_N)}{B(t)} &= P(0, T_N) e^{-\int_0^t \sigma_P(s, T_N)^\top dW_s^\mathbb{Q} - \frac{1}{2} \int_0^t \sigma_P(s, T_N)^\top \sigma_P(s, T_N) ds} \end{aligned}$$

D'après le théorème de Girsanov,

$$\boxed{W_t^{T_N} = W_t^\mathbb{Q} + \int_0^t \sigma_P(s, T_N) ds} \quad \text{est un MB sous } \mathbb{Q}^{T_N}$$

Rappelons les hypothèses du modèle, $\sigma_f(t, T) = g(t)h(T)$ et $\sigma_P(t, T) = \int_t^T \sigma_f(t, u) du = g(t) \int_t^T h(u) du = \sigma_r(t)G(t, T)$

Ainsi

$$\begin{aligned} dx(t) &= (y(t) - \chi(t)x(t)) dt + \sigma_r(t) (dW_t^{T_N} - \sigma_P(t, T_N)dt) \\ &= (y(t) - \sigma_r(t)^2 G(t, T_N) - \chi(t)x(t)) dt + \sigma_r(t) dW_t^{T_N} \end{aligned}$$

On observe dans la pratique une variance plus élevée sous cette mesure car le payoff de la swaption semble s'annuler plus fréquemment. Cela pourrait provenir du drift de x auquel on a amputé un terme positif. On apporte une preuve partielle dans l'annexe 7.

3.3 Discrétisation des EDS

3.3.1 Le schéma d'Euler

On se propose à présent de réfléchir à la méthode de simulation des processus stochastiques intervenant dans le payoff de la swaption, à savoir x et y , mais également l'intégrale du taux court $I(T_0) = -\int_0^{T_0} x(s) ds$.

Pour cela, il convient de mettre en place un schéma de discrétisation. La technique classique consiste en un schéma d'Euler mais nous verrons qu'il existe une variante intéressante. Notre but est donc de discrétiser le système suivant:

$$\begin{cases} dx(t) &= (y(t) - \chi(t)x(t)) dt + \sigma_r(t) dW_t \\ dy(t) &= (\sigma_r(t)^2 - 2\chi(t)y(t)) dt \\ dI(t) &= -x(t) dt \end{cases}$$

Mean reversion constante et volatilité locale linéaire: Premièrement, posons pour la simplicité d'interprétation

$$\begin{aligned} \chi(t) &= \chi \in \mathbb{R} \\ \text{et } \sigma_r(t) &= \sigma_r(t, x(t), y(t)) = \lambda(a + bx(t)) \end{aligned}$$

On définit une grille de discrétisation $0 = t_0 < t_1 < \dots < t_N = T_0$ et $\forall i \in [0, N-1]$, $\Delta_i = t_{i+1} - t_i$.

Nous allons simuler $x(T_0)$ et $y(T_0)$ grâce au schéma:

$$\begin{cases} x_{t_{i+1}} &= x_{t_i} + (y_{t_i} - \chi x_{t_i}) \Delta_i + \lambda(a + bx_{t_i}) \sqrt{\Delta_i} Z_i \\ y_{t_{i+1}} &= y_{t_i} + (\lambda^2(a + bx_{t_i})^2 - 2\chi y_{t_i}) \Delta_i \\ I_{t_{i+1}} &= I_{t_i} - x_{t_i} \Delta_i \end{cases}, Z_i \sim \mathcal{N}(0, 1)$$

Remarque: Ci-dessus est décrit le schéma d'Euler dans sa version explicite, il est possible d'implémenter un schéma d'Euler implicite pour éviter les instabilités. Voyons comment s'écrit le schéma implicite:

$$\begin{aligned} &\begin{cases} x_{t_{i+1}} &= x_{t_i} + (y_{t_{i+1}} - \chi x_{t_{i+1}}) \Delta_i + \lambda(a + bx_{t_i}) \sqrt{\Delta_i} Z_i \\ y_{t_{i+1}} &= y_{t_i} + (\lambda^2(a + bx_{t_{i+1}})^2 - 2\chi y_{t_{i+1}}) \Delta_i \end{cases}, Z_i \sim \mathcal{N}(0, 1) \\ &\iff \begin{pmatrix} x_{t_{i+1}} \\ y_{t_{i+1}} \end{pmatrix} = \begin{pmatrix} x_{t_i} \\ y_{t_i} \end{pmatrix} + \begin{pmatrix} f_1(t_{i+1}, x_{t_{i+1}}, y_{t_{i+1}}) \\ f_2(t_{i+1}, x_{t_{i+1}}, y_{t_{i+1}}) \end{pmatrix} \Delta_i + \begin{pmatrix} \lambda(a + bx_{t_i}) \\ 0 \end{pmatrix} \sqrt{\Delta_i} Z_i \\ &\iff \begin{pmatrix} x_{t_{i+1}} - f_1(t_{i+1}, x_{t_{i+1}}, y_{t_{i+1}}) \Delta_i \\ y_{t_{i+1}} - f_2(t_{i+1}, x_{t_{i+1}}, y_{t_{i+1}}) \Delta_i \end{pmatrix} = \begin{pmatrix} x_{t_i} \\ y_{t_i} \end{pmatrix} + \begin{pmatrix} \lambda(a + bx_{t_i}) \\ 0 \end{pmatrix} \sqrt{\Delta_i} Z_i \\ &\quad \text{où } f_1(t, x, y) = y - \chi x \quad \text{et} \quad f_2(t, x, y) = \lambda^2(a + bx)^2 - 2\chi y \end{aligned}$$

On est donc contraint de résoudre à chaque étape de l'algorithme une inversion de fonction ou une recherche de racine. Ce qui augmenterait le temps d'exécution d'une part et les imprécisions d'autre part.

Le schéma d'Euler classique implique des incréments gaussiens, c'est pourquoi la diffusion discrétisée de x peut prendre des valeurs dans $] -\infty, +\infty[$. La volatilité de x n'est donc pas nécessairement positive. Nous pouvons nous demander ce que cela donnerait avec une volatilité de la forme $\sqrt{\lambda(a + bx)}$. Le processus x deviendrait un processus CIR borné inférieurement sous la condition de Feller que nous avons démontrée en annexe 8. Toutefois, on va montrer qu'il est possible de trouver un résultat similaire dans notre modèle.

On va montrer la propriété suivante:

Proposition 2

$$\boxed{\text{Si on a } \begin{cases} \bullet & x(0) + \frac{a(0)}{b(0)} \geq 0 \\ \bullet & t \rightarrow \frac{a(t)}{b(t)} e^{\int_0^t \chi(u) du} \text{ est croissante} \end{cases} \text{ Alors } \forall t \geq 0, \quad x(t) + \frac{a(t)}{b(t)} \geq 0}$$

Preuve: On commence par exprimer y sous forme explicite. On pose $Y_t = \exp\left(2 \int_0^t \chi(u) du\right) y(t)$. D'abord, on a:

$$\begin{aligned} d(Y_t) &= 2\chi(t)Y_t dt + \exp\left(2 \int_0^t \chi(u) du\right) dy(t) + d\langle \exp\left(2 \int_0^t \chi(u) du\right), y(\cdot) \rangle_t \\ &= \exp\left(2 \int_0^t \chi(u) du\right) \sigma_r(t)^2 dt \end{aligned}$$

Par intégration et puisque $y(0) = 0$,

$$\begin{aligned} Y_t &= \int_0^t \exp\left(2 \int_0^s \chi(u) du\right) \sigma_r(s)^2 ds \\ \iff y(t) &= \int_0^t \exp\left(-2 \int_s^t \chi(u) du\right) \sigma_r(s)^2 ds \end{aligned}$$

Ainsi $\forall t \geq 0, y(t) \geq 0$.

On considère x suivant la dynamique $dx(t) = (y(t) - \chi(t)x(t)) dt + (a(t) + b(t)x(t)) dW_t^{\mathbb{Q}}$

De surcroît, on pose $X_t = \exp\left(\int_0^t \chi(u) du + \frac{1}{2} \int_0^t b(s)^2 ds - \int_0^t b(s) dW_s\right)$

D'après la formule d'Itô toujours, il vient:

$$\begin{aligned} dX_t &= X_t \left[\chi(t)dt + \frac{1}{2}b(t)^2dt - b(t)dW_t^{\mathbb{Q}} + \frac{1}{2}d\langle \int_0^t \chi(u) du + \frac{1}{2} \int_0^t b(s)^2 ds - \int_0^t b(s) dW_s \rangle_t \right] \\ &= X_t \left[\chi(t)dt + b(t)^2dt - b(t)dW_t^{\mathbb{Q}} \right] \end{aligned} \tag{9}$$

Par la formule d'intégration par parties et avec (9), on a:

$$\begin{aligned}
d\left(X_t x(t)\right) &= x(t)dX_t + X_t dx(t) + d\langle X, x(\cdot) \rangle_t \\
&= x(t)X_t \left(\chi(t)dt + b(t)^2 dt - b(t)dW_t^{\mathbb{Q}} \right) + X_t \left((y(t) - \chi(t)x(t))dt \right. \\
&\quad \left. + (a(t) + b(t)x(t)) dW_t^{\mathbb{Q}} \right) - b(t)X_t(a(t) + b(t)x(t)) dt \\
&= X_t y(t)dt - \frac{a(t)}{b(t)} X_t \left(b(t)^2 dt - b(t)dW_t^{\mathbb{Q}} \right) \\
&= X_t y(t)dt - \frac{a(t)}{b(t)} \left(dX_t - X_t \chi(t)dt \right) \\
&= X_t y(t)dt - \frac{a(t)}{b(t)} \exp \left(\int_0^t \chi(u) du \right) d \left(\exp \left(\int_0^t \chi(u) du \right) X_t \right)
\end{aligned} \tag{10}$$

Ainsi, en utilisant la formule d'intégration par parties une nouvelle fois, on obtient:

$$\begin{aligned}
X_t x(t) &= x(0) + \int_0^t X_u y(u) du - \int_0^t \frac{a(s)}{b(s)} \exp \left(\int_0^s \chi(u) du \right) d \left(\exp \left(\int_0^s \chi(u) du \right) X_s \right) \\
&= x(0) + \int_0^t X_u y(u) du + \frac{a(0)}{b(0)} - \frac{a(t)}{b(t)} X_t + \int_0^t \exp \left(\int_0^s \chi(u) du \right) X_s d \left(\frac{a(s)}{b(s)} \exp \left(\int_0^s \chi(u) du \right) \right)
\end{aligned}$$

D'où

$$X_t \left(x(t) + \frac{a(t)}{b(t)} \right) = x(0) + \frac{a(0)}{b(0)} + \int_0^t X_u y(u) du + \int_0^t \exp \left(\int_0^s \chi(u) du \right) X_s d \left(\frac{a(s)}{b(s)} \exp \left(\int_0^s \chi(u) du \right) \right)$$

Ou encore

$$x(t) + \frac{a(t)}{b(t)} = \frac{1}{X_t} \left[x(0) + \frac{a(0)}{b(0)} + \int_0^t X_u y(u) du + \int_0^t \exp \left(\int_0^s \chi(u) du \right) X_s d \left(\frac{a(s)}{b(s)} \exp \left(\int_0^s \chi(u) du \right) \right) \right]$$

On conclut trivialement avec le résultat évoqué dans la proposition 2 si les hypothèses sont vérifiées.

3.3.2 Le schéma log-Euler

Afin de produire une approximation de x qui satisfait la propriété démontrée précédemment, il est possible d'implémenter ledit schéma **log-Euler**.

La méthode consiste en une application du schéma d'Euler classique sur une transformation inversible de x . On pose $x(t) = f(u(t))$ telle que $\forall t, f(u(t)) \geq -\frac{a(t)}{b(t)}$.

Par exemple,

$$x(t) = e^{u(t)} - \frac{a(t)}{b(t)} \iff u(t) = \ln \left(x(t) + \frac{a(t)}{b(t)} \right)$$

Si

$$dx(t) = \mu(t, x(t))dt + \sigma(t, x(t)) dW_t,$$

on a d'après le lemme d'Itô :

$$\begin{aligned}
du(t) &= \frac{\left(\frac{a(t)}{b(t)}\right)'}{x(t) + \frac{a(t)}{b(t)}} dt + \frac{1}{x(t) + \frac{a(t)}{b(t)}} dx(t) - \frac{1}{2} \frac{1}{\left(x(t) + \frac{a(t)}{b(t)}\right)^2} \sigma^2(t, x(t)) dt \\
&= \left[\left(\frac{a(t)}{b(t)}\right)' + \mu(t, x(t)) - \frac{1}{2} \frac{\sigma^2(t, x(t))}{x(t) + \frac{a(t)}{b(t)}} \right] \frac{1}{x(t) + \frac{a(t)}{b(t)}} dt + \frac{\sigma(t, x(t))}{x(t) + \frac{a(t)}{b(t)}} dW_t \\
&= \alpha(t, x(t)) dt + \gamma(t, x(t)) dW_t
\end{aligned}$$

Alors

$$\begin{aligned}
\hat{u}_{i+1} &= \hat{u}_i + \alpha(t_i, \hat{x}_i) \Delta_i + \gamma(t_i, \hat{x}_i) \sqrt{\Delta_i} Z_i \\
\iff \ln \left(\hat{x}_{i+1} + \frac{a(t_{i+1})}{b(t_{i+1})} \right) &= \ln \left(\hat{x}_i + \frac{a(t_i)}{b(t_i)} \right) + \alpha(t_i, \hat{x}_i) \Delta_i + \gamma(t_i, \hat{x}_i) \sqrt{\Delta_i} Z_i
\end{aligned}$$

Finalement

$$\boxed{\hat{x}_{i+1} = -\frac{a(t_{i+1})}{b(t_{i+1})} + \hat{x}_i + \frac{a(t_i)}{b(t_i)} + e^{\alpha(t_i, \hat{x}_i) \Delta_i + \gamma(t_i, \hat{x}_i) \sqrt{\Delta_i} Z_i}}$$

De plus, on voit que si $x(0) > -\frac{a(0)}{b(0)}$, par récurrence, notre approximation est au-dessus de la borne attendue, c'est-à-dire

$$\forall i \in [0, N], \quad \hat{x}_i > -\frac{a(t_i)}{b(t_i)}$$

Cas de la volatilité locale linéaire à coefficients constants:

Dans nos hypothèses, $\mu(t, x) = y(t) - \chi x$, $\sigma(t, x) = \sigma_r(t, x) = \lambda(a + bx)$ ce qui conduit à

$$\begin{cases} \alpha(t, x(t)) = \frac{b\lambda}{\sigma_r(t, x(t))} [y(t) - \chi x(t)] - \frac{(b\lambda)^2}{2} \\ \gamma(t, x(t)) = b\lambda \end{cases}$$

Le schéma log-Euler s'écrit alors

$$\hat{u}_{i+1} = \hat{u}_i + \left(\frac{b\lambda}{\sigma_r(t_i, \hat{x}_i)} [\hat{y}_i - \chi \hat{x}_i] - \frac{(b\lambda)^2}{2} \right) \Delta_i + b\lambda \sqrt{\Delta_i} Z_i$$

Et

$$\hat{u}_0 = \ln a/b, \quad \hat{x}_i = e^{\hat{u}_i} - \frac{a}{b}$$

Remarque: Pour éviter les problèmes lorsque λ devient proche de zéro, il est préférable de calculer $\frac{b}{(a+b\hat{x}_i)}$ au lieu de $\frac{b\lambda}{\sigma_r(t_i, \hat{x}_i)}$ dans le code.

Illustrons ce propos avec des trajectoires du processus x simulées avec le schéma d'Euler d'une part et le schéma log-Euler d'autre part. On représente notamment la limite inférieure théorique et le fait que le schéma classique n'empêche pas de la franchir.

Le fait que l'on injecte \hat{u} dans une fonction exponentielle à chaque étape conduit dans des cas fréquents à l'explosion du processus. Cette solution est donc théorique et pourrait probablement être améliorée pour éviter le problème cité mais nous n'explorerons pas cette piste de recherche. On générera dans la suite les processus grâce au schéma classique, car il diverge moins rapidement, mais que l'on va borner artificiellement.

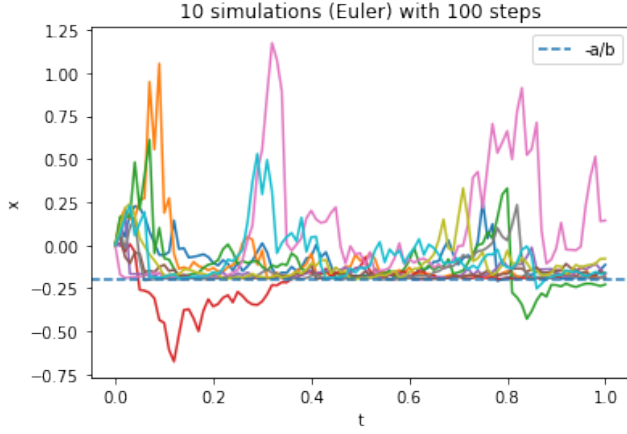


Figure 1: Schéma d'Euler

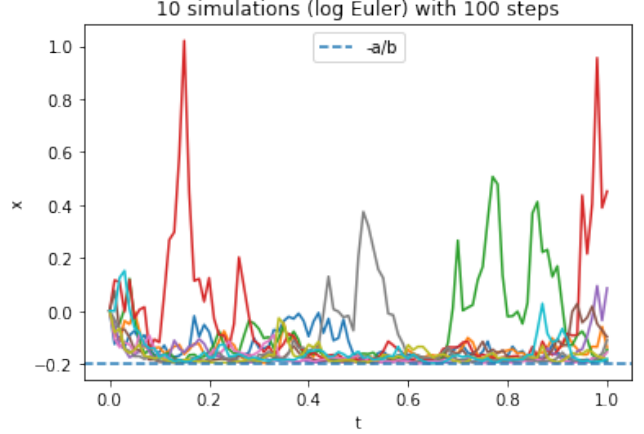


Figure 2: Schéma log-Euler

3.4 Courbe des prix et choix de simulation

Choix des paramètres Monte Carlo

Il convient de simuler $3M$ chemins nécessitant chacun N itérations, d'où une complexité en $\mathcal{O}(M * N)$.

Dans un premier temps, nous nous proposons de traiter la question de l'erreur statistique liée à l'approximation de l'espérance mathématique par l'estimateur qu'est la moyenne empirique. Autrement dit, $\mathcal{E}_{stat} = \mathbb{E}[f(\hat{x}(T_0), \hat{y}(T_0), \hat{I}(T_0))] - \frac{1}{M} \sum_{m=1}^M f(x_m, y_m, I_m)$. (On note \hat{X} l'approximation de X par le schéma discret).

On rappelle que le théorème central limite nous permet de construire un intervalle de confiance asymptotique de l'espérance recherchée.

On a

$$\lim_{M \rightarrow +\infty} \mathbb{Q}\left(\sqrt{\frac{M}{\sigma^2}} |\mathcal{E}_{stat}| \leq q\right) = 2\Phi(q) - 1 = 1 - \alpha$$

où $1 - \alpha$ est le seuil de confiance désiré.

Pour un seuil $\alpha = 0.05$, il vient $q = \Phi^{-1}(1 - \frac{\alpha}{2}) \approx 1.96$.

Ce résultat reste valable en substituant l'écart-type σ inconnu par l'écart-type empirique $\hat{\sigma}_M$ tel que $\hat{\sigma}_M^2 = \frac{1}{M} \sum_{m=1}^M (X_m - \bar{X}_M)^2$ si X est la variable aléatoire du problème (ici $f(x_m, y_m, I_m)$).

Ainsi, l'intervalle de confiance au seuil de confiance 95% s'écrit:

$$IC_\alpha = \left[\bar{X}_M - \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \sqrt{\frac{\hat{\sigma}_M^2}{M}} ; \bar{X}_M + \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \sqrt{\frac{\hat{\sigma}_M^2}{M}} \right]$$

Nous faisons le choix d'une erreur maximale au seuil de confiance 95% de l'ordre de 2% du prix de swaption, dont les valeurs varient selon la maturité, le tenor et les paramètres de

diffusion entre 10^{-2} et 10^0 dans la plupart des cas traités ici. Cela représente donc une erreur approximative de 2.10^{-4} . Enfin, nous observons généralement une variance empirique $\hat{\sigma}_M^2$ de l'ordre de 10^{-4} à 10^{-2} selon les paramètres d'entrée, nous prendrons donc 10^{-3} comme référence. Si nous calculons le nombre de simulations à réaliser afin d'atteindre cet objectif d'erreur, il vient:

$$\epsilon = 2.10^{-4} \approx 1.96 \frac{\hat{\sigma}_M}{\sqrt{M}} \implies M = \frac{1.96^2}{4} * \frac{\hat{\sigma}_M^2}{\epsilon^2} \approx 10^5$$

Dans un second temps, nous allons justifier le choix du nombre de pas de discrétisation N intervenant dans l'implémentation du schéma discret des EDS.

Pour cela, nous nous devons de citer le résultat important ci-dessous:

Soit X une diffusion et \hat{X} l'approximation discrète associée. Si les coefficients de diffusion de X sont homogènes et satisfont l'hypothèse de Lipschitz et de croissance au plus linéaire (\mathcal{H}), Alors:

$$\exists C > 0, \quad \forall p \geq 1, \quad \mathbb{E} \left[\sup_{0 \leq t \leq T_0} |X_t - \hat{X}_t|^p \right]^{\frac{1}{p}} \leq \frac{C}{N^{\frac{1}{2}}}$$

De plus, si f est K -Lipschitzienne, l'inégalité de Jensen nous permet de conclure sur la majoration de l'erreur faible:

$$\begin{aligned} |\mathcal{E}_{weak}| &= |\mathbb{E}[f(X_{T_0})] - \mathbb{E}[f(\hat{X}_{T_0})]| \leq K \mathbb{E}[|X_{T_0} - \hat{X}_{T_0}|] \\ &\leq K \mathbb{E}[|X_{T_0} - \hat{X}_{T_0}|^2]^{\frac{1}{2}} \\ &\leq \frac{C}{\sqrt{N}} \end{aligned}$$

Dans notre cas, l'objectif est d'estimer $\mathbb{E}[f(x(T_0), y(T_0), I(T_0))]$ où

$$f : (x, y, I) \rightarrow e^I \left(1 - P(T_0, T_{term}, x, y) - K \sum_{T \in \{T_1, T_{term}\}} \tau P(T_0, T, x, y) \right)^+$$

Cette fonction payoff n'est pas globalement Lipschitz mais peut être considérée localement Lipschitz sur un compact de \mathbb{R}^3 afin d'appliquer les résultats précédents.

Pour une erreur faible de l'ordre de 0.1, $N = 100$ itérations sont nécessaires et $N = 400$ sont suffisantes pour atteindre une erreur de 0.05. Nous choisirons une valeur intermédiaire de $N = 250$ qui conciliera la précision du schéma discret avec un temps de calcul raisonnable puisque l'on effectuera $100\,000 * N$ opérations lors du pricing d'une option.

Courbes de Prix

Pour illustrer notre propos nous présenterons ci-dessous certains résultats numériques sur les courbes de prix de swaptions. Premièrement, voici les courbes de Prix d'une swaption de tenor 5Y en fonction du strike (allant de 50% à 150% du strike ATM S_0) et de la maturité pour les paramètres (pris comme exemple et ne correspondant pas à une réalité de marché) $(\lambda, a, b, \chi) = (1, 1, 1, 1)$.

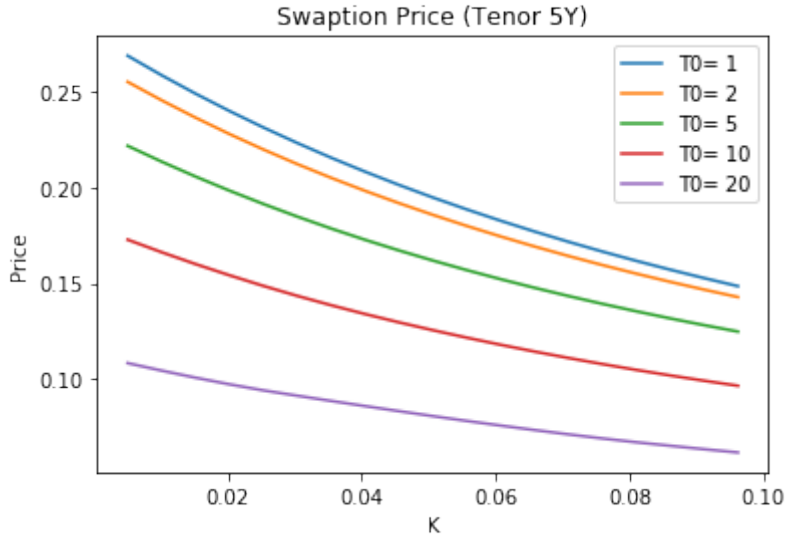


Figure 3: Courbe des prix de swaption générés

Visualisons à présent l'effet des paramètres de volatilité du taux court sur la courbe de prix de la swaption. Nous allons représenter, pour chaque paramètre λ, b et χ (l'effet de a est moins pertinent puisque l'on réduira la calibration du modèle aux seuls paramètres λ, b), le graphique du Prix de swaption à la monnaie en faisant varier le paramètre étudié. Nous ajouterons à cela les courbes de prix de la swaption en fonction du strike pour différentes valeurs du paramètres en question.

Niveau de volatilité λ

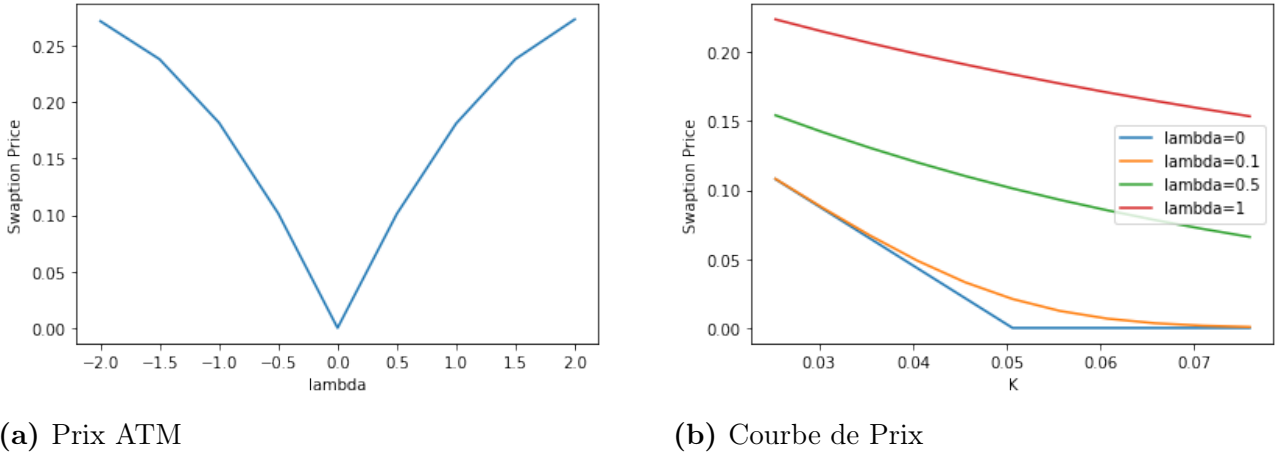


Figure 4: Variations du prix en fonction de λ

La première remarque que l'on peut faire concerne la symétrie des prix par rapport à $\lambda = 0$. Cette propriété est immédiate lorsque l'on regarde les dynamiques des facteurs x, y , en sachant que $-W \stackrel{\text{loi}}{=} W$.

Le deuxième phénomène à souligner est le fait que λ détermine le niveau de la volatilité mais également celui des prix. On remarque effectivement que la courbe de prix semble subir un décalage verticale autour de la monnaie lorsque λ varie. Les prix (Figure b) forment donc des courbes quasiment équidistantes localement au sens où la pente varie de manière négligeable devant le niveau de la courbe.

Pente de volatilité b

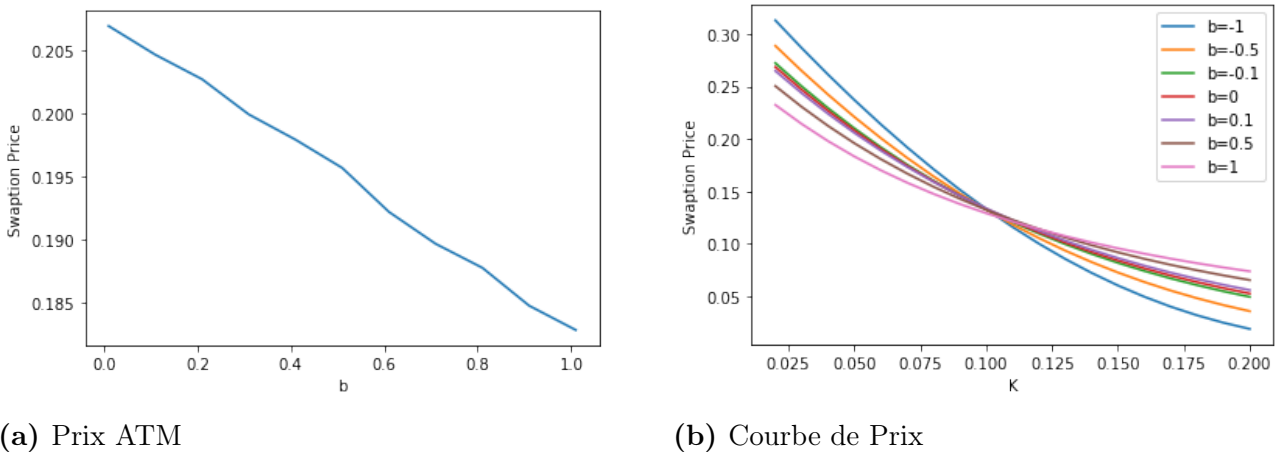
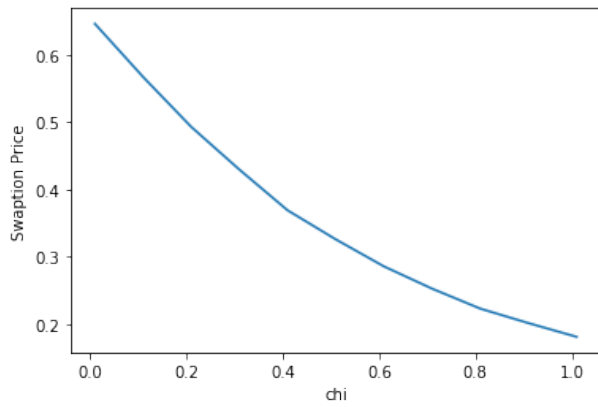


Figure 5: Variations du prix en fonction de b

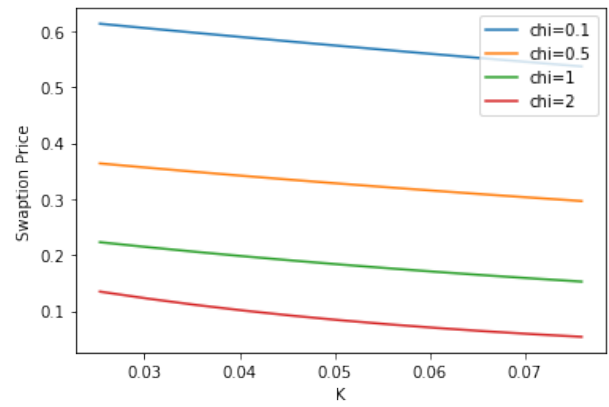
Le graphique (a) nous montre que le prix de la swaption diminue avec b autour de la monnaie. Pour mieux comprendre l'impact de b sur la courbe de prix, il convient de l'observer sur un intervalle de strike plus large. On dégage ainsi un lien direct entre b et la pente de la courbe

de prix. λ et b apparaissent donc comme complémentaires.

Mean reversion χ



(a) Prix ATM



(b) Courbe de Prix

Figure 6: Variations du prix en fonction de χ

Comme son nom l'indique, le paramètre χ tend à garder les facteurs x, y autour d'un niveau moyen et empêche leur explosion. Si ces derniers restent autour d'un niveau faible, l'annuité n'en est que plus importante et le prix de la swaption plus faible.

3.5 Approximation par un modèle "displaced log-normal"

Si nous avons déjà étudié la méthode de pricing par l'estimation de Monte Carlo, qui semble être une des méthodes les plus naturelles et réalistes de faire, il n'en est pas moins qu'elle reste contraignante d'un point de vue de la complexité numérique. En effet, nous rappelons la complexité en $\mathcal{O}(M * N)$, ce qui représente dans notre cas un total de plus de 25 millions d'opérations pour un seul tuple de paramètres. De ce fait, il s'agit d'explicitier dans cette section une méthode d'approximation du prix de la swaption développée dans Andersen 2010b[Section 13.1].

Suite à de nombreuses simplifications sur la forme complexe de la volatilité du taux swap S moyennant des techniques de projection et d'approximation d'espérances plus ou moins fines, le prix d'une swaption dans ce modèle de Cheyette peut être assimilé au prix d'un call dans un modèle simple dont l'expression mathématique est explicite. Bien que la méthode de Monte Carlo s'avère plus précise, nous mettons en exergue ici l'efficacité de calcul au détriment de la précision. Il existe bien entendu une multitude de méthodes de réduction de variance pour l'estimateur de Monte Carlo mais elles ne feront pas l'objet de notre étude.

Le fait de privilégier une formule de pricing fermée sera d'un grand intérêt pratique lors de l'étape de calibration que nous présenterons dans la section 5.

3.5.1 La swaption comme Call sur taux swap

Rappelons, pour entamer cette section légèrement plus technique, la forme du taux swap S qui annule le payoff du swap européen (payant aux dates T_0, \dots, T_N) en l'instant t :

$$S(t, x, y) = \frac{P(t, T_0, x, y) - P(t, T_N, x, y)}{\sum_{n=0}^{N-1} \tau_n P(t, T_{n+1}, x, y)} \quad (11)$$

$$= \frac{P(t, T_0, x, y) - P(t, T_N, x, y)}{A(t, x, y)} \quad (12)$$

Observons à présent qu'une simple division de son payoff par l'annuité fait de la swaption une option similaire à un Call sur le taux swap comme sous-jacent, effectivement d'après (7),

$$V_0 = \mathbb{E}_{\mathbb{Q}} \left[e^{-\int_0^{T_0} r(s) \, ds} A(T_0) (S(T_0, x(T_0), y(T_0)) - K)^+ \right]$$

Pour en faire un vrai Call, il suffit de se placer sous une mesure appropriée, ce que nous allons détailler ci-après.

3.5.2 La projection Markovienne pour la volatilité locale

On remarque avec ce qui précède que la dynamique du taux swap peut être exprimée de manière simple sous une mesure martingale associée à l'annuité A . L'annuité est définie comme la combinaison linéaire de zéro coupons, c'est donc un processus strictement positif représentant la valeur d'un actif financier. De plus, le processus actualisé \tilde{A} est évidemment une \mathbb{Q} -martingale, ce qui fait de l'annuité un numéraire adéquat.

On définit la probabilité martingale \mathbb{Q}^A associée à ce numéraire (*The annuity measure*) de la manière suivante:

$$\left. \frac{d\mathbb{Q}^A}{d\mathbb{Q}} \right|_{\mathcal{F}_t} = \frac{A(t, x(t), y(t))}{A(0)} e^{-\int_0^t r(s) ds}$$

Si on se souvient de la dynamique des zéro coupons dans notre modèle sous la forme $dP(t, T, x(t), y(t)) = P(t, T, x(t), y(t)) \left(r(t) dt - \left(\int_t^T \sigma_f(t, u) du \right) dW_t^{\mathbb{Q}} \right)$, alors on a:

$$\frac{dA(t, x(t), y(t))}{A(t, x(t), y(t))} = r(t) dt - \sum_{n=0}^{N-1} \tau_n P(t, T_{n+1}, x(t), y(t)) \frac{\left(\int_t^{T_{n+1}} \sigma_f(t, u) du \right)}{A(t, x(t), y(t))} dW_t^{\mathbb{Q}}$$

On considère les conditions réunies pour en déduire l'existence d'un mouvement Brownien W^A sous \mathbb{Q}^A . Sous cette probabilité, le taux swap est une martingale continue et de carré intégrable en tant que quotient d'un actif échangeable et de l'annuité. Par application du théorème de représentation des martingales et de la formule d'Itô sur le taux swap, fonction \mathcal{C}^2 de x et y sur un intervalle de temps $[t, T]$ fini, il vient:

$$\boxed{dS(t, x(t), y(t)) = \frac{\partial S}{\partial x}(t, x(t), y(t)) \sigma_r(t, x(t), y(t)) dW_t^A} \quad (13)$$

Les dérivées partielles de S par rapport à x sont d'ailleurs facilement calculables du fait de leur forme exponentielle, après simplifications, on a:

$$\begin{aligned} \frac{\partial S}{\partial x}(t, x, y) &= \frac{-1}{A(t, x, y)} \left(G(t, T_0) P(t, T_0, x, y) - G(t, T_N) P(t, T_N, x, y) \right) \\ &\quad + \frac{S(t, x, y)}{A(t, x, y)} \sum_{n=0}^{N-1} \tau_n G(t, T_{n+1}) P(t, T_{n+1}, x, y) \end{aligned} \quad (14)$$

$$\begin{aligned} \frac{\partial^2 S}{\partial x^2}(t, x, y) &= \frac{1}{A(t, x, y)} \left[G(t, T_0)^2 P(t, T_0, x, y) - G(t, T_N)^2 P(t, T_N, x, y) - S(t, x, y) \frac{\partial^2 A}{\partial x^2} \right] \\ &\quad - \frac{\partial A}{\partial x} \left[\frac{\frac{\partial S}{\partial x}}{A(t, x, y)} - \frac{G(t, T_0) P(t, T_0, x, y) - G(t, T_N) P(t, T_N, x, y)}{A(t, x, y)^2} \right] + \left(\frac{\partial A}{\partial x} \right)^2 \frac{S(t, x, y)}{A(t, x, y)^2} \\ &= \frac{1}{A(t, x, y)} \left[G(t, T_0)^2 P(t, T_0, x, y) - G(t, T_N)^2 P(t, T_N, x, y) - S(t, x, y) A''(t, x, y) \right] \\ &\quad + \frac{2A'(t, x, y)}{A(t, x, y)^2} \left(A'(t, x, y) S(t, x, y) + (G(t, T_0) P(t, T_0, x, y) - G(t, T_N) P(t, T_N, x, y)) \right) \end{aligned} \quad (15)$$

Notons que la dynamique du taux swap fait intervenir une volatilité fonction des facteurs stochastiques x et y . Cela entraîne des complications du point de vue de la modélisation, c'est

pourquoi l'objectif de ce qui suit est d'approximer cette structure de volatilité par une volatilité locale qui dépend directement de la valeur de S . Pour ce faire, il s'agit d'abord d'exploiter la projection markovienne décrite dans Piterbarg 2006[Théorème 2.1].

Ce résultat nous permet d'écrire la dynamique du taux swap sous la forme suivante:

$$\begin{cases} dS(t, x(t), y(t)) &= \phi(t, S(t, x(t), y(t))) dW_t^A \\ \text{Où } \phi(t, s)^2 &= \mathbb{E}_{\mathbb{Q}^A} \left[\left(\frac{\partial S}{\partial x}(t, x(t), y(t)) \sigma_r(t, x(t), y(t)) \right)^2 | S(t, x(t), y(t)) = s \right] \end{cases}$$

Puisque le terme dans l'espérance conditionnelle ne dépend pas de S mais de x et y , les calculs n'en sont que plus complexes. Cependant, nous chercherons à approximer y par une fonction déterministe \bar{y} de sorte à considérer un lien fort entre x et S tel que $x(t) = X(t, S(t))$. Cette opération résoudrait le problème de l'espérance conditionnelle en ce sens que la connaissance de X conduirait à

$$\phi(t, s) \approx \frac{\partial S}{\partial x}(t, X(t, s), \bar{y}(t)) \sigma_r(t, X(t, s), \bar{y}(t))$$

Dans ce cas, on se ramènerait aux méthodes classiques de pricing d'un call dans un modèle à volatilité locale déterministe.

Afin d'approcher la fonction X , il convient de procéder, comme pour y , à une approximation déterministe \bar{x} de x . Cette approximation va servir de point de fixation sur lequel on va appliquer un développement de Taylor dans le but d'inverser la fonction S .

Si on écrit le développement de S au second ordre et qu'on injecte cette quantité dans l'expression

$$S(t, X(t, s), \bar{y}(t)) = s, \tag{16}$$

il vient (On notera l'approximation également X ici) :

$$S(t, \bar{x}(t), \bar{y}(t)) + \frac{\partial S}{\partial x}(t, \bar{x}(t), \bar{y}(t)) (X(t, s) - \bar{x}(t)) + \frac{1}{2} \frac{\partial^2 S}{\partial x^2}(t, \bar{x}(t), \bar{y}(t)) (X(t, s) - \bar{x}(t))^2 = s$$

Il nous reste à développer les méthodes d'approximation déterministes employées dans les calculs de \bar{x}, \bar{y} .

a) Estimation de l'espérance de y

Tout d'abord, il est convenable d'utiliser l'espérance d'une variable aléatoire comme valeur de référence lorsque l'on souhaite remplacer cette variable par une fonction déterministe à des fins de simplification.

En prenant l'espérance sous \mathbb{Q}^A dans la dynamique de y et en faisant l'approximation

$$\mathbb{E}_{\mathbb{Q}^A} [\sigma_r(t, x(t), y(t))^2] \approx \sigma_r(t, 0, 0)^2,$$

on trouve que l'espérance de y satisfait l'équation différentielle ordinaire suivante:

$$d\mathbb{E}_{\mathbb{Q}^A}[y(t)] \approx (\sigma_r(t, 0, 0)^2 - 2\chi\mathbb{E}_{\mathbb{Q}^A}[y(t)]) dt, \quad y(0) = 0$$

Finalement,

$$\bar{y}(t) = h(t)^2 \int_0^t \sigma_r(t, 0, 0)^2 h(s)^{-2} ds$$

Dans notre cas avec coefficients constants, cela donne:

$$\bar{y}(t) = \frac{\lambda^2 a^2}{2\chi} (1 - e^{-2\chi t})$$

b) Estimation de l'espérance de x au premier ordre

Si $(t, s) \rightarrow X(t, s)$ est la solution de $S(t, X(t, s), \bar{y}(t)) = s$ et si on note $x_0(t) = X(t, S(0))$, on peut trouver numériquement la valeur de x_0 grâce à un algorithme de Newton-Raphson du fait que le taux swap est continue, monotone et deux fois différentiable en x .

Prouvons la monotonie de S :

Preuve:

Soit $0 \leq t \leq T_0$, $(x, y) \in \mathbb{R}^2$.

On reprend l'expression de la dérivée du taux swap par rapport à x (14) et on note

$$B(t, x, y) = \sum_{n=0}^{N-1} \tau_n P(t, T_{n+1}, x, y) G(t, T_{n+1})$$

.

Par substitution et ré-arrangement des termes, on trouve:

$$\begin{aligned} \frac{\partial S}{\partial x}(t, x, y) &= -\frac{1}{A(t, x, y)} (P(t, T_0, x, y) G(t, T_0) - P(t, T_N, x, y) G(t, T_N)) + \frac{P(t, T_0, x, y) - P(t, T_N, x, y)}{A(t, x, y)^2} B(t, x, y) \\ &= \frac{P(t, T_0, x, y)}{A(t, x, y)} \left(\frac{B(t, x, y)}{A(t, x, y)} - G(t, T_0) \right) - \frac{P(t, T_N, x, y)}{A(t, x, y)} \left(\frac{B(t, x, y)}{A(t, x, y)} - G(t, T_N) \right) \end{aligned}$$

Puisque h est une fonction positive, $\forall n \in [1, N]$, $G(t, T_0) \leq G(t, T_n) \leq G(t, T_N)$. Cela entraîne :

$$\sum_{n=0}^{N-1} \tau_n P(t, T_{n+1}, x, y) G(t, T_0) \leq \sum_{n=0}^{N-1} \tau_n P(t, T_{n+1}, x, y) G(t, T_{n+1}) \leq \sum_{n=0}^{N-1} \tau_n P(t, T_{n+1}, x, y) G(t, T_N)$$

Ou encore

$$G(t, T_0) \leq \frac{B(t, x, y)}{A(t, x, y)} \leq G(t, T_N) \implies \frac{B(t, x, y)}{A(t, x, y)} - G(t, T_0) \geq 0 \quad \text{et} \quad \frac{B(t, x, y)}{A(t, x, y)} - G(t, T_N) \leq 0$$

On conclut la preuve en remarquant que $\frac{\partial S}{\partial x}(t, x, y) \geq 0$.

Pour les couples (t, s) tels que $\frac{\partial S}{\partial x}(t, X(t, s), \bar{y}(t)) \neq 0$, X est aussi 2 fois différentiable. Ainsi par la formule de Taylor

$$x(t) = X(t, S(t)) = x_0(t) + \frac{\partial X}{\partial s}(t, S(0)) \times (S(t) - S(0)) + \mathcal{O}(S(t) - S(0))^2$$

et en prenant l'espérance sous l'annuity measure, comme le taux swap est une \mathbb{Q}^A martingale on a

$$\boxed{\mathbb{E}_{\mathbb{Q}^A}[x(t)] - x_0(t) = \mathcal{O}(\mathbb{E}_{\mathbb{Q}^A}(S(t) - S(0))^2)}$$

c) Estimation de l'espérance de x au second ordre

Le développement asymptotique au second ordre est donné par:

$$x(t) - x_0(t) = \frac{\partial X}{\partial s}(t, S(0)) \times (S(t) - S(0)) + \frac{1}{2} \frac{\partial^2 X}{\partial s^2}(t, S(0)) \times (S(t) - S(0))^2 + \mathcal{O}((S(t) - S(0))^3)$$

En prenant l'espérance à nouveau, cela donne:

$$\boxed{\mathbb{E}_{\mathbb{Q}^A}[x(t)] = x_0(t) + \frac{1}{2} \frac{\partial^2 X}{\partial s^2}(t, S(0)) \times Var^A[S(t)] + \mathcal{O}(\mathbb{E}_{\mathbb{Q}^A}(S(t) - S(0))^3)}$$

Nous avons besoin d'exprimer la dérivée seconde de X en dérivant deux fois l'expression (16) par rapport à s

$$\frac{\partial^2 X}{\partial s^2}(t, s) = - \frac{\frac{\partial^2 S}{\partial x^2}(t, X(t, s), \bar{y}(t))}{\left(\frac{\partial S}{\partial x}(t, X(t, s), \bar{y}(t))\right)^3}$$

Remarque: La difficulté réside dans l'expression de la variance en raison de la stochasticité de la volatilité. On a donc

$$Var^A[S(t)] = \mathbb{E}_{\mathbb{Q}^A} \left[\int_0^t \frac{\partial S}{\partial x}(s, x(s), \bar{y}(s))^2 \sigma_r(s, x(s))^2 ds \right]$$

que l'on pourrait approcher par un algorithme de Monte Carlo par exemple. Mais une possibilité est de considérer une dynamique gaussienne pour le taux swap avec une volatilité déterministe $\frac{\partial S}{\partial x}(s, 0, 0) \sigma_r(s, 0)$ telle que $Var^A[S(t)] \approx \int_0^t \frac{\partial S}{\partial x}(s, 0, 0)^2 \sigma_r(s, 0)^2 ds$

Discussion

Nous avons souligné le fait que la fonction x_0 pouvait être évaluée à chaque instant t assez simplement grâce à une méthode de Newton-Raphson. Néanmoins, cette méthode présente une limite que nous allons mettre en lumière ici. En effet, on peut dégager une erreur d'approximation

de la racine qui croît lorsque le tenor de la swaption étudiée diminue, c'est-à-dire lorsque la maturité de la swaption T_0 tend vers la date de paiement terminale du swap sous-jacent T_N . Pour prendre conscience de ce problème, nous tracerons la courbe de la fonction $x \rightarrow S(x) - S_0$ dont nous devons trouver la racine. Nous choisirons d'abord un tenor de $15Y$ et une maturité de $T_0 = 1Y$, puis nous diminuerons le tenor en augmentant la maturité T_0 sans changer la date terminale $T_N = 16Y$. Voici, ci-dessous (Figure 7) les graphiques observés (à l'instant t fixé à 0.95).

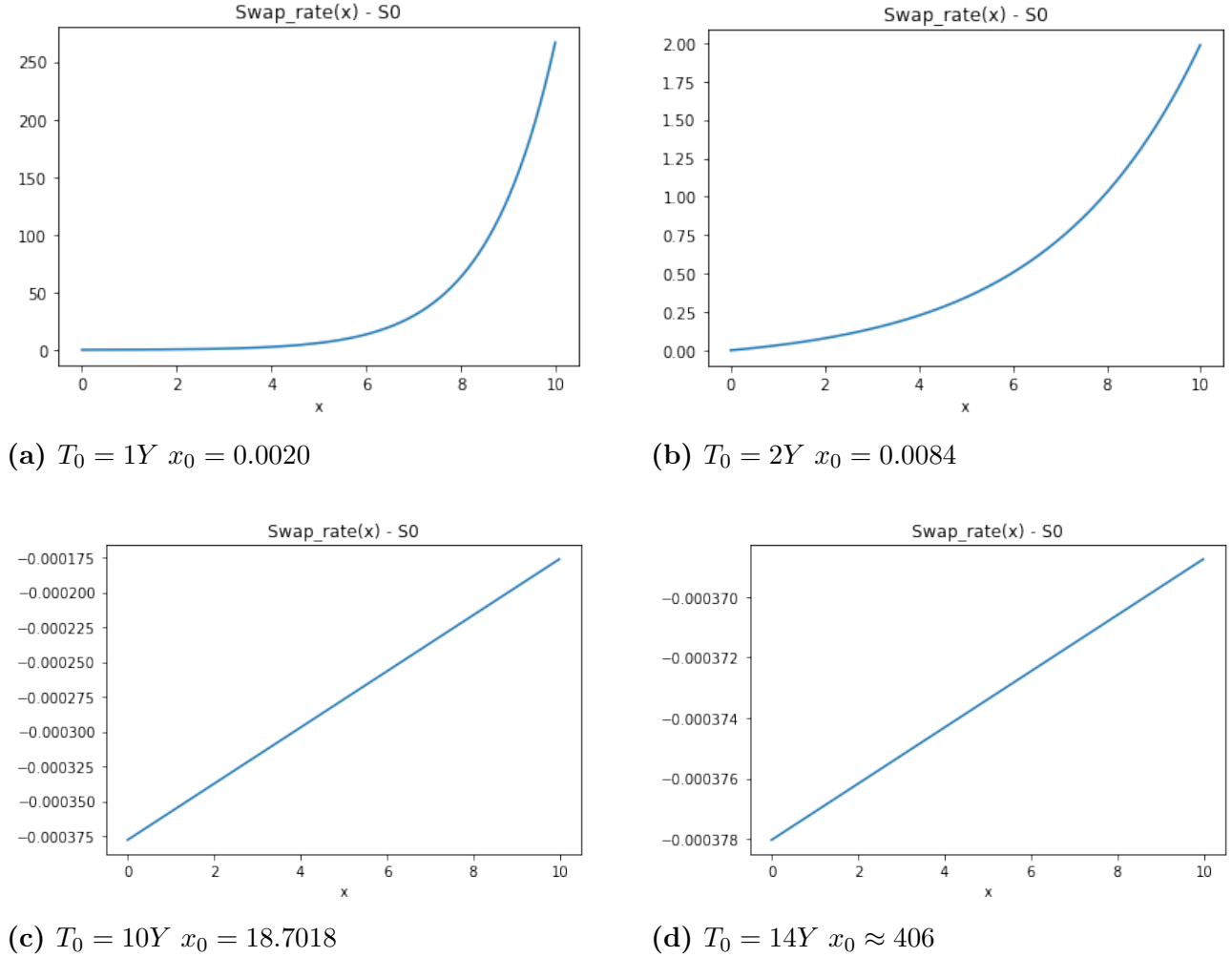
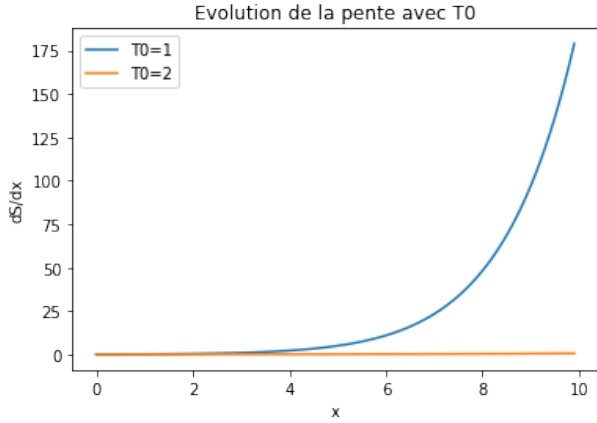


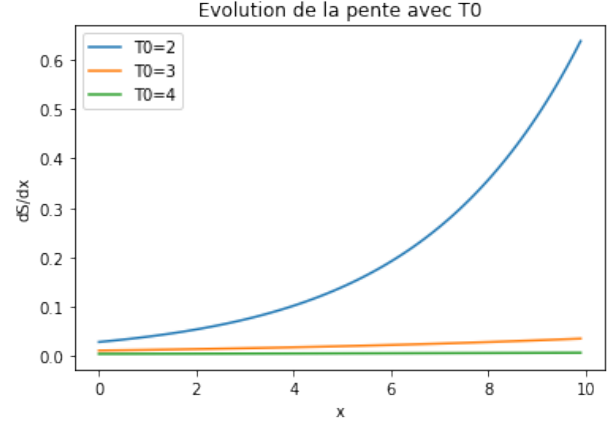
Figure 7: Evolution de la courbe de taux swap avec un tenor décroissant

Nous pouvons noter ici que la pente du taux swap tend à décroître fortement lorsque la maturité T_0 augmente, à date terminale T_N fixe. Leif B.G Andersen et Vladimir Piterbarg soulignent d'ailleurs dans leur livre Andersen 2010b[Page 546] que le taux swap peut empiriquement être approximé par une fonction quadratique de x , ce que l'on observe sur les graphique (On pourrait penser à une fonction de la forme $\alpha_{T_0} x^2$ où le coefficient α_{T_0} diminue avec T_0).

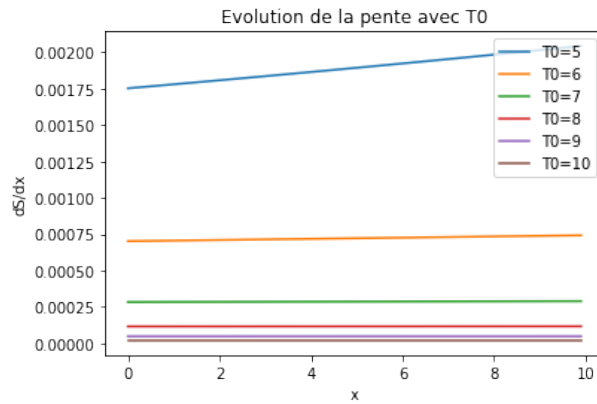
La figure 8 nous montre ce que l'on observe si l'on trace la dérivée $\frac{\partial S}{\partial x}$.



(a) Tenor 14Y à 15Y



(b) Tenor 14Y à 12Y



(c) Tenor 11Y à 6Y

Figure 8: Evolution de la pente $\frac{\partial S}{\partial x}$

On remarque bien la décroissance exponentielle de la pente, toutefois nous ne sommes parvenus à aucune démonstration mathématique malgré la forme explicite de la dérivée étudiée. (14).

Remarque: Le problème que cela implique est que l'algorithme de Newton ne converge pas ou très lentement pour des fonctions avec des pentes très faibles comme on peut l'observer ici. Ainsi, les tests pour les courts tenor peuvent renvoyer une erreur comme c'est le cas pour $T_0 = 14Y$ et $T_N = 16Y$. Enfin, dans le cas où la racine est renvoyée par l'algorithme et est correcte, on peut voir que cette approximation perd son sens financier dans des cas limites présentés plus haut où x_0 atteint des valeurs bien trop élevées.

3.5.3 La formule de Pricing

A l'aide de ce qui précède, Leif B.G. Andersen et V. Piterbarg développent quelques approximations supplémentaires afin de simplifier encore la volatilité du taux swap (voir Andersen 2010b[Section 13.1.5] pour le détail des calculs). L'objectif est d'aboutir en définitive à ce qui ressemble à un modèle bien connu, en l'occurrence un modèle displaced log-normal que l'on

définit à partir de maintenant par la dynamique: $dS(t) = \sigma(S(t) - \theta) dW_t$, $S_0 > \theta$, $\sigma > 0$.

Dans un premier temps, il s'agit d'approcher la dynamique du taux swap par la dynamique d'un processus displaced log-normal avec des paramètres dépendants du temps. Ainsi vient la dynamique suivante:

$$dS(t) \approx \lambda_S(t) (b_S(t)S(t) + (1 - b_S(t))S(0)) dW^A(t)$$

$$\text{où} \quad \begin{cases} \lambda_S(t) = \lambda_{\frac{1}{S(0)} \frac{\partial S}{\partial x}}(t, \bar{x}(t), \bar{y}(t)) (a + b\bar{x}(t)) \\ b_S(t) = \frac{S(0)}{(a + b\bar{x}(t))} \frac{b}{\frac{\partial S}{\partial x}(t, \bar{x}(t), \bar{y}(t))} + \frac{S(0) \frac{\partial^2 S}{\partial x^2}(t, \bar{x}(t), \bar{y}(t))}{\left(\frac{\partial S}{\partial x}(t, \bar{x}(t), \bar{y}(t))\right)^2} \end{cases}$$

Dans un second temps, il reste à contourner l'obstacle que représente la dépendance en temps des paramètres. Puisque l'on cherche à obtenir une formule fermée et que la formule de pricing d'un Call est bien connue dans le cas d'une dynamique log-displaced à paramètres constants (voir Annexe 9), nous allons nous ramener au cas des paramètres constants. Pour faire cela de manière judicieuse, il est commun de faire appel à des techniques de moyennage (ou *Averaging methods*). Ces techniques ont le bon goût de nous éviter des formules complexes et moins réductibles. Pour illustrer un cas concret et faire face aux difficultés du pricing avec coefficients dépendants du temps, nous pouvons nous intéresser au cas displaced log-normal avec des paramètres en escalier. Une esquisse de ce calcul sur deux dates est décrite en annexe 10. Dans le cadre de notre étude, nous allons reprendre la méthode de moyennage suggérée dans Andersen 2010b[Section 13.1.5] que voici:

$$\begin{aligned} \bar{\lambda}_S &= \left(\frac{1}{T_0} \int_0^{T_0} \lambda_S(t)^2 dt \right)^{1/2} \\ \bar{b}_S &= \int_0^{T_0} b_S(t) w_S(t) dt \\ w_S(t) &= \frac{\lambda_S(t)^2 \int_0^t \lambda_S(s)^2 ds}{\int_0^{T_0} (\lambda_S(u)^2 \int_0^u \lambda_S(s)^2 ds) du} \end{aligned}$$

Cela nous amène, par le biais de la formule du call dans le modèle displaced log-normal (annexe 9), à la formule de pricing suivante

Proposition 3

$$\boxed{\begin{aligned} V_0 &= A(0) \left((S(0) + \frac{1-\bar{b}_S}{\bar{b}_S} S(0)) \Phi(d^+) - (K + \frac{1-\bar{b}_S}{\bar{b}_S} S(0)) \Phi(d^-) \right) \\ d^{+/-} &= \frac{\ln \left(\frac{S(0) + S(0) \frac{1-\bar{b}_S}{\bar{b}_S}}{K + S(0) \frac{1-\bar{b}_S}{\bar{b}_S}} \right) + /- \frac{\bar{\lambda}_S^2 \bar{b}_S^2 T_0}{2}}{\bar{\lambda}_S \bar{b}_S \sqrt{T_0}} \end{aligned}}$$

À titre de comparaison, voici ci-dessous la courbe de prix de la swaption de tenor 5Y et de maturité $T_0 = 1Y$ obtenue d'une part grâce à la méthode de Monte Carlo et d'autre part grâce à l'approximation décrite dans cette section.

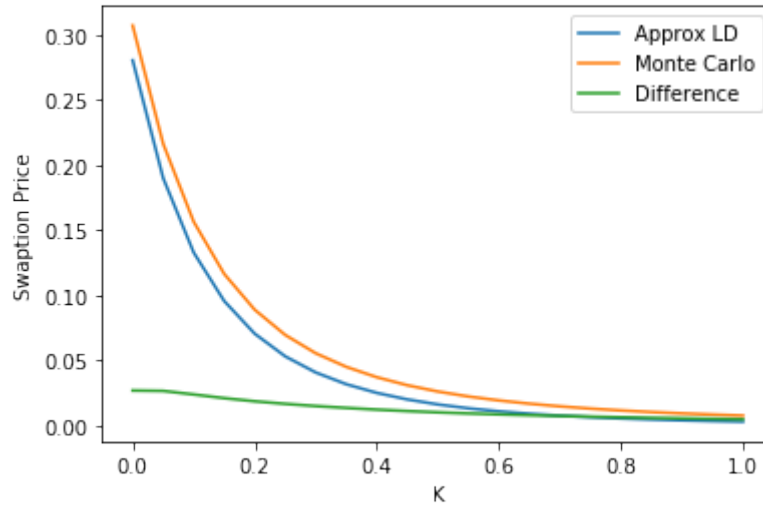


Figure 9: Comparaison des méthodes de pricing

Sur cette figure, la différence a été tracée sur une très large gamme de strikes. Le strike ATM se situe vers 0.05 dans la région où la différence absolue entre les deux méthodes est la plus grande certes, mais où la différence relative est la plus faible. L'aspect qui nous intéresse également est la vitesse de calcul, qui est uniquement freinée ici par les estimations d'intégrales dans les paramètres $\bar{\lambda}, \bar{b}, w$ et par les méthodes de recherche de racine pour x_0 . Plus la maturité T_0 augmente, plus le nombre de calculs est important. Il n'en reste pas moins que cette méthode est bien moins gourmande en temps de calcul que la méthode de Monte Carlo. D'ailleurs, nous conseillons d'utiliser une structure de dictionnaire dans l'implémentation de \bar{x}, \bar{y} afin de mémoriser les valeurs des appels récurrents.

| T_0 (Y) | Temps d'exécution (sec) |
|-----------|-------------------------|
| 1 | 4 |
| 5 | 15 |
| 10 | 45 |

| M | Temps d'exécution (sec) | $\hat{\sigma}_M^2$ |
|---------|-------------------------|--------------------|
| 1 000 | 0.8 | 0.080 |
| 10 000 | 8 | 0.083 |
| 100 000 | 87 | 0.083 |

Performances de l'approximation Log-displaced (Gauche) et Monte Carlo (Droite)

Tenor 5Y avec paiements semi-annuels

4 Courbe de volatilité implicite

Dans cette section, il s'agira d'expliciter la méthode de construction de la surface de volatilité implicite, étape qui s'avérera cruciale pour calibrer notre modèle sur des prix observés. La volatilité implicite dépend du modèle de pricing dans lequel elle est calculée. Le modèle généralement choisi pour impliciter la volatilité est celui de Black & Scholes (ou log-normal) mais nous verrons que la volatilité implicite dans le modèle normal peut être intéressante dans certains cas.

4.1 Volatilité implicite log-normale

Commençons par rappeler le modèle de Black & Scholes et la formule de valorisation de la swaption dans celui-ci.

On considère le taux swap S log-normal satisfaisant la dynamique suivante

$$dS_{T_0, \dots, T_N}(t) = S_{T_0, \dots, T_N}(t) \sigma dW_t.$$

À l'aide des détails des sections 3.1.6 et 3.6.2 et par la formule de Bayes, il vient

$$V_0^{BS} = \mathbb{E}_{\mathbb{Q}} \left[e^{-\int_0^t r(s) ds} Swap(T_0)^+ \right] = A(0) \mathbb{E}_{\mathbb{Q}^A} \left[(S(T_0) - K) \mathbf{1}_{S(T_0) > K} \right]$$

On pose

$$\frac{d\tilde{\mathbb{Q}}}{d\mathbb{Q}^A} = \frac{S(T_0)}{\mathbb{E}[S(T_0)]} = e^{-\frac{\sigma^2}{2} T_0 + \sigma W_{T_0}}$$

et d'après le théorème de Girsanov,

$$\tilde{W}_t = W_t^A - \sigma t$$

est un mouvement Brownien sous $\tilde{\mathbb{Q}}$. Cette mesure va nous permettre de développer les calculs ci-dessous.

On se place dans le cas où $S(0) > 0$, $K > 0$.

Cela nous mène à la formule de Black Scholes suivante:

$$\begin{aligned} V_0^{BS} &= A(0) S(0) \tilde{\mathbb{Q}} \left(S(0) e^{\frac{\sigma^2}{2} T_0 + \sigma \tilde{W}_{T_0}} > K \right) - K A(0) \mathbb{Q}^A \left(\sigma W_{T_0} > \ln \frac{K}{S(0)} + \frac{\sigma^2}{2} T_0 \right) \\ &= \boxed{A(0) S(0) \Phi(d^+) - K A(0) \Phi(d^-)} \end{aligned}$$

$$\begin{cases} d^+ &= \frac{\ln \frac{S(0)}{K} + \frac{\sigma^2}{2} T_0}{\sigma \sqrt{T_0}} \\ d^- &= d^+ - \sigma \sqrt{T_0} \end{cases}$$

Disjonction de cas:

$$\bullet S_0 < 0 \text{ \& } K > 0 \implies V_0^{BS} = 0$$

- $S_0 > 0$ & $K < 0 \implies V_0^{BS} = A(0)S(0) - KA(0)$
- $S_0 < 0$ & $K < 0 \implies V_0^{BS} = A(0)S(0)\tilde{\mathbb{Q}}\left(e^{\frac{\sigma^2}{2}T_0 + \sigma\tilde{W}_{T_0}} < \frac{K}{S(0)}\right) - KA(0)\mathbb{Q}^A\left(e^{\frac{-\sigma^2}{2}T_0 + \sigma W_{T_0}^A} < \frac{K}{S(0)}\right) = A(0)S(0)\Phi(-d^+) - KA(0)\Phi(-d^-)$
- Enfin si $S(0) = 0$ ou $K = 0$ ou $\sigma = 0$, nous obtenons $V_0^{BS} = A(0)(S(0) - K)\mathbf{1}_{S(0) > K}$

Le premier cas traité est le plus fréquent et c'est celui que nous rencontrerons dans les tests numériques. Dans ce cas, on peut noter en premier lieu le comportement asymptotique du prix:

$$\begin{aligned} V_0^{BS} &\xrightarrow[\sigma \rightarrow \infty]{} A(0)S(0) \\ \text{et } V_0^{BS} &\xrightarrow[\sigma \rightarrow 0^+]{} A(0)(S(0) - K)^+ \end{aligned} \quad (17)$$

En second lieu, on remarque la monotonie en σ . En effet, $\frac{\partial d^+}{\partial \sigma} = \frac{\partial d^-}{\partial \sigma} + \sqrt{T_0}$ et

$$vega = \frac{\partial V_0^{BS}}{\partial \sigma} = A(0)S(0)\frac{\partial d^+}{\partial \sigma}\phi(d^+) - KA(0)\frac{\partial d^-}{\partial \sigma}\phi(d^-) = A(0)S(0)\sqrt{\frac{T_0}{2\pi}}e^{-\frac{(d^+)^2}{2}} > 0$$

Visualisons cela de manière graphique:

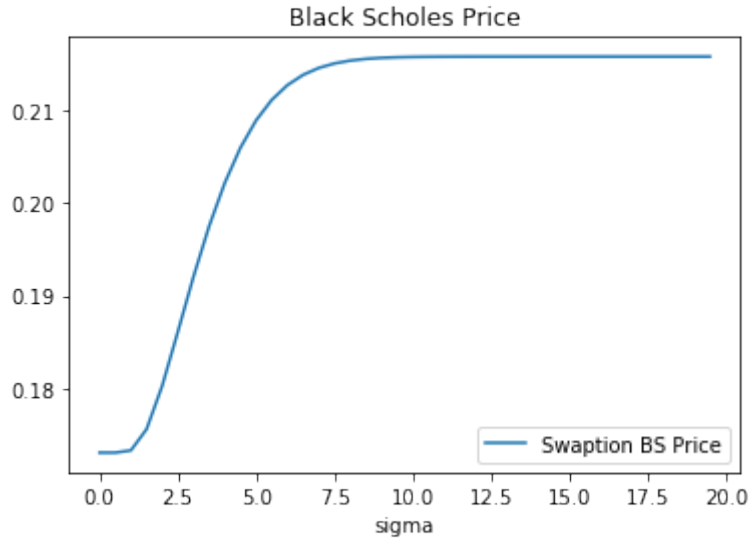


Figure 10: Courbe de prix Black Scholes

Si on observe sur le marché un prix \hat{V} pour le couple (T_0, K) , la volatilité implicite σ_{imp} est définie telle que

$$V_0^{BS}(T_0, K, \sigma_{imp}(T_0, K)) = \hat{V}.$$

Ce que l'on a montré précédemment sur la monotonie des prix par rapport à la volatilité Black Scholes nous permet d'inverser un prix observé \hat{V} compris entre les limites du prix Black Scholes

(17). Pour inverser le prix, il est possible d'utiliser une méthode classique de bissectrice, ou alors la méthode de Newton Raphson qui consiste à effectuer l'itération:

$$\begin{cases} \sigma_0 & \text{choisi} \\ \sigma_{n+1} = \sigma_n - \left(\frac{V_0(\sigma_n) - \hat{V}}{\frac{\partial V_0}{\partial \sigma}(\sigma_n)} \right) \end{cases}$$

La forme du vega entraîne la possibilité d'obtenir des valeurs très faibles pour la dérivées calculée (notamment lorsque la maturité est petite et le strike loin de la monnaie). Cela provoque des instabilités dans la méthode de Newton et nous recommandons ici d'impliciter la volatilité grâce à la méthode de la bissectrice.

Ainsi, en notant V_0^{MC} la fonction de pricing par la méthode de Monte Carlo et en définissant la volatilité implicite σ_{imp} telle que

$$V_0^{BS}(T_0, K, \sigma_{imp}(T_0, K)) = V_0^{MC}(T_0, K),$$

il nous est possible de construire une surface de volatilité $(\sigma_{imp}(T_0, K))_{T_0, K}$ associée à notre modèle de Cheyette.

En prenant un tenor de 5Y, voici ce que l'on peut visualiser:

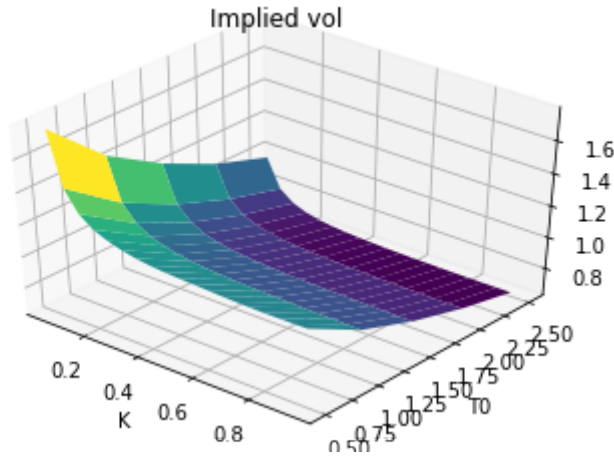


Figure 11: Surface de volatilité implicite dans le modèle log-normal

On peut noter la décroissance de la volatilité implicite par rapport au strike, et ce pour toutes les maturités. On ne peut le démontrer lorsque l'on inverse des prix Monte Carlo. Voici ce que l'on peut mettre en équation.

On sait que

$$\frac{\partial V_0^{MC}(K)}{\partial K} \leq 0, \frac{\partial V_0^{BS}}{\partial \sigma}(S_0, T, \sigma, K) > 0$$

On fixe T_0, S_0 .

On va utiliser la définition de la volatilité implicite:

$$\forall K > 0, V_0^{BS}(\sigma_{Imp}(K), K) = V_0^{MC}(K)$$

En dérivant cette équation par rapport au strike, il vient

$$\frac{\partial \sigma_{Imp}(K)}{\partial K} = \frac{\left(\frac{\partial V_0^{MC}(K)}{\partial K} - \frac{\partial V_0^{BS}(\sigma_{Imp}(K), K)}{\partial K} \right)}{\frac{\partial V_0^{BS}(\sigma_{Imp}(K), K)}{\partial \sigma}}.$$

Cette quantité a le même signe que son numérateur. Pour cette raison, la pente de la volatilité implicite par rapport à K peut être négative si la pente du prix Monte Carlo n'est pas aussi forte que la pente de la fonction de pricing dans le modèle log-normal. On peut observer cela de manière empirique:

On note

$$Diff = \left(\frac{\partial V_0^{MC}(K)}{\partial K} - \frac{\partial V_0^{BS}(\sigma_{Imp}(K), K)}{\partial K} \right)$$

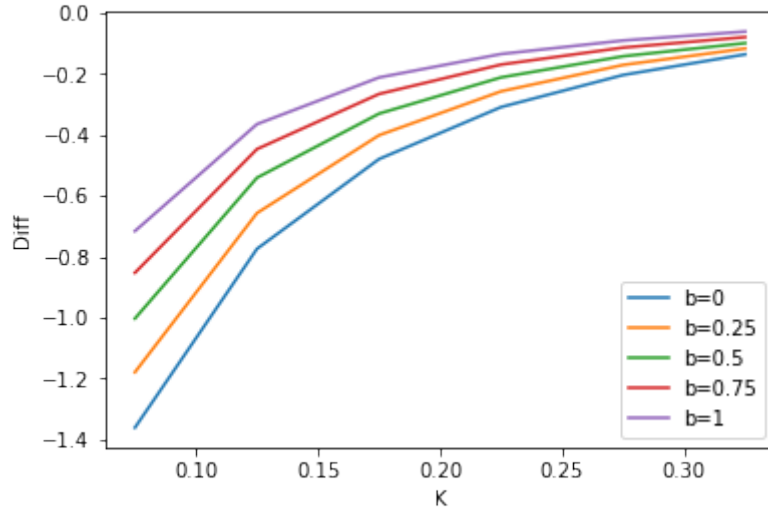


Figure 12: Signe du skew en fonction de b ($T_0 = 1Y$)

Ce phénomène de skew négatif peut en revanche être démontré dans le cas où on inverse des prix générés dans un modèle displaced log-normal, ce qui revient à peu de choses près à notre modèle au vue de l'approximation de prix faite dans la section 3.6. On peut d'ailleurs retrouver la preuve dans le théorème 1 de Roger Lee 2009, qui est le suivant:

Proposition 4 *Pour un taux swap suivant une dynamique displaced log-normal de décalage θ , $dS(t) = \sigma(S(t) - \theta)dW_t$ on a:*

$$\forall T > 0, \forall K \in \{K > \theta^+, \text{Call}^{BS}(S_0 - \theta, K - \theta, \sigma, T) < S_0\}, \quad \text{sign}\left(\frac{\partial \sigma_{imp}(K)}{\partial K}\right) = \text{sign}(\theta)$$

Dans notre cas, on note $\theta = -\frac{1-b}{b}S_0 < 0$, ce qui explique nos observations.

Remarque: On dégage un autre phénomène intéressant, celui de la décroissance de la pente en valeur absolue lorsque l'on augmente b . On y reviendra dans la section 4.3 avec une démonstration.

Contraintes de modèle

En raison de ses similitudes avec un modèle displaced log-normal, le modèle de Cheyette génère des prix ne pouvant exister dans le modèle de Black Scholes pour une même maturité. En d'autres termes, pour une maturité T_0 et des dates de paiement T_1, \dots, T_N données, certains prix issus du modèle de Cheyette sont en dehors des limites du modèles Black Scholes décrites en (17). Cela entraîne une incapacité à calculer la volatilité implicite sur les régions où cela survient.

Dans un premier temps, nous nous intéresserons au franchissement de la borne supérieure du modèle de Black Scholes, à savoir $\bar{L} = A(0)S(0)$. Pour ce faire, observons numériquement ce phénomène en représentant les courbes de prix Monte Carlo pour différentes maturités ainsi que les droites horizontales d'équation $y = \bar{L}(T_0)$ (en pointillés sur le graphique).

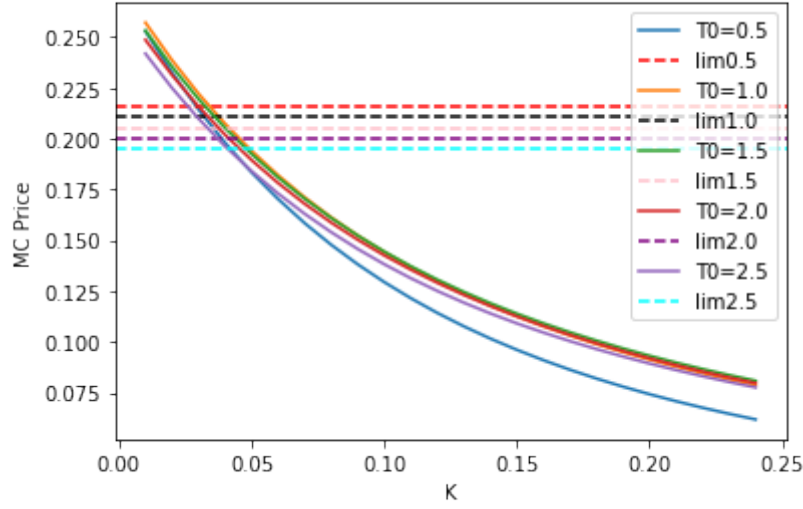


Figure 13: Franchissement de \bar{L}

D'une part, on remarque effectivement une région de strikes entre 0 et une certaine valeur positive sur laquelle la courbe de prix Monte Carlo est au-dessus de la valeur de \bar{L} associée. D'autre part, la limite $\bar{L} = P(0, T_0) - P(0, T_N)$ décroît lorsque T_0 augmente, à tenor $T_N - T_0$ constant. On peut le voir de manière immédiate dans le cas où le taux forward initial est constant par exemple. Les problèmes d'inversion des prix surviennent donc sur une plus large région de strikes pour des maturités élevées.

Une fois de plus, prouver ces observations dans le cas du modèle de Cheyette peut s'avérer épineux, c'est pourquoi nous allons nous placer dans le cas d'un modèle displaced log-normal. **On veut montrer que le prix d'un call dans ce modèle peut dépasser la limite supérieure du modèle Black Scholes.**

Proposition 5 $\exists K > 0, \mathbb{E}[(S(T) - K)^+] > S(0)$

Preuve:

On considère un sous jacent S qui suit la dynamique suivante

$$dS(t) = \sigma(S(t) - \theta) dW_t \quad \left(S(0) > 0, S(0) > \theta \quad \sigma > 0 \right)$$

où W est un mouvement Brownien sous la probabilité risque neutre.

On a donc

$$S(t) = \theta + (S(0) - \theta)e^{-\frac{\sigma^2}{2}t + \sigma W_t}$$

ce qui implique que le sous jacent prend ses valeurs dans $]\theta, +\infty[$.

Si $\theta > 0$, on montre qu'il n'existe pas de strike vérifiant la proposition.

On remarque que si S suit une dynamique displaced log-normal définie comme ci-dessus, le processus $(S(t) - \theta)_t$ suit une dynamique log-normal de volatilité σ .

On a donc

$$C^{LD}(T, K, \sigma, \theta) = \mathbb{E}[(S(T) - K)^+] = \mathbb{E}[(S(T) - \theta - (K - \theta))^+] = C^{BS}(S(0) - \theta, K - \theta, \sigma, T)$$

Comme un Call Black Scholes est borné par S_0 pour les strikes positifs, on a ici

$$\forall K > \theta, \mathbb{E}[(S_T - K)^+] < S(0) - \theta < S(0).$$

De plus, si $0 < K < \theta$, $\mathbb{E}[(S_T - K)^+] = \mathbb{E}[S_T - K] = S(0) - K < S(0)$

Si $\theta < 0$ et $K = 0$, on a:

$$\mathbb{E}[(S(T))^+] = \int_0^{+\infty} x f_{S(T)}(x) dx$$

où $f_{S(T)}$ est la densité de probabilité de $S(T)$.

Or

$$S(0) = \mathbb{E}[S(T)] = \int_{-\infty}^{+\infty} x f_{S(T)}(x) dx = \int_0^{+\infty} x f_{S(T)}(x) dx + \int_{-\infty}^0 x f_{S(T)}(x) dx.$$

Comme $f_{S(T)}$ est positive et continue sur \mathbb{R} , il faudrait que $\forall x \in]-\infty, 0[, f_{S(T)}(x) = 0$ pour que $\int_{-\infty}^0 x f_{S(T)}(x) dx = 0$.

Or

$$\begin{aligned} \mathbb{Q}(S(T) < 0) &= \mathbb{Q}\left(\mathcal{N}(0, 1) < \frac{\log(\frac{-\theta}{S(0)-\theta}) + \frac{\sigma^2}{2}T}{\sigma\sqrt{T}}\right) = 1 - \Phi\left(\frac{\log(\frac{-\theta}{S(0)-\theta}) + \frac{\sigma^2}{2}T}{\sigma\sqrt{T}}\right) \\ &= \int_{-\infty}^0 f_{S(T)}(x) dx > 0 \end{aligned}$$

Cela prouve que $\exists x \in]-\infty, 0[, f_{S(T)}(x) \neq 0$ et que $\int_{-\infty}^0 x f_{S(T)}(x) dx < 0$

Donc

$$Call(S(0), 0) > S(0)$$

On montre que pour tout $S(0)$, la fonction $K \rightarrow \text{Call}(S(0), K)$ est dérivable sur $]0, +\infty[$. Par le théorème de dérivabilité des intégrales à paramètres, on peut effectuer ce qui suit:

$$\begin{aligned}
\frac{\partial \text{Call}}{\partial K}(S(0), K) &= \frac{\partial}{\partial K} \int_K^{+\infty} (x - K) f_{S_T}(x) \, dx \\
&= -(K - K) f_{S_T}(K) + \int_K^{+\infty} \frac{\partial}{\partial K} (x - K) f_{S_T}(x) \, dx \quad (\text{R\`egle de Leibniz}) \\
&= - \int_K^{+\infty} f_{S_T}(x) \, dx \\
&= -\mathbb{P}(S_T > K).
\end{aligned}$$

Le Call est donc une fonction dérivable du strike et à plus forte raison continue.

Le Call étant une fonction décroissante et continue du strike, puisque l'image réciproque de $]S(0), +\infty[$ par la fonction $\text{Call}(S(0), \cdot)$ est un ouvert, on en déduit

$$\boxed{\exists K > 0, \text{Call}(S(0), K) > S(0)}$$

Dans un second temps, nous allons nous intéresser aux problèmes numériques rencontrés dus à la borne inférieure des prix du modèle Black Scholes. On sait que les prix Black Scholes sont au-dessus de la borne $\underline{L} = A(0)(S(0) - K)^+$. Cependant, le modèle displaced log-normal choisi avec une faible valeur de λ et une valeur de b proche de 1 (un modèle Black Scholes avec un faible décalage) produit une courbe de prix ressemblant à celle du modèle Black Scholes. Celle-ci peut être vue comme la valeur \underline{L} à laquelle on ajoute une valeur temps positive, croissante en fonction de T_0 et qui s'estompe lorsque l'on s'éloigne de la monnaie.

Dans la pratique, on observe que le niveau de volatilité implicite retombe à 0 pour des strikes éloignés de la monnaie et des maturités courtes.

Si on regarde les prix de swaption dans le modèle displaced log-normal $(\lambda, b) = (0.2, 0.9)$ avec une maturité de 6 mois, on visualise ce genre de graphique:

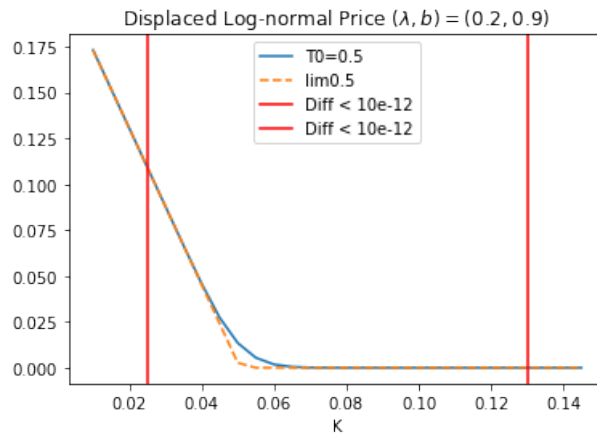


Figure 14: Atteinte d'un voisinage de \underline{L}

- En pointillés oranges, c'est la limite inférieure des prix Black Scholes \underline{L} , c'est à dire le prix qui expire aujourd'hui (sans la valeur temps), ou le prix si $\sigma = 0$.
- En bleu, c'est le prix du swaption avec maturité 6M (borne inférieure + valeur temps).
- En rouge, ce sont les limites de strikes à partir desquelles $V_0^{LD} - A(0) * (S(0) - K)^+ \leq 10^{-12}$.

Ce qu'il se passe dans l'algorithme d'inversion pour trouver la volatilité implicite, c'est que si le prix d'entrée est supérieure à la borne inférieure Black Scholes avec une différence de moins de 10^{-12} , le langage de programmation va effectuer un arrondi et la fonction renvoie 0 pour la volatilité implicite. En fait cette limite est atteinte en théorie pour une volatilité Black Scholes $\sigma = 0$ ou alors pour $\sigma \neq 0$ et $K \rightarrow 0$ ou $K \rightarrow \infty$. L'algorithme qui calcule la volatilité implicite par inversion des prix ne distingue pas ces deux cas de figure, c'est la raison pour laquelle la volatilité implicite calculée est nulle sur des régions de strikes éloignées de la monnaie pour de faibles maturités.

4.2 Volatilité implicite normale

De manière analogue à la section précédente, commençons par rappeler succinctement la formule de valorisation dans le modèle normal. Dans celui-ci, la dynamique du taux swap est simplement $dS_t = \sigma dW_t$.

Dans ce cas, on a

$$\begin{aligned}
\mathbb{E}[(S_T - K)^+ | \mathcal{F}_t] &= \mathbb{E}[(S_T - K)^+ | S_t = S] \\
&= \mathbb{E}\left[\left(S - K + \sigma(W_T - W_t)\right) \mathbf{1}_{S_T > K} | S_t = S\right] \\
&= (S - K)\Phi\left(\frac{S-K}{\sigma\sqrt{v}}\right) + \sigma\sqrt{T-t} \int_{\frac{K-S}{\sigma\sqrt{v}}}^{\infty} \frac{x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
&= (S - K)\Phi(d) + \sigma\sqrt{T-t}\phi(d)
\end{aligned}$$

De ce fait, nous avons la valeur de la swaption V_0^N suivante

$$\boxed{
\begin{aligned}
V_0^N &= A(0) \left[(S(0) - K)\Phi(d) + \sigma\sqrt{T}\phi(d) \right] \\
d &= \frac{S(0) - K}{\sigma\sqrt{T}}
\end{aligned}
}$$

La volatilité implicite dans ce modèle est donc

$$V_0^N(S_0, T, K, \sigma_{Imp}(K)) = V_0^{MC}(K).$$

L'intérêt d'utiliser ce modèle plutôt que le modèle de Black Scholes serait de pouvoir impliquer des prix générés dans le modèle de Cheyette sans faire de considération sur la borne supérieure des prix dans le modèle normal puisqu'il n'y en a pas. En effet, ce modèle est borné inférieurement (par \underline{L}) mais pas supérieurement car $V_0^N(S_0, T, K, \sigma) \xrightarrow[\sigma \rightarrow \infty]{} +\infty$.

De surcroît, on conserve la stricte monotonie par rapport à la volatilité puisque

$$\frac{\partial V_0^N}{\partial \sigma}(S_0, T, K, \sigma) = A(0)\sqrt{T}\phi(d) > 0.$$

Avec ce qui précède, on peut construire une surface de volatilité implicite pour des strikes proches de 0, où cela était potentiellement impossible dans le modèle log-normal. Voici ci-dessous une surface de volatilité implicite pour un tenor de 5 ans.

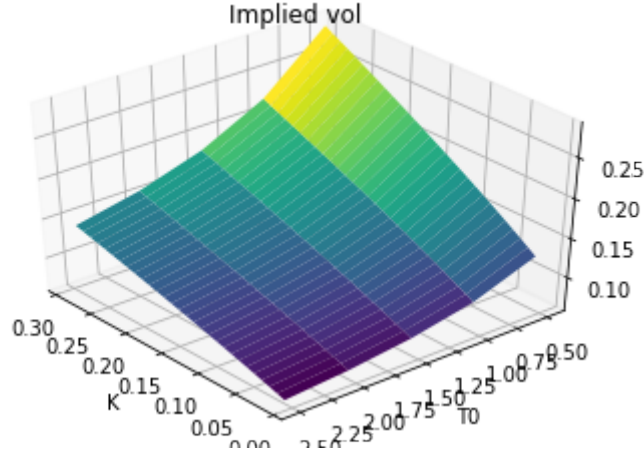


Figure 15: Surface de volatilité implicite dans le modèle normal

À l'inverse du modèle log-normal, le modèle normal semble fournir une surface de volatilité implicite au skew positif. En effet, si l'on garde à l'esprit l'approximation des prix Monte Carlo par un modèle displaced log-normal, il est possible de démontrer cette déduction empirique.

Proposition 6 *Si on définit la volatilité implicite ici par $\tilde{\sigma}$ telle que*

$$V_0^N(S(0), K, \tilde{\sigma}, T) = V_0^{LD}(S(0), K, T, \sigma, \theta)$$

où V_0^N est le prix du call dans le modèle normal et V_0^{LD} le prix du call dans le modèle displaced log-normal. Alors

$$\forall T > 0, K > 0 \quad \text{s.t.} \quad C^{BS}(S(0) - \theta, K - \theta, \sigma, T) < S(0), \quad \frac{\partial \tilde{\sigma}}{\partial K}(T, K) > 0$$

Preuve:

On se propose d'abord de démontrer le résultat en se plaçant dans le cadre suivant:

Le pricing se fait dans le modèle displaced log-normal où le sous-jacent suit la dynamique

$$dS(t) = \sigma(S(t) - \theta) dW_t \quad \left(S(0) > 0, S(0) > \theta \quad \sigma > 0 \right).$$

On définit cette fois la volatilité implicite calculée dans le modèle displaced log-normal, c'est-à-dire σ_{LD} telle que

$$\forall K, T \geq 0, \quad C^{LD}(T, K, \sigma_{LD}(T, K), \theta_{LD}) = C^{LD}(T, K, \sigma, \theta) \quad (\star)$$

On a donc

$$C^{LD}(T, K, \sigma, \theta) = \mathbb{E}[(S(T) - K)^+] = \mathbb{E}[(S(T) - \theta - (K - \theta))^+] = C^{BS}(S(0) - \theta, K - \theta, \sigma, T)$$

L'équation (\star) devient alors

$$C^{BS}(S(0) - \theta_{LD}, K - \theta_{LD}, \sigma_{LD}(T, K), T) = C^{BS}(S(0) - \theta, K - \theta, \sigma, T)$$

On pose

$$\begin{cases} \tilde{S}(0) = S(0) - \theta_{LD} \\ \tilde{K} = K - \theta_{LD} \end{cases}$$

Il vient:

$$\boxed{C^{BS}(\tilde{S}(0), \tilde{K}, \sigma_{LD}(T, K), T) = C^{BS}(\tilde{S}(0) - (\theta - \theta_{LD}), \tilde{K} - (\theta - \theta_{LD}), \sigma, T)}$$

D'après **Théorème 1** de Roger Lee 2009,

$$\forall T > 0, K > 0 \quad s.t \quad C^{BS}(S(0) - \theta, K - \theta, \sigma, T) < \tilde{S}(0),$$

$$sgn \frac{\partial \sigma_{LD}}{\partial K}(T, K) = sgn(\theta - \theta_{LD})$$

Dans notre cas, les dynamique log-displaced sont sous la forme

$$dS(t) = \lambda(bS(t) + (1 - b)S(0)) dW_t$$

On pose donc $(b_1, b_2) \in]0, 1]^2$ tels que

$$\begin{cases} \theta = -\frac{1-b_1}{b_1}S(0) \\ \theta_{LD} = -\frac{1-b_2}{b_2}S(0) \end{cases}$$

Donc

$$\theta - \theta_{LD} = S(0) \left[\frac{1-b_2}{b_2} - \frac{1-b_1}{b_1} \right]$$

Comme $x \rightarrow \frac{1-x}{x} = \frac{1}{x} - 1$ est décroissante sur \mathbb{R}^{+*} ,

$$\text{sgn}(\theta - \theta_{LD}) = \text{sgn}\left(\frac{1-b_2}{b_2} - \frac{1-b_1}{b_1}\right) = \text{sgn}(b_1 - b_2)$$

Conclusion: Initialement, nous voulions impliciter la volatilité dans un modèle normal, soit un modèle displaced log-normal avec $b_2 = 0$ et avec des prix sous un modèle displaced log-normal $b_1 \in [0, 1]$. On conclut finalement avec ce qui précède que si $b_2 \rightarrow 0^+$, la pente de la volatilité implicite est bien positive.

Choix du modèle pour le calcul de volatilité implicite

Maintenant que nous avons présenté deux modèles de calcul de la volatilité implicite, il convient de faire un choix pour l'étape de calibration. Malgré les contraintes de modèle explicitées plus haut, le calcul de la volatilité implicite dans le modèle de Black Scholes reste le plus utilisé et le plus présent sur le marché. Pour cette raison nous choisirons d'impliquer la volatilité dans ce modèle, tout en étant conscient que les intervalles de strikes et de maturités devront possiblement être réduits pour construire la surface de volatilité implicite.

4.3 Observations et propriétés sur le skew

Nous veillerons à ne pas confondre les paramètres de la volatilité dans le modèle de Cheyette, que nous noterons à présent λ_{Chey}, b_{Chey} , avec les paramètres du modèle approximatif log-displaced $\bar{\lambda}, \bar{b}$.

Effet de b_{Chey} sur la pente de volatilité implicite

Nous avons déjà noté la décroissance en valeur absolue du numérateur de la volatilité implicite en fonction de b . Il est possible d'estimer la pente de la courbe de volatilité implicite par une différence finie centrée $\frac{\partial f}{\partial x}(x_0) \approx \frac{f(x_0+h)-f(x_0-h)}{2h}$, $h = 5\%S_0$ afin d'avoir une approche visuelle de l'effet de b_{Chey} .

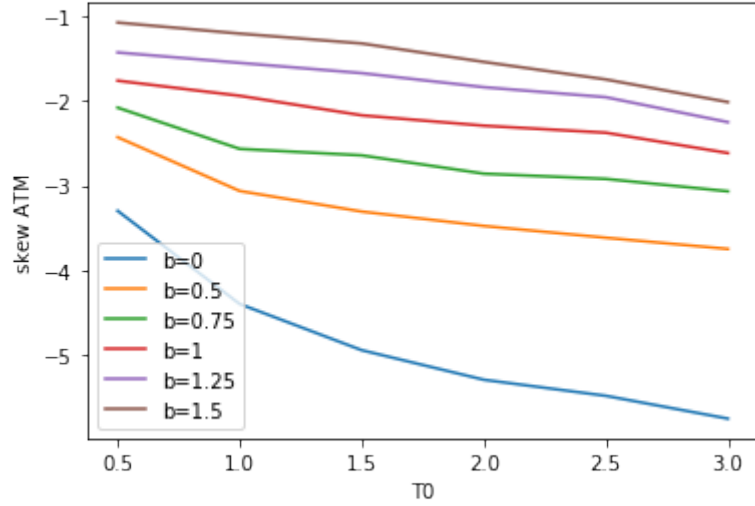


Figure 16: Skew ATM en fonction de b_{Chey}

Malgré les observations que l'on peut faire, b_{Chey} va intervenir dans la diffusion des facteurs x, y et il semble difficile de dégager une déduction mathématique de son impact sur le skew de la volatilité implicite. Si l'on étudie cet impact sur les prix générés dans l'approximation log-displaced de la section 3.6, on voit que b_{Chey} est omniprésent dans les termes $\bar{\lambda}, \bar{b}$. Il est donc complexe de démontrer clairement l'effet de b_{Chey} sur les prix et indirectement sur la volatilité implicite. En revanche, nous observons la croissance de \bar{b} par rapport à b_{Chey} .

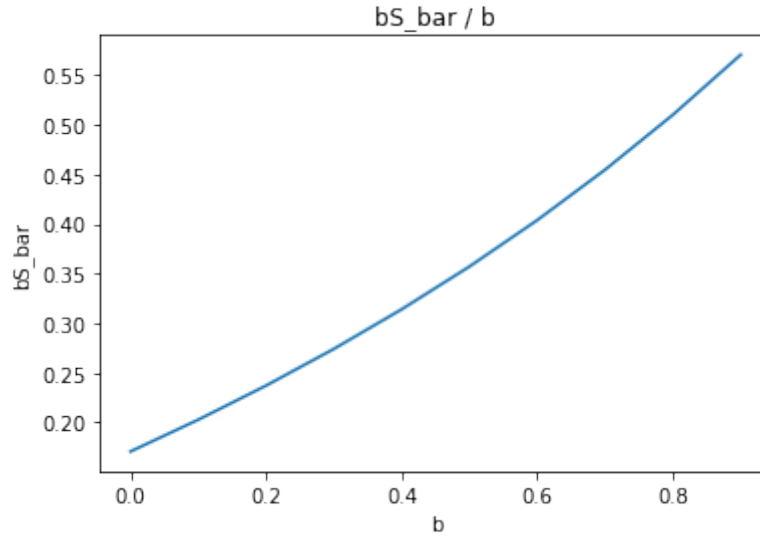


Figure 17: \bar{b} en fonction de b_{Chey}

Par ailleurs, il nous est possible d'apporter une preuve de la croissance du skew ATM lorsque \bar{b} croît dans le modèle displaced log-normal.

On fixe la maturité T_0 .

Définissons d'abord la fonction F qui renvoie la volatilité implicite calculée d'un modèle Log-displaced de paramètre b_1 dans un modèle Log-displaced de paramètre b_2 pour un strike donné. On a donc:

$$F^{b_1 \rightarrow b_2}(\sigma_K^{b_1}, K) \stackrel{\text{def}}{=} \sigma_K^{b_2}$$

où $\sigma_K^{b_1}, \sigma_K^{b_2}$ sont telles que

$$Call^{LD}(S_0, b_1, \sigma_K^{b_1}, K) = Call^{LD}(S_0, b_2, \sigma_K^{b_2}, K).$$

On veut montrer l'assertion suivante:

Proposition 7

$$\text{Si } 0 < b_1 < b_2 < 1, \text{ alors } \left. \frac{\partial F^{b_1 \rightarrow 1}}{\partial K} \right|_{\sigma_{S_0}^{b_1}, S_0} < \left. \frac{\partial F^{b_2 \rightarrow 1}}{\partial K} \right|_{\sigma_{S_0}^{b_2}, S_0}$$

Preuve:

On rappelle que $b = 1$ correspond au modèle log-normal.

D'abord, on a

$$\begin{aligned} F^{b_1 \rightarrow 1}(\sigma_K^{b_1}, K) &= \sigma_K^1 = F^{b_2 \rightarrow 1}(\sigma_K^{b_2}, K) \\ &= F^{b_2 \rightarrow 1}(F^{b_1 \rightarrow b_2}(\sigma_K^{b_1}, K), K) \end{aligned}$$

Ainsi

$$\begin{aligned} \left. \frac{\partial F^{b_1 \rightarrow 1}}{\partial K} \right|_{\sigma_{S_0}^{b_1}, S_0} &= \left. \frac{\partial F^{b_2 \rightarrow 1}}{\partial K} \right|_{F^{b_1 \rightarrow b_2}(\sigma_{S_0}^{b_1}, S_0), S_0} + \left. \frac{\partial F^{b_1 \rightarrow b_2}}{\partial K} \right|_{\sigma_{S_0}^{b_1}, S_0} \times \left. \frac{\partial F^{b_2 \rightarrow 1}}{\partial \sigma} \right|_{F^{b_1 \rightarrow b_2}(\sigma_{S_0}^{b_1}, S_0), S_0} \\ &= \left. \frac{\partial F^{b_2 \rightarrow 1}}{\partial K} \right|_{\sigma_{S_0}^{b_2}, S_0} + \left. \frac{\partial F^{b_1 \rightarrow b_2}}{\partial K} \right|_{\sigma_{S_0}^{b_1}, S_0} \times \left. \frac{\partial F^{b_2 \rightarrow 1}}{\partial \sigma} \right|_{\sigma_{S_0}^{b_2}, S_0} \end{aligned}$$

On a déjà démontré que

$$\left. \frac{\partial F^{b_1 \rightarrow b_2}}{\partial K} \right|_{\sigma_{S_0}^{b_1}, S_0} < 0 \quad \text{pour } b_1 < b_2$$

De plus, on sait que

$$Call^{LD}(S_0, b, \sigma, K) = Call^{BS}(S_0 + \frac{1-b}{b}S_0, \sigma, K + \frac{1-b}{b}S_0)$$

et comme le prix du call dans le modèle de Black Scholes est strictement croissant par rapport à la volatilité σ , il vient:

$$\left. \frac{\partial F^{b_1 \rightarrow b_2}}{\partial \sigma} \right|_{\sigma_{S_0}^{b_1}, S_0} > 0, \quad \forall (b_1, b_2) \in [0, 1]^2$$

En effet, la volatilité implicite du modèle 1 vers le modèle 2 augmente lorsque la volatilité du modèle 1 d'entrée augmente.

On en déduit:

$$\left. \frac{\partial F^{b_1 \rightarrow 1}}{\partial K} \right|_{\sigma_{S_0}^{b_1}, S_0} - \left. \frac{\partial F^{b_2 \rightarrow 1}}{\partial K} \right|_{F^{b_1 \rightarrow b_2}(\sigma_{S_0}^{b_1}, S_0), S_0} < 0$$

et cela achève la preuve.

Effet de λ_{Chey} sur le niveau de volatilité implicite

On va raisonner de la même manière en ce qui concerne l'effet de λ_{Chey} sur la volatilité. On peut voir que λ_{Chey} contrôle le niveau globale de la surface de volatilité implicite et plus particulièrement le niveau de la volatilité implicite à la monnaie.

Si l'on prête attention à la forme de $\bar{\lambda}$, on voit que l'on peut considérer ce paramètre comme une fonction quasiment linéaire de λ_{Chey} . On peut d'ailleurs observer une droite presque parfaite si l'on trace $\lambda_{Chey} \rightarrow \bar{\lambda}(\lambda_{Chey})$.

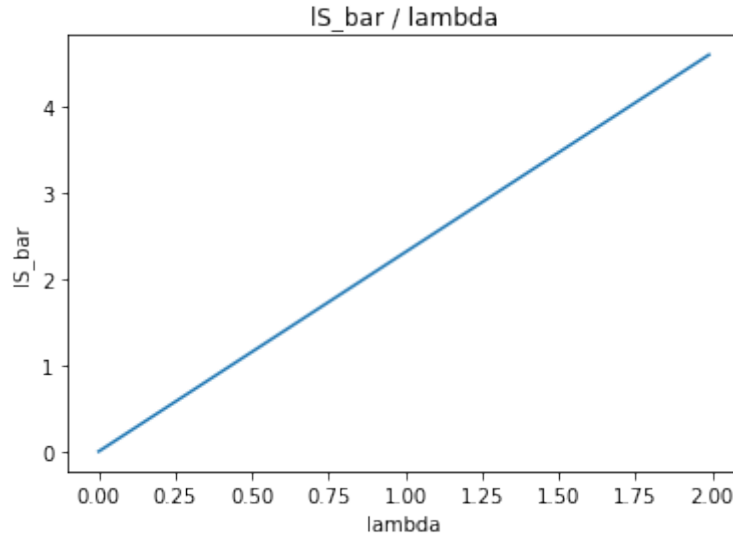


Figure 18: $\bar{\lambda}$ en fonction de λ_{Chey}

Enfin, on sait d'après le théorème 2.1 de Roger Lee 2009 que la volatilité implicite dans le modèle Black Scholes pour des prix displaced log-normaux de paramètres (λ, b) satisfait:

$$\forall T > 0, \quad \sigma_{imp}(ATM) \geq (1 + |\frac{1-b}{b}|)\lambda b$$

Donc dans notre étude, le niveau de la volatilité implicite augmente avec $\bar{\lambda}$, qui est elle-même une fonction croissante de λ_{Chey} .

Remarque: On peut souligner en parallèle que l'approximation des prix par un modèle displaced log-normal a ses limites. Les explications précédentes sont donc acceptables dans la mesure où la courbe de prix Monte Carlo et la courbe de prix log-displaced ne présentent pas

de différence notable. En effet les courbes de volatilité implicite dans ce modèle ont une décroissance 'rigide' qui ne reproduit pas nécessairement les smiles observés sur le marché ou même les smiles générés dans le modèle de Cheyette par la méthode Monte Carlo (voir p. 456 de D Brigo 2006).

5 Calibration de volatilité locale en dimension 1

Dans toute cette section, nous considérerons un swaption strip composé de $N - 1$ swaptions avec une grille de maturités $0 = T_0, T_1, \dots, T_N$. C'est-à-dire un ensemble de $N - 1$ swaptions où la n -ième swaption expire à la date T_n pour $n \in [1, N - 1]$ et dont le swap sous-jacent paye tous les 6 mois entre T_n et T_N . En d'autres termes, la structure tenor du n -ième swap sous-jacent est $(T_n, T_n + 6M, \dots, T_N)$. Nous veillerons à ne pas confondre le nombre N de dates considérées avec le paramètre du schéma d'Euler dénoté également N dans une section précédente.

5.1 Paramétrisation du problème

Nous arrivons maintenant à l'objectif principale de l'étude, à savoir la calibration du modèle de Cheyette sur des prix observés. Nous rappelons que la calibration se fait sur la structure de volatilité du taux court $\sigma_r(t, x(t)) = \lambda(t)(a(t) + b(t)x(t))$ où les paramètres λ, a, b sont constants par morceaux. Nous noterons ces paramètres cibles ainsi:

$$\begin{aligned}\lambda(t) &= \sum_{n=1}^{N-1} \lambda_n \mathbf{1}_{[T_{n-1}, T_n]}(t) \\ a(t) &= \sum_{n=1}^{N-1} S_n(0) \mathbf{1}_{[T_{n-1}, T_n]}(t) \\ b(t) &= \sum_{n=1}^{N-1} b_n D_n \mathbf{1}_{[T_{n-1}, T_n]}(t)\end{aligned}$$

D'abord, la notation S_n désigne le taux swap associé aux dates de paiements $(T_n, T_n + 6M, \dots, T_N)$ et permet de distinguer les taux swap dont les maturités diffèrent. Ensuite, on note $D_n = \frac{\partial S_n}{\partial x}(t, 0, 0)$. Nous soulignerons le fait que a possède une forme choisie de sorte à réduire la calibration aux deux paramètres λ, b . De plus, on voit apparaître un terme supplémentaire dans l'expression de $b(t)$. Nous justifierons ces choix grâce à la formule suivante, déduite de l'expression de $b_S(t)$ dans la section 3.6.3.

$$b_{S_n}(t) \approx \frac{S_n(0)}{a(t)} \frac{b(t)}{\frac{\partial S}{\partial x}(t, \bar{x}(t), \bar{y}(t))}$$

Les choix de a et b sont donc judicieux dans le sens où l'on se ramène à un paramètre b_{S_n} , sur lequel on calibre, du même ordre de grandeur que le paramètre b à calibrer.

Ensemble des données observées

Dans l'optique de construire l'ensemble des données d'entrée, considérons que l'on a des données sous la forme de volatilités implicites Black Scholes du marché. Pour une swaption qui expire à T_n et des valeurs de strikes K_1, \dots, K_M , on a le smile de volatilité implicite

$$\hat{\sigma}(T_n, K_m), \quad n = 1, N - 1 \quad m = 1, M.$$

En fait, les données forment ce que l'on appelle un cube de volatilité de marché, soit les volatilités implicites calculées pour un tenor donné, une maturité donnée et un strike donné.

Dans un premier temps, nous allons chercher à déterminer les paramètres $(\hat{\lambda}_n, \hat{b}_n)_{n=1, N-1}$ tels que le n -ième prix de swaption observé coïncide avec le prix d'une swaption dans le modèle displaced log-normal de paramètres $(\hat{\lambda}_n, \hat{b}_n)$:

$$dS_n(t) = \hat{\lambda}_n(\hat{b}_n S_n(t) + (1 - \hat{b}_n)S_n(0)) dW_t$$

Nous voulons de ce fait les paramètres empiriques $(\hat{\lambda}, \hat{b})$ tels que

$$\forall n, (\hat{\lambda}_n, \hat{b}_n) = \underset{(\lambda, b) \in \mathbb{R} \times \mathbb{R}^*}{\operatorname{argmin}} \sum_{m=1}^M \left(\operatorname{Swaption}^{BS-1}(\operatorname{Swaption}^{LD}(\lambda, b, T_n, K_m), T_n, K_m) - \hat{\sigma}(T_n, K_m) \right)^2$$

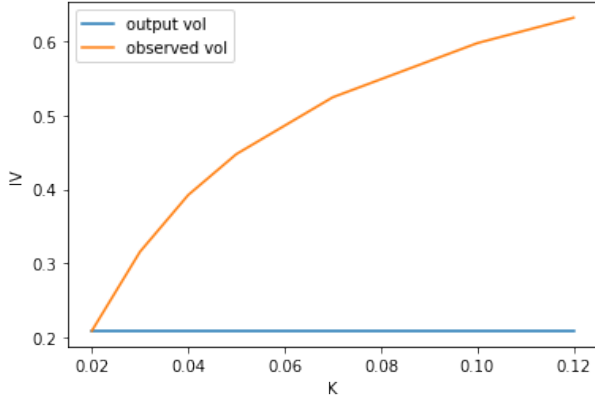
$\operatorname{Swaption}^{BS-1}(\Pi, T, K)$ est la volatilité implicite calculée dans le modèle de Black Scholes à partir d'un prix Π . Pour ce calcul, nous réaliserons à chaque itération n une étape d'optimisation par la méthode des moindres carrés. Une solution plus simple et discutable est de prendre pour chaque swaption $(\hat{\lambda}_n, \hat{b}_n) = (\text{IV ATM}, \text{IV slope ATM})$. Ce couple constitue d'ailleurs un très bon choix pour le point de départ de l'algorithme de minimisation. En revanche, si les données ne sont pas assez fournies, le calcul de la pente de volatilité implicite à la monnaie peut être grossier, c'est pourquoi nous fixerons le point de départ pour b à 1, ce qui correspond au modèle de Black Scholes. Ce point de départ permet de 'matcher' la courbe de volatilité à la monnaie.

Calibration sur les volatilités implicites au lieu des prix

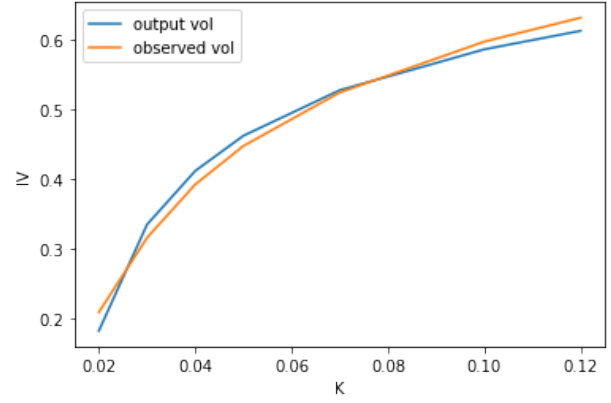
La construction des données d'entrée via la minimisation ci-dessus est réalisable sur les volatilités implicites mais également sur les prix des swaptions. On pourrait effectivement employer la minimisation suivante:

$$\forall n, (\hat{\lambda}_n, \hat{b}_n) = \underset{(\lambda, b) \in \mathbb{R} \times \mathbb{R}^*}{\operatorname{argmin}} \sum_{m=1}^M \left(\operatorname{Swaption}^{LD}(\lambda, b, T_n, K_m) - \operatorname{Swaption}^{BS}(T_n, K_m, \hat{\sigma}(T_n, K_m)) \right)^2$$

Néanmoins, nous savons que les prix de swaptions peuvent atteindre des niveaux que l'on considérera proche de 0, de l'ordre de 10^{-4} par exemple. Ainsi sur ces échelles de prix au niveau très bas, l'algorithme de minimisation des moindres carrés perd son efficacité puisqu'il atteint après une ou quelques itérations son seuil de tolérance. Il peut donc renvoyer dans certains cas le point de départ et la courbe de volatilité implicite générée n'est autre qu'une droite horizontale d'ordonnée la volatilité implicite observée à la monnaie. La méthode des moindres carrés sur les volatilités implicites, dont les niveaux sont plus hauts que les prix, apparaît donc comme un moyen efficace de palier ce problème numérique.



(a) Minimisation sur les prix



(b) Minimisation sur les volatilités implicites

Figure 19: Choix de la méthode des moindres carrés

Remarque: Il est concevable de se demander si la cette procédure de construction des données sous cette forme de paramètres d'un modèle displaced log-normal est nécessaire. Nous pourrions en effet calibrer le modèle de Cheyette, par exemple, en minimisant directement l'erreur entre des volatilités implicites calculées avec la méthode de Monte Carlo et les volatilités implicites observées. Cependant cette façon de faire impliquerait M calcul de prix de swaption (pour l'étape de minimisation sur les strikes) et cela pour chaque maturité T_n . Cela deviendrait de ce fait extrêmement coûteux en temps de calcul. C'est donc à cet instant que l'on prend conscience de la puissance de calcul qu'apporte l'approximation des prix dans le modèle displaced log-normal. Ainsi, nous avons la possibilité de calibrer le modèle sur les paramètres log-displaced $\bar{\lambda}, \bar{b}$ et éviter tout calcul de prix de swaption.

5.2 Algorithme de calibration

En considérant ce qui précède, nous avons à faire à une structure de volatilité de la forme

$$\sigma_r(t, x, y) = \sum_{n=1}^{N-1} \lambda_n (S_n(0) + b_n D_n x) \mathbf{1}_{t \in [T_{n-1}, T_n]}, \quad D_n = \frac{\partial S_n}{\partial x}(t, 0, 0)$$

Donc $\bar{y}(t)$ se ré-écrit

$$\bar{y}(t) = \int_0^t e^{-2\chi(t-s)} \sigma_r(s, 0, 0)^2 ds = \frac{1}{2\chi} \sum_{n=1}^{N-1} \lambda_n^2 S_n(0)^2 (e^{-2\chi(t \wedge T_n)} - e^{-2\chi(t \wedge T_{n-1})})$$

Voici une version de l'algorithme en langage naturel:

Posons $(\hat{\lambda}_n, \hat{b}_n)_{n=1, N-1}$ un ensemble de paramètres observés.

Pour $n = 1, \dots, N - 1$:

- Deviner une valeur pour (λ_n, b_n)
- Calculer $\bar{x}(t), \bar{y}(t)$ pour $t \in [0, T_n]$

- Calculer $\bar{\lambda}_{S_n}$ and \bar{b}_{S_n} avec les paramètres calibrés $(\lambda_i, b_i)_{i=1, n-1}$ et les paramètres non calibrés (λ_n, b_n)
- Ajuster (λ_n, b_n) tel que $\|(\bar{\lambda}_{S_n}, \bar{b}_{S_n}) - (\hat{\lambda}_n, \hat{b}_n)\| \leq \epsilon$ ou $(\bar{\lambda}_{S_n} - \hat{\lambda}_n)^2 \leq \frac{\epsilon^2}{2}$ & $(\bar{b}_{S_n} - \hat{b}_n)^2 \leq \frac{\epsilon^2}{2}$

En pratique, on va directement implémenter une recherche de racine ou une minimisation de l'erreur. En effet, on note

$$\begin{aligned}\psi : ((\lambda_i)_{i=1, n-1}, (b_i)_{i=1, n-1}, \lambda, b, T_n) &\rightarrow \bar{\lambda}_{S_n} \\ \phi : ((\lambda_i)_{i=1, n-1}, (b_i)_{i=1, n-1}, \lambda, b, T_n) &\rightarrow \bar{b}_{S_n}\end{aligned}$$

On va minimiser la fonction

$$(\lambda, b) \rightarrow \left\| \left(\psi((\lambda_i)_{i=1, n-1}, (b_i)_{i=1, n-1}, \lambda, b, T_n), \phi((\lambda_i)_{i=1, n-1}, (b_i)_{i=1, n-1}, \lambda, b, T_n) \right) - (\hat{\lambda}_n, \hat{b}_n) \right\|^2$$

Variante exploitant une relation à la monnaie

Nous nous sommes intéressés à un moyen de simplifier l'étape de minimisation de sorte à optimiser les paramètres du modèles de Cheyette sur la fonction \bar{b} uniquement. Pour cela, il convient de trouver une relation entre les deux variables $\hat{\lambda}$ et \hat{b} . Il est effectivement possible de trouver une relation explicite entre $\hat{\lambda}$ et \hat{b} qui est exacte à la monnaie mais pas pour tous les strikes. Soit T_0 fixée. On suppose $S_0 = K$ et on note $\hat{\sigma}$ la volatilité implicite observée à la monnaie.

D'une part,

$$\begin{aligned}V_0^{BS}(S_0, S_0, T_0, \hat{\sigma}) &= A(0)S_0 \left(\Phi\left(\frac{\hat{\sigma}\sqrt{T_0}}{2}\right) - \Phi\left(\frac{-\hat{\sigma}\sqrt{T_0}}{2}\right) \right) \\ &= A(0)S_0 \left(2\Phi\left(\frac{\hat{\sigma}\sqrt{T_0}}{2}\right) - 1 \right)\end{aligned}$$

D'autre part,

$$\begin{aligned}V_0^{LD}(S_0, S_0, T_0, \hat{\lambda}, \hat{b}) &= A(0)\frac{S_0}{b} \left(\Phi\left(\frac{\hat{\lambda}\hat{b}\sqrt{T_0}}{2}\right) - \Phi\left(\frac{-\hat{\lambda}\hat{b}\sqrt{T_0}}{2}\right) \right) \\ &= A(0)\frac{S_0}{b} \left(2\Phi\left(\frac{\hat{\lambda}\hat{b}\sqrt{T_0}}{2}\right) - 1 \right)\end{aligned}$$

Donc

$$\begin{aligned}V_0^{LD}(S_0, S_0, \hat{\lambda}, \hat{b}) &= V_0^{BS}(S_0, S_0, \hat{\sigma}) \\ \iff \frac{1}{b} \left(2\Phi\left(\frac{\hat{\lambda}\hat{b}\sqrt{T_0}}{2}\right) - 1 \right) &= \left(2\Phi\left(\frac{\hat{\sigma}\sqrt{T_0}}{2}\right) - 1 \right) \\ \iff \Phi\left(\frac{\hat{\lambda}\hat{b}\sqrt{T_0}}{2}\right) &= \hat{b}\Phi\left(\frac{\hat{\sigma}\sqrt{T_0}}{2}\right) + \frac{1}{2}(1 - \hat{b})\end{aligned}$$

Et on a finalement

$$\boxed{\hat{\lambda} = \frac{2}{b\sqrt{T_0}}\Phi^{-1}\left(b\Phi\left(\frac{\hat{\sigma}\sqrt{T_0}}{2}\right) + \frac{1}{2}(1 - b)\right) = \psi(T_0, b, \hat{\sigma})}$$

On se propose alors de construire un data set $(\hat{b}_i)_i$ en appliquant la méthode des moindres carrés aux quantités suivantes

$$\forall K_i, \quad Swaption^{BS-1} \left(Swaption^{LD} \left(S_0, K_i, T_n, \psi(T_n, \hat{b}, \hat{\sigma}), \hat{b} \right), T_n, K_i \right) - \hat{\sigma}(T_n, K_i)$$

Maintenant que les paramètres d'entrée sont disponibles, il reste à calibrer les paramètres de volatilité du modèle de Cheyette $(\lambda_n, b_n)_{n=1, N-1}$ de telle sorte que

$$\forall n \in [1, N-1], \quad (\lambda_n, b_n) = \operatorname{argmin}_{\lambda, b} (\bar{b}(\lambda, b) - \hat{b}_n)^2$$

Une fois la calibration effectuée, nous comparons la courbe de volatilité implicite observée et celle de la volatilité implicite dans l'approximation displaced log-normal avec les paramètres \bar{b} et $\bar{\lambda}$ tels que

$$\bar{\lambda} = \psi(T_n, \hat{\sigma}_{ATM}, \bar{b}((\lambda_i, b_i)_{i=1, n})).$$

Nous remarquons que cette méthode n'est efficace que dans une région de strike proche de la monnaie donc nous n'exploiterons pas cette variante dans notre étude.

5.3 Tests et résultats

Tests

Dans l'optique de tester le bon fonctionnement de notre algorithme de calibration, nous avons réalisé une batterie de tests dont le principe va être décrit ci-dessous:

Entrée: Nous allons générer des données d'entrée sous la forme de prix de swaptions dans un modèle displaced log-normal avec des paramètres choisis et fixés $(\lambda_i, b_i)_{i=1, \dots, N-1}$. Le test porte donc sur $N-1$ swaptions de maturités T_1, \dots, T_{N-1} où $T_i = iY$ avec une même date terminale T_N . Ces prix sont calculés pour chaque maturité T_i sur l'intervalle de strikes $[50\% S_0, 60\% S_0, \dots, 150\% S_0]$. Une fois les prix calculés, on applique notre méthode d'inversion afin d'avoir les volatilités implicites dans le modèle de Black Scholes correspondantes (courbe 'LD(λ, b)').

Ensuite, nous allons suivre les étapes naturelles de la calibration dans notre algorithme.

1. Premièrement, il s'agit de construire les paramètres $(\hat{\lambda}_i, \hat{b}_i)_{i=1, \dots, N-1}$ sur lesquels on va calibrer notre modèle.
2. Deuxièmement, la calibration avec pour arguments $(\hat{\lambda}_i, \hat{b}_i)_{i=1, \dots, N-1}$ fournit les paramètres calibrés du modèle de Cheyette $(\lambda_{Chey}^{(i)}, b_{Chey}^{(i)})_{i=1, \dots, N-1}$.
3. Troisièmement, nous générons des trajectoires pour x, y avec les paramètres calibrés $(\lambda_{Chey}^{(i)}, b_{Chey}^{(i)})_{i=1, \dots, N-1}$ afin de calculer les prix ainsi que les volatilités implicites des swaptions par la méthode de Monte Carlo (courbe 'MC').

4. Enfin, nous allons produire une troisième courbe de volatilité implicite, celle calculée à partir des prix du modèle displaced log-normal de paramètres $\bar{\lambda}((\lambda_{Chey}^{(i)}, b_{Chey}^{(i)}))$ et $\bar{b}((\lambda_{Chey}^{(i)}, b_{Chey}^{(i)}))$ (courbe 'LD(lbar,bbar)').

Cette démarche nous permet, d'une part, de juger de l'efficacité de l'étape de calibration sur des données $\hat{\lambda}, \hat{b}$ produisant une courbe de volatilité implicite quasiment égale en tout point à celle que l'on a en entrée. Effectivement, la première étape qui consiste à appliquer une méthode des moindres carrés sur les données d'entrée devrait être très efficace pour retrouver les paramètres log-displaced de départ $(\lambda_i, b_i)_{i=1, \dots, N-1}$. D'autre part, la construction des trois courbes de volatilité implicite nous permet de visualiser graphiquement la précision de la calibration en elle-même grâce à l'écart entre la courbe d'entrée 'LD(λ, b)' et la courbe calibrée 'LD(lbar,bbar)'. De même, nous pouvons observer la capacité de notre algorithme à reproduire avec le modèle de Cheyette la courbe de volatilité implicite d'entrée. Cette observation sera faite à l'aide de l'écart entre la courbe d'entrée 'LD(λ, b)' et la courbe Monte Carlo 'MC'.

Résultats

Pour commencer, nous pouvons signaler que les strikes doivent être réduits pour les tests effectués sur des courbes de prix d'entrée générées dans un modèle displaced log-normal où $\lambda \leq 0.125$. Cela provient toujours des problèmes d'inversion de la volatilité implicite cités dans la section 4. Voilà ce que l'on peut observer pour un test sur une seule maturité et des paramètres d'entrée $(\lambda, b) = (0.1, 0.9)$:

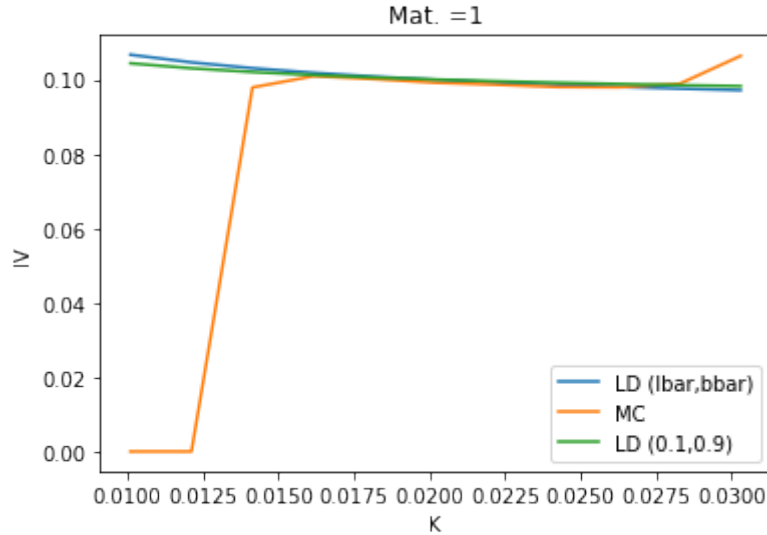


Figure 20: Test sur une maturité pour $(\lambda, b) = (0.1, 0.9)$

```
Diff ATM (LD(1,b) - LD(lbar,bbar)) -7.381629984593019e-05
Diff ATM (MC - LD(1,b)) = -0.000860760019349982
Diff. max |LD(1,b) - LD(lbar,bbar)| 0.0023058056831359863
Diff. max |MC - LD(1,b)| = 0.10436177253723145
```

On note que la calibration fournit un rendu acceptable mais la méthode des moindres carrés peine à retrouver $b = 0.9$ et renvoie plutôt $\hat{b} = 0.84$.

Nous allons maintenant présenter les résultats du test décrit ci-dessus appliqué au cas à 3 maturités $T_1 = 1Y, T_2 = 2Y, T_3 = 3Y$ et une date terminale $T_4 = 6Y$. Voici ce que nous obtenons pour les paramètres d'entrée $(\lambda_1, \lambda_2, \lambda_3) = (0.6, 0.7, 0.8)$ et $(b_1, b_2, b_3) = (0.5, 0.5, 0.5)$:

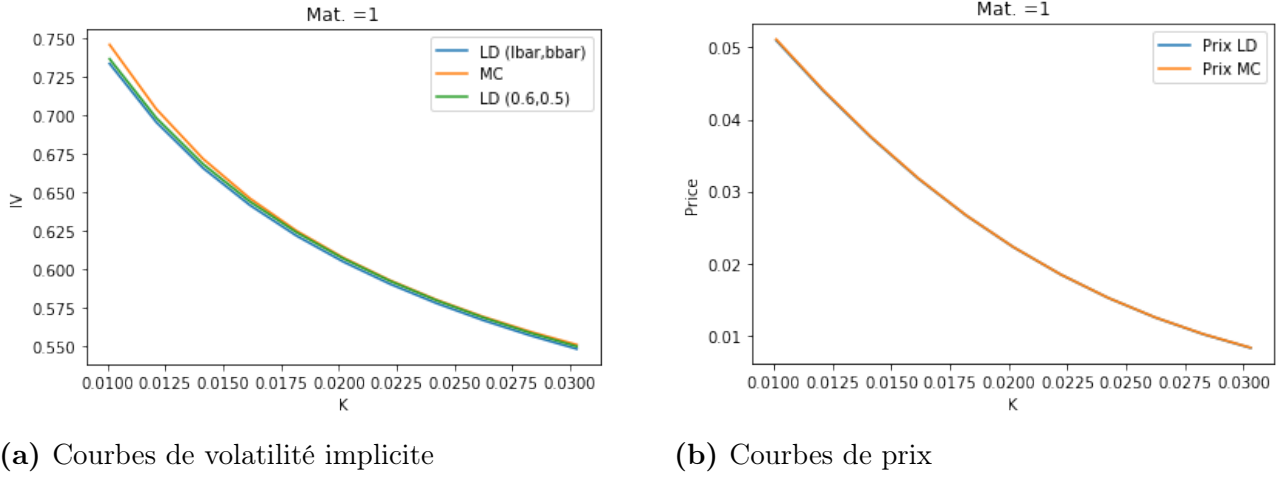


Figure 21: Représentation graphique pour T_1

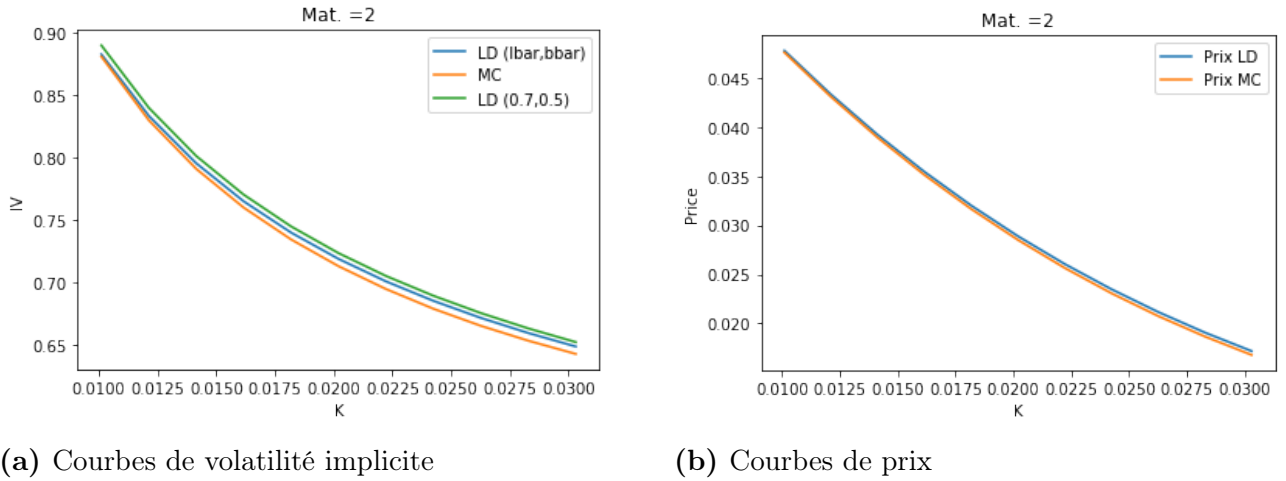
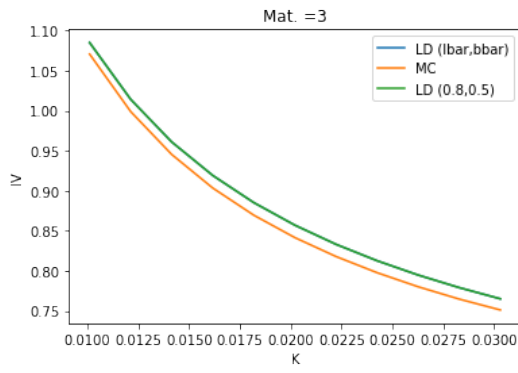
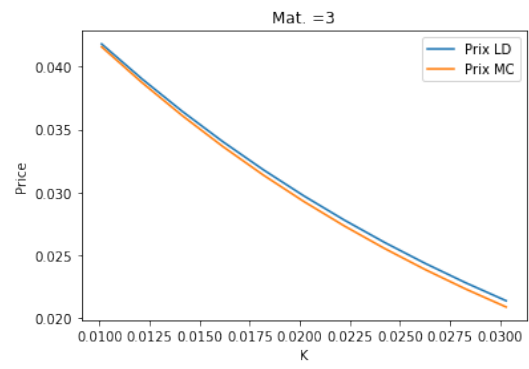


Figure 22: Représentation graphique pour T_2



(a) Courbes de volatilité implicite



(b) Courbes de prix

Figure 23: Représentation graphique pour T_3

Nous remarquons un phénomène particulier d'amplification de l'erreur entre la courbe Monte Carlo et la courbe d'entrée lorsque la maturité augmente. Nous pensions d'abord à une erreur provenant de l'algorithme de calibration en ce sens que le fait de calibrer uniquement sur $\lambda_{Chey}^{(3)}, b_{Chey}^{(3)}$ avec $\lambda_{Chey}^{(1)}, b_{Chey}^{(1)}, \lambda_{Chey}^{(2)}, b_{Chey}^{(2)}$ fixés pour la maturité T_3 ajouterait des contraintes supplémentaires à l'optimisation. En effet, on peut écrire $\bar{\lambda}$ comme :

$$\bar{\lambda}_n = \left(\frac{1}{T_n} \left(\psi(\lambda_1, \dots, \lambda_{n-1}, b_1, \dots, b_{n-1}) + \int_{T_{n-1}}^{T_n} \lambda_{S_n}(t)^2 dt \right) \right)^{\frac{1}{2}}$$

et l'on voit qu'un terme est fixé à l'étape n , ce qui pourrait limiter l'efficacité de l'optimisation si le niveau de la cible $\hat{\lambda}_n$ n'est pas atteignable. Néanmoins, nous voyons que la courbe 'LD(lbar,bbar)' est calibrée de manière satisfaisante puisqu'elle se superpose avec la courbe d'entrée pour toutes les maturités. De plus, nous avons réalisé ce test pour différentes configurations des paramètres d'entrée : λ croissant, décroissant à b fixé et b croissant et décroissant à λ fixé. Voici les tableaux résumant les erreurs résultant de ces tests (avec l'erreur relative en pourcentage $Erreur_{max}/IV_{ATM} * 100$).

| | T_1 | T_2 | T_3 |
|-----------------------|---------------|---------------|---------------|
| λ croissant | 0.0079 (1.3%) | 0.0123 (1.7%) | 0.0219 (2.6%) |
| λ décroissant | 0.0080 (0.9%) | 0.0139 (1.9%) | 0.0212 (3.5%) |
| b croissant | 0.0121 (2.4%) | 0.0180 (3.6%) | 0.0231 (4.6%) |
| b décroissant | 0.0088 (1.7%) | 0.0131 (2.6%) | 0.0134 (2.7%) |

$$Max|MC - LD(l, b)|$$

| | T_1 | T_2 | T_3 |
|-----------------------|----------------|---------------|--------------|
| λ croissant | 0.0031 (0.5%) | 0.0072 (1%) | 6.72e-6 (0%) |
| λ décroissant | 0.0077 (0.9%) | 0.0063 (0.9%) | 2.2e-6 (0%) |
| b croissant | 0.0017 (0.3%) | 0.0031 (0.6%) | 2.38e-6 (0%) |
| b décroissant | 0.0033 (0.66%) | 0.0031 (0.6%) | 1.71e-6 (0%) |

$$Max|LD(\bar{\lambda}, \bar{b}) - LD(\lambda, b)|$$

Nous avons ainsi écarté cette hypothèse pour nous focaliser sur l'effet que la maturité peut avoir sur l'approximation du modèle de Cheyette par un modèle displaced log-normal. L'objectif est de vérifier si l'écart entre les volatilités implicites générées par la méthode de Monte Carlo et celles générées par l'approximation dans un modèle displaced log-normal augmente avec la maturité (en fait lorsque l'on réduit le tenor). Pour corroborer notre hypothèse, il suffit de tracer la surface de la différence de volatilité implicite entre le modèle de Cheyette et son approximation avec les paramètres de Cheyette calibrés dans le test présenté plus haut (Figure 23), ce qui nous donne:

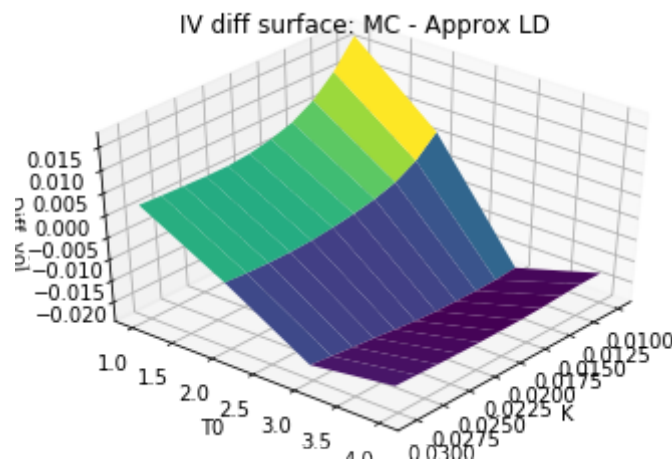


Figure 24: Différence de volatilité implicite entre le modèle de Cheyette et son approximation

La figure 24 est pertinente puisqu'elle nous permet de visualiser la croissance de la différence en valeur absolue et notamment une différence avoisinant les -0.02 pour une maturité $T = 3Y$. Cela explique donc l'écart maximal de 0.0219 que l'on observe sur (Figure 23). Cela nous amène donc à la conclusion que l'algorithme de calibration fournit une précision à la hauteur de nos attentes. Cependant, les écarts observés entre les courbes d'entrée et les courbes générées par la méthode de Monte Carlo après calibration sont intrinsèquement liés à la méthode de calibration choisie et proviennent essentiellement de l'approximation des prix dans un modèle displaced log-normal dont la précision s'épuise avec la maturité.

Pour finir, voici le tableau des temps d'exécution de l'algorithme de calibration pour 1,2 et 3 maturités:

| Nombre de maturités | Temps (sec) |
|---------------------|-------------|
| 1 | 408.5 |
| 2 | 1 912.3 |
| 3 | 4 523.9 |

6 Conclusion

Pour conclure notre étude, nous nous proposons de résumer le travail accompli ainsi que les discussions mis en lumière tout au long de ce document. Notre développement orienté autour des trois axes de recherche que sont le pricing, le calcul de la volatilité implicite et la calibration nous a permis d'exposer au lecteur le cadre particulier du modèle de Cheyette unidimensionnel à volatilité locale paramétrique, mais également les difficultés rencontrées sur les plans numérique, mathématique et financier.

Les principaux enjeux relevés furent le fait de justifier l'utilisation du modèle de Cheyette dans un cadre financier, de l'implémenter et enfin de le tester afin de définir le meilleur moyen d'en faire usage. Ces enjeux nous ont conduit à la compréhension des rouages du modèle, entre autres l'effet des paramètres de volatilité sur la courbe de volatilité implicite. Nous avons tenté d'apporter à cette recherche une démarche mathématique et déductive rigoureuse, cependant la plupart des preuves fournies sont restreintes à des modèles plus simplistes que le modèle de Cheyette, notamment le modèle displaced log-normal. Si ce modèle displaced log-normal s'est avéré très utile pour interpréter nos observations et d'un point de vue pratique pour approcher le prix d'une swaption européenne dans le modèle de Cheyette, il n'en reste pas moins qu'il a ses limites. Ces dernières portent essentiellement sur l'imprécision de l'approximation des prix des swaptions lorsque leur tenor décroît, ce qui a été mis en exergue grâce à nos tests de calibration. Par ailleurs, ces tests ont été effectués sur des données que nous avons générées. Cette façon de procéder a un sens puisqu'elle nous a permis de juger de la capacité de notre algorithme de calibration à reproduire une courbe de volatilité implicite connue. Néanmoins, il est regrettable de ne pas avoir eu la possibilité de calibrer notre modèle sur des prix de marché réels.

En guise d'ouverture, nous invitons le lecteur à s'intéresser aux multiples extensions du modèle de Cheyette unidimensionnel à volatilité locale paramétrique. (Nous noterons que la suite logique de ce travail est la calibration du paramètre de modèle χ . Elle nécessite des considérations et un algorithme à part, ce qui n'a pas été traité ici). D'une part, il est possible d'étendre ce modèle de manière naturelle à sa version multidimensionnelle à volatilité locale paramétrique, ce que nous allons présenter brièvement dans l'annexe 11. D'autre part, il semble également pertinent de se concentrer sur la version du modèle à volatilité stochastique. Dans ce cas, il s'agirait d'ajouter à la structure de volatilité σ_r un facteur stochastique $z(t)$ qui suit la dynamique CIR: $dz(t) = \theta(z(0) - z(t))dt + \nu(t)\sqrt{z(t)}dW_t$. Pour finir, le lecteur peut s'orienter vers une structure de volatilité non paramétrique et la calibrer à l'aide des méthodes de Dupire ou des méthodes de particules.

References

- [1] Piterbarg Andersen. *Interest rate modeling*. Vol. 1. 2010.
- [2] Piterbarg Andersen. *Interest rate modeling*. Vol. 2. 2010.
- [3] F Mercurio D Brigo. *Interest rate models-Theory and Practice. With Smile, Inflation and Credit*. 2006.
- [4] E.Chevalier.et.al. “Optimal exit strategies for investment projects”. In: *J. Math. Anal. Appl.* (2015).
- [5] J. Healy. “Equivalence between forward rate interpolations and discount factor interpolations for the yield curve construction”. In: *arXiv preprint arXiv:2005.13890* (2019).
- [6] O.Cheyette. “Markov Representation of the Heath-Jarrow-Morton Model”. In: 1 (1995).
- [7] Vladimir Piterbarg. “Markovian Projection Method for Volatility Calibration”. In: (2006).
- [8] Dan Wang Roger Lee. “Displaced Lognormal Volatility Skews: Analysis and Applications to Stochastic Volatility Simulations”. In: (2009).

Annexe

1. Le Lemme d'Itô

On note M_{loc}^2 l'ensemble des fonctions progressivement mesurables ϕ de $\mathbb{R}^+ \times \Omega$ telles que

$$\int_0^T \phi(s, \omega)^2 ds < \infty \quad p.s., \forall T.$$

Soit Φ de classe $\mathcal{C}^{1,2}$ sur $\mathbb{R}^+ \times \mathbb{R}^d$, W un mouvement Brownien de \mathbb{R}^d .

Si $dX_t = \mu(t, X_t) dt + \sigma(t, X_t) dW_t$ où μ, σ sont dans M_{loc}^2 et prennent leurs valeurs respectivement dans \mathbb{R}^d et $\mathbb{R}^{d \times k}$. Alors on a :

$$\begin{aligned} d\Phi(t, X_t) &= \frac{\partial \Phi}{\partial t}(t, X_t) dt + \nabla \Phi(t, X_t) \cdot dX_t + \frac{1}{2} tr(\nabla^2 \Phi \sigma \sigma^\top) dt \\ &= \frac{\partial \Phi}{\partial t}(t, X_t) dt + \sum_{i=1}^d \frac{\partial \Phi}{\partial x_i}(\mu_i(t, X_t) dt + \sigma_i(t, X_t) \cdot dW_t) + \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 \Phi}{\partial x_i \partial x_j}(t, X_t) d\langle X_i, X_j \rangle_t \end{aligned}$$

Remarque: La condition de bornitude des dérivées de Φ par rapport à X est exigée pour l'existence de l'écriture sous forme intégrale et pour la propriété de martingale de l'intégrale stochastique. La formule reste valable si X prend p.s ses valeurs dans un ouvert convexe $D \subset \mathbb{R}^d$ sur lequel Φ est \mathcal{C}^2 par rapport à x .

2. Intersion dérivée et différentielle

Nous allons détailler ici l'intersion des signes dérivée et différentielle effectuée dans la construction du modèle HJM. Nous avons la dynamique suivante:

$$d \ln(P(t, T)) = \mu(t, T) dt + \sigma(t, T) dW_t.$$

On veut modéliser $t \rightarrow P(t, T)$ de manière continue, donc on suppose le processus $(P(t, T))_t$ à trajectoires continues pour tout $T > 0$. De plus, on considère que P prend ses valeurs dans un ouvert de \mathbb{R} tel que $\forall t \leq T, P(t, T) \in]\epsilon, P_\infty]$, ouvert sur lequel $x \rightarrow \ln(x)$ est \mathcal{C}^2 , bornée et à dérivées bornées. Ce qui nous permet d'écrire la forme intégrale suivante:

$$\forall T > 0, \ln(P(t, T)) = \ln(P(0, T)) + \int_0^t \mu(s, T) ds + \int_0^t \sigma(s, T) dW_s$$

Si μ et σ sont dérivables par rapport à T et que $\frac{\partial \mu}{\partial T}, \frac{\partial \sigma}{\partial T}$ sont continues et dominées en valeurs absolue par une fonction intégrable par rapport à t , alors :

$$\frac{\partial \ln(P(t, T))}{\partial T} = \frac{\partial \ln(P(0, T))}{\partial T} + \int_0^t \frac{\partial \mu}{\partial T}(s, T) ds + \int_0^t \frac{\partial \sigma}{\partial T}(s, T) dW_s.$$

Dans notre cas, on a $\mu(t, T) = r(t) - \frac{1}{2} \sigma_P(t, T)^\top \sigma_P(t, T)$ et $\sigma(t, T) = \sigma_P(t, T)$. Si σ_P est dérivable par rapport à T , alors $\sigma_f(t, T) = \frac{\partial \sigma_P}{\partial T}(t, T)$. Comme la forme intégrale de $P(t, T)$ est

bien définie, $\int_0^T \sigma_P(t, u) dt = \int_0^T \int_t^T \sigma_f(t, u) du dt$ est bien définie. par Fubini, on en déduit que σ_f est intégrable par rapport à t et les conditions évoquées plus haut sont satisfaites. On a donc un processus écrit sous la forme d'Itô, il vient:

$$d\left(\frac{\partial \ln(P(t, T))}{\partial T}\right) = \frac{\partial \mu}{\partial T}(t, T) dt + \frac{\partial \sigma}{\partial T}(t, T) dW_t$$

3. Le Théorème de représentation des martingales

Pour tout processus $X \in \mathbb{L}^2(\mathbb{P})$ et \mathcal{F}_T -mesurable, il existe un processus $H \in M_{loc}^2$ tel que

$$X = \mathbb{E}[X] + \int_0^T H(s)^2 dW_s$$

4. Le Théorème de Girsanov

Soit $\phi \in M_{loc}^2$ et

$$\forall t \in [0, T], \quad Z(t) \stackrel{\text{def}}{=} \exp \left\{ \int_0^t \phi(s) dW_s - \frac{1}{2} \int_0^t \|\phi(s)\|^2 ds \right\}$$

Soit le changement de probabilité défini par

$$\frac{d\mathbb{Q}}{d\mathbb{P}} \Big|_{\mathcal{F}_t} = Z(t).$$

D'après le théorème de Girsanov, on a le résultat suivant:

$$\mathbb{E}_{\mathbb{P}}[Z(T)] = 1 \implies W_t^{\mathbb{Q}} = W_t - \int_0^t \phi(s) ds \quad \text{est un mouvement Brownien sous } \mathbb{Q}$$

5. Le Théorème d'existence et d'unicité de la solution forte d'une EDS

Soit un drift $\mu : \mathbb{R}^+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ et une volatilité $\sigma : \mathbb{R}^+ \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times k}$ vérifiant l'hypothèse (\mathcal{H}) . Si X vérifie l'EDS

$$dX_t = \mu(t, X_t) dt + \sigma(t, X_t) dW_t, \quad X_0 \in \mathbb{L}^2$$

alors il existe une unique solution $X \in M^2$ à cette équation et c'est un processus de Markov.

6. Construction de la courbe Zero Coupon

La construction de la courbe zero coupon est une étape cruciale dans l'entreprise de pricing des swaptions en raison de la nécessité des prix de zero coupon à chaque maturité $\left(P(0, T_i)\right)_{i=0, N}$. Il est possible de les calculer à l'aide de l'expression

$$P(0, T_i) = e^{-\int_0^{T_i} f(0, u) \, du}$$

à condition d'avoir accès à la courbe de taux forward initiale $t \rightarrow f(0, t)$.

Dans l'optique de simplifier au maximum les calculs et d'assurer une certaine cohérence des résultats, il est possible de choisir la forme suivante $P(0, T) = e^{-rT}$, en d'autres termes on considère le taux forward initial $f(0, t)$ constant. Cette approche simpliste sera utilisée à de nombreuses reprises dans les tests lorsque les calculs requièrent la courbe zero coupon pour un nombre important de maturités. Cela évite en effet des calculs et peut s'avérer pratique lorsque l'on dispose de données de marché insuffisantes.

Présentons maintenant une méthode de bootstrapping sur le taux swap au pair, qui va nous servir à construire une courbe zero coupon plus réaliste de 0 à 50 ans avec une base semi-annuelle.

Données: Considérons que nous avons accès à des données historiques de taux swap pour N maturités $T_1 < \dots < T_N$. On peut construire à partir de cela la courbe zéro coupon en M points $(t_j)_{j=1, M}$, admettons que l'on veuille une valeur de zéro coupon tous les 6 mois, on aurait alors $t_j = 0.5 \times j$ Y.

Nous avons d'abord récupéré les données de taux swap suivantes dans (Andersen 2010a, Section 6.2):

| Maturity (Y) | Swap par rate |
|-----------------|------------------|
| 1 | 4.20% |
| 2 | 4.30% |
| 3 | 4.70% |
| 5 | 5.40% |
| 7 | 5.70% |
| 10 | 6.00% |
| 12 | 6.10 % |
| 15 | 5.90% |
| 20 | 5.60% |
| 25 | 5.55% |

Voyons à présent comment procéder.

Rendement continuellement composé linéaire par morceaux

Le rendement continuellement composé ou *Continuously compounded yield* est défini comme la fonction $y : [0, T_{max}] \rightarrow \mathbb{R}^+$ telle que

$$e^{-y(T)T} = P(0, T) = e^{-\int_0^T f(0, u) du}$$

On note que

$$f(0, T) = y(T) + \frac{dy}{dT}(T) \times T$$

Cette méthode consiste à considérer le **rendement linéaire par morceaux** sur $\{T_i\}_i$, en d'autres termes

$$\forall i \in [1, N], \quad y(T) = y(T_i) \frac{T_{i+1} - T}{T_{i+1} - T_i} + y(T_{i+1}) \frac{T - T_i}{T_{i+1} - T_i}, \quad T \in [T_i, T_{i+1}]$$

On se doit également de fixer le taux initial tel que $y(t) \equiv y(T_1)$, $t < T_1$.

Il est essentiel de rappeler d'après (6) que:

$$S(T_i) = \frac{1 - P(T_i)}{\sum_{j, t_j \leq T_i} \tau P(t_j)}$$

L'algorithme de Bootstrapping pour nos données est le suivant:

i = 1:

- On fixe le yield sur le premier intervalle $y(t) = y(T_1)$, $t \leq T_1$:

On a dans notre cas $y(0.5) = y(1)$

Pour i = 2, ..., N:

- Connaissant $P(t), t \leq T_{i-1}$, on va chercher la valeur de $P(T_i)$.

On pose $y(T_i) = x$ et $\forall t_j \in [T_{i-1}, T_i]$, $y(t_j) = y(T_{i-1}) \frac{T_i - t_j}{T_i - T_{i-1}} + x \frac{t_j - T_{i-1}}{T_i - T_{i-1}}$

- On trouve x par une méthode de recherche de racine telle que

$$S(T_i) = \frac{1 - P(T_i)}{\sum_{j=1, t_j \leq T_i} \tau P(t_j, x)}$$

Avec le tableau de données d'entrée présenté plus haut, voici les courbes de taux générées grâce à cette méthode:

Taux forward constant par morceaux

Nous aborderons dans ce paragraphe une méthode similaire de construction de la courbe zéro coupon par Bootstrapping sur le taux swap, mais cette fois-ci en considérant une forme particulière du taux forward.

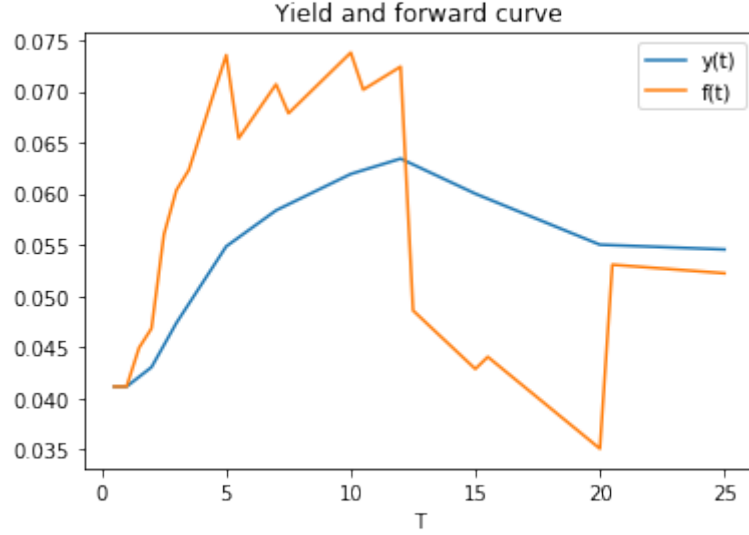


Figure 25: Courbe de taux linéaire par morceaux

Il s'agit ici d'assumer un taux forward constant par morceaux, c'est-à-dire:

$$\forall i \in [0, N], \quad f(T) \equiv f(T_i), \quad T \in [T_i, T_{i+1}[$$

On rappelle que (voir page précédente)

$$\forall T \in [0, T_N], \quad y(T) \times T = \int_0^T f(0, u) \, du$$

D'où

$$y(T) = \frac{y(T_i)T_i + f(T_i)(T - T_i)}{T}$$

et de manière équivalente

$$y(T) = \frac{1}{T} \left(T_i y(T_i) \frac{T_{i+1} - T}{T_{i+1} - T_i} + T_{i+1} y(T_{i+1}) \frac{T - T_i}{T_{i+1} - T_i} \right)$$

Algorithme de Bootstrapping:

- On initialise $y(t) \equiv y(T_1)$, $t \leq T_1$ et on déduit $P(0.5)$ et $P(1)$ à partir de $S(1)$.

Pour i= 2,...,N:

- Sachant $P(t), t \leq T_{i-1}$ on va faire déduire la valeur de $P(T_i)$.

On pose $y(T_i) = x$ et on détermine x par recherche de racine comme précédemment avec

$$\forall t_j \in [T_{i-1}, T_i], \quad y(t_j) = \frac{1}{t_j} \left(y(T_{i-1})T_{i-1} \frac{T_i - t_j}{T_i - T_{i-1}} + xT_i \frac{t_j - T_{i-1}}{T_i - T_{i-1}} \right)$$

Voici le rendu ci-dessous:

On peut ainsi construire une courbe jusqu'à l'horizon désiré, du moment que les données sont suffisamment fournies. Pour l'application de la courbe zéro coupon au pricing d'instruments

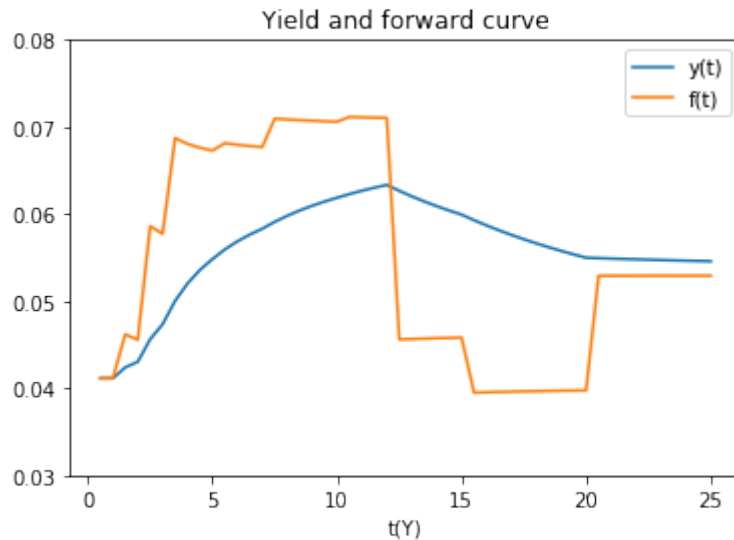


Figure 26: Courbe de taux forward constant par morceaux

financiers, il est, dans la majorité des cas, nécessaire d'aller au delà de 25 ans et 50 ans est un horizon assez lointain. Nous allons donc construire une courbe zéro coupon allant de 0 à 50 ans en utilisant des deux méthodes présentées précédemment afin de réaliser les travaux numériques par la suite.

Voici une table de données disponible dans l'article (Healy 2019) datant de Novembre 2019.

| | | |
|--------------|------------|---------|
| OIS swap 1Y | 2020/11/09 | 0.01455 |
| OIS swap 2Y | 2021/11/08 | 0.01373 |
| OIS swap 3Y | 2022/11/08 | 0.01354 |
| OIS swap 4Y | 2023/11/08 | 0.01347 |
| OIS swap 5Y | 2024/11/08 | 0.01355 |
| OIS swap 6Y | 2025/11/10 | 0.01375 |
| OIS swap 7Y | 2026/11/09 | 0.01398 |
| OIS swap 8Y | 2027/11/08 | 0.01429 |
| OIS swap 9Y | 2028/11/08 | 0.01451 |
| OIS swap 10Y | 2029/11/08 | 0.01484 |
| OIS swap 12Y | 2031/11/10 | 0.01534 |
| OIS swap 15Y | 2034/11/08 | 0.01591 |
| OIS swap 20Y | 2039/11/08 | 0.01645 |
| OIS swap 25Y | 2044/11/08 | 0.01662 |
| OIS swap 30Y | 2049/11/08 | 0.01672 |
| OIS swap 40Y | 2059/11/10 | 0.01650 |
| OIS swap 50Y | 2069/11/08 | 0.01617 |

Maintenant que deux méthodes plus élaborées ont été présentées, il est intéressant de revenir à la première forme du prix de l'obligation zéro coupon $P(0, T) = e^{-rT}$. On peut effectivement constater que l'hypothèse d'un taux forward initial constant n'est pas délirant en ce sens qu'il permet de reproduire une courbe zéro coupon proche de la réalité sur le court terme, sous réserve que la constante soit évidemment bien choisie. Ce constat peut être fait en comparant la courbe zéro coupon bootstrappée et celle induite par le taux forward initial constant.

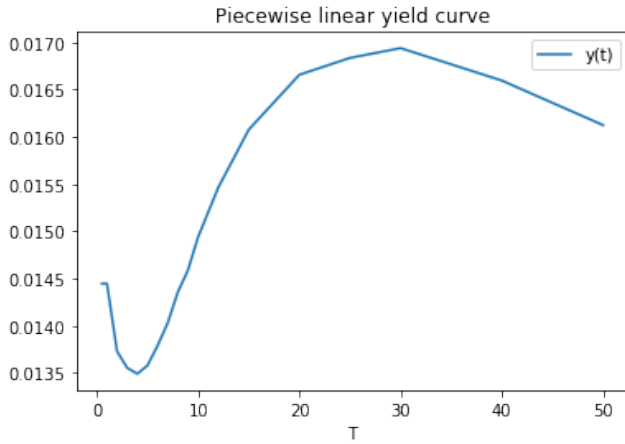


Figure 27: Yield linéaire par morc.

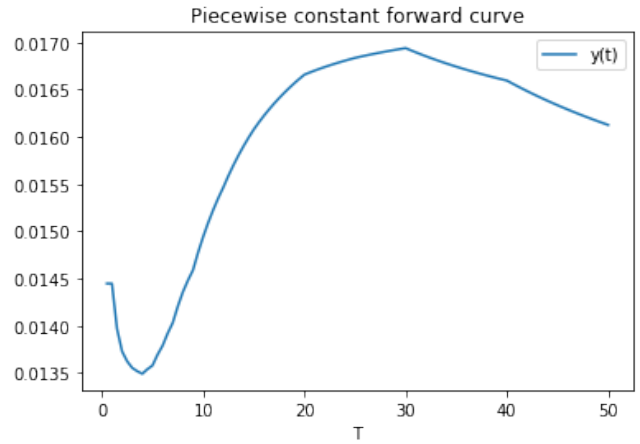
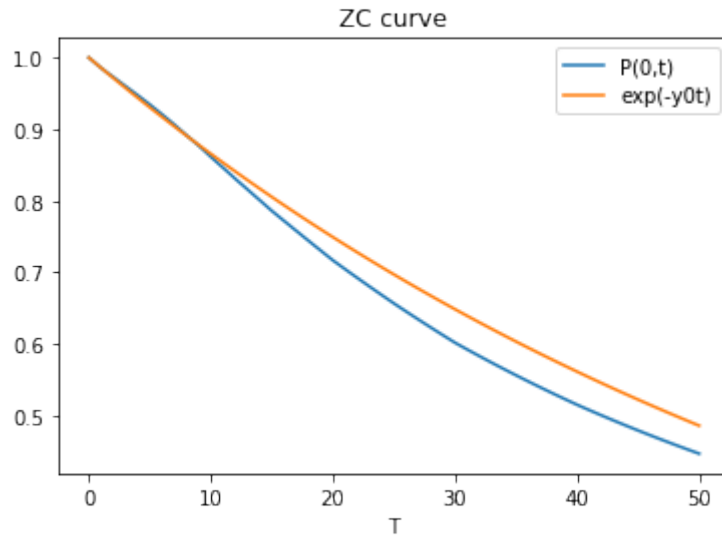


Figure 28: Forward constant par morc.



7. Différence de distribution entre probabilité risque neutre et terminale

Le payoff du swaption se note

$$(S(T_0, T_N, x(T_0), y(T_0)) - K)^+$$

. On a vu que sous cette probabilité, la dynamique de x s'écrit sous la forme

$$dx(t) = (y(t) - \chi(t)x(t)) dt + \sigma_r(t) dW_t$$

où W est un mouvement Brownien sous \mathbb{Q} .

Le payoff conserve la même expression, à un facteur indépendant de x près, sous la probabilité terminale \mathbb{Q}^{T_N} . Néanmoins, la dynamique de x peut se réécrire sous la forme

$$dx(t) = (y(t) - \sigma_r(t)^2 G(t, T_N) - \chi(t)x(t)) dt + \sigma_r(t) dW_t^{T_N}$$

où W^{T_N} est également un mouvement Brownien mais sous la probabilité terminale.

On veut montrer ici que la probabilité que le payoff s'annule est plus grande sous la mesure terminale.

On note

$$\begin{aligned}\mu_1(t) &= (y(t) - \chi(t)x(t)) \\ \mu_2(t) &= (y(t) - \sigma_r(t)^2 G(t, T_N) - \chi(t)x(t))\end{aligned}$$

et $\epsilon(t) \geq 0$ telle que $\mu_1(t) = \mu_2(t) + \epsilon(t)$.

Comme W sous \mathbb{Q} a la même loi que W^{T_N} sous \mathbb{Q}^{T_N} , on a:

$$\begin{aligned}\forall A \in \mathcal{B}(\mathbb{R}), \quad \mathbb{Q}(\int_0^{T_0} \mu_1(t)dt + \int_0^{T_0} \sigma_r(t)dW_t \in A) \\ = \mathbb{Q}^{T_N}(\int_0^{T_0} \mu_1(t)dt + \int_0^{T_0} \sigma_r(t)dW_t^{T_N} \in A) \\ = \mathbb{Q}^{T_N}(\int_0^{T_0} (\mu_2(t) + \epsilon(t))dt + \int_0^{T_0} \sigma_r(t)dW_t^{T_N} \in A)\end{aligned}$$

Hypothèse: Si on considère que ϵ est déterministe, la quantité précédente se réécrit:

$$\mathbb{Q}^{T_N}(\int_0^{T_0} \mu_2(t)dt + \int_0^{T_0} \sigma_r(t)dW_t^{T_N} \in \{x - \int_0^{T_0} \epsilon(t)dt, x \in A\})$$

On note

$$O = S(T_0, T_N, \cdot, y(T_0))^{-1}(]-\infty, K])$$

, qui est un fermé car image réciproque d'un fermé par une application continue.

On sait que S est une fonction monotone croissante et continue de x .

De plus, $\lim_{x \rightarrow \infty} S(T_0, T_N, x, y(T_0)) = +\infty$ et $\lim_{x \rightarrow -\infty} S(T_0, T_N, x, y(T_0)) = -\infty$ par croissance de $T \rightarrow G(t, T)$.

On en déduit que $O =]-\infty, S(T_0, T_N, \cdot, y(T_0))^{-1}(K)]$. Donc $O \in \mathcal{B}(\mathbb{R}) = \sigma(]-\infty, b], b \in \mathbb{R})$ et $\tilde{O} = \{x - \int_0^{T_0} \epsilon(t)dt, x \in O\} \subset O$.

Ce qui implique que

$$\begin{aligned}\mathbb{Q}^{T_N}(\int_0^{T_0} \mu_2(t)dt + \int_0^{T_0} \sigma_r(t)dW_t^{T_N} \in \{x - \int_0^{T_0} \epsilon(t)dt, x \in O\}) \\ \leq \mathbb{Q}^{T_N}(\int_0^{T_0} \mu_2(t)dt + \int_0^{T_0} \sigma_r(t)dW_t^{T_N} \in O)\end{aligned}$$

Enfin

$$\boxed{\mathbb{Q}(\int_0^{T_0} \mu_1(t)dt + \int_0^{T_0} \sigma_r(t)dW_t \in O) \leq \mathbb{Q}^{T_N}(\int_0^{T_0} \mu_2(t)dt + \int_0^{T_0} \sigma_r(t)dW_t^{T_N} \in O)}$$

Il reste à montrer le résultat si ϵ n'est pas déterministe. Si ϵ est un processus aléatoire, il n'en reste pas moins que

$$\forall (\omega, t) \in \Omega \times \mathbb{R}^+, \epsilon(\omega, t) \geq 0 \implies \forall \omega \in \Omega, \int_0^{T_0} \epsilon(\omega, t) dt \geq 0$$

Si on note $Y = \int_0^{T_0} \mu_2(t) dt + \int_0^{T_0} \sigma_r(t) dW_t^{T_N}$. On a

$$\begin{aligned} & \mathbb{Q}^{T_N} \left(Y + \int_0^{T_0} \epsilon(t) dt \in O \right) \\ &= \mathbb{Q}^{T_N} \left(\{\omega, Y(\omega) + \int_0^{T_0} \epsilon(\omega, t) dt \in O\} \right) \\ &= \mathbb{Q}^{T_N} \left(\{\omega, Y(\omega) + \int_0^{T_0} \epsilon(\omega, t) dt \in] - \infty, S(T_0, T_N, \cdot, y(T_0))^{-1}(K)]\} \right) \\ &= \mathbb{Q}^{T_N} \left(\{\omega, Y(\omega) \in] - \infty, S(T_0, T_N, \cdot, y(T_0))^{-1}(K) - \int_0^{T_0} \epsilon(\omega, t) dt]\} \right) \end{aligned}$$

Or

$$\forall \omega \in \Omega,] - \infty, S(T_0, T_N, \cdot, y(T_0))^{-1}(K) - \int_0^{T_0} \epsilon(\omega, t) dt] \subset] - \infty, S(T_0, T_N, \cdot, y(T_0))^{-1}(K)]$$

Donc

$$\begin{aligned} & \{\omega, Y(\omega) \in] - \infty, S(T_0, T_N, \cdot, y(T_0))^{-1}(K) - \int_0^{T_0} \epsilon(\omega, t) dt]\} \\ & \subset \{\omega, Y(\omega) \in] - \infty, S(T_0, T_N, \cdot, y(T_0))^{-1}(K)]\} \end{aligned}$$

Et on conclut de manière analogue au cas ϵ déterministe.

8.Preuve de la condition de Feller

Cette preuve est tirée de l'exercice 34 du livre "Introduction au calcul stochastique et à la finance" de D. Lamberton et B. Lapeyre.

Si X suit un processus CIR tel que

$$dX_t = (a - bX_t) dt + \sigma \sqrt{X_t} dW_t$$

qui commence en $x \in \mathbb{R}$, on note X^x .

Posons

$$s(x) = \int_1^x e^{y \frac{2b}{\sigma^2}} y^{-2a/\sigma^2} dy$$

et

$$\tau_M = \inf \{t \geq 0, X_t = M\}$$

On a facilement

$$\frac{\sigma^2}{2} x \frac{\partial^2 s}{\partial x^2} + (a - bx) \frac{\partial s}{\partial x} \equiv 0$$

Soit $0 < \epsilon < x < M$.

En appliquant le lemme d'Itô à $s(x_t)$ (s est \mathcal{C}^2 et bornée sur un interval $[\epsilon, M]$, $0 < \epsilon < x < M$):

$$s\left(X_{t \wedge \tau_{\epsilon, M}^x}^x\right) = s(x) + \int_0^{t \wedge \tau_{\epsilon, M}^x} s'(X_s^x) \sigma \sqrt{X_s^x} dW_s$$

On veut calculer la variance dans cette dernière équation. Comme X est à trajectoires continues et que $x \rightarrow s'(x)^2 x$ est également continue sur \mathbb{R} , cette fonction est bornée sur un segment $[\epsilon, M]$. Donc $t \rightarrow \int_0^t s'(X_s^x) \sigma \sqrt{X_s^x} dW_s$ est une martingale et

$$\mathbb{V}\left[s(X_{t \wedge \tau_{\epsilon, M}^x})\right] = \mathbb{E}\left[\int_0^{t \wedge \tau_{\epsilon, M}^x} s'(X_s^x)^2 \sigma^2 X_s^x ds\right]$$

De plus,

$$\frac{\partial s}{\partial x} = e^{x \frac{2b}{\sigma^2}} x^{-\frac{2a}{\sigma^2}} \geq 0 \quad \text{on } x \geq 0$$

Donc $s(\epsilon) \leq s(X_{t \wedge \tau_{\epsilon, M}^x}) \leq s(M)$

D'où

$$\mathbb{V}\left[s(X_{t \wedge \tau_{\epsilon, M}^x})\right] \leq s(M)^2$$

et

$$\mathbb{E}\left[\int_0^{t \wedge \tau_{\epsilon, M}^x} s'(X_s^x)^2 \sigma^2 X_s^x ds\right] \geq \sigma^2 (s')_{inf}^2 \epsilon \quad \mathbb{E}[t \wedge \tau_{\epsilon, M}^x]$$

où $(s')_{inf} = \inf \{s'(x), x \in [\epsilon, M]\}$

Par le théorème de convergence monotone, il vient $\mathbb{E}[t \wedge \tau_{\epsilon, M}^x] \rightarrow_{t \rightarrow \infty} \mathbb{E}[\tau_{\epsilon, M}^x]$. One has that

$$\mathbb{E}[\tau_{\epsilon, M}^x] < \infty$$

which implies that $\tau_{\epsilon, M}^x$ is finite almost surely.

Une fois que l'on sait que le temps d'arrêt $\tau_{\epsilon, M}^x$ est fini presque sûrement, on peut appliquer le théorème d'arrêt pour les martingales:

$$\begin{aligned} s(x) &= s(\epsilon) \mathbf{P}(\tau_{\epsilon}^x < \tau_M^x) + s(M) \mathbf{P}(\tau_{\epsilon}^x > \tau_M^x) \\ &= \mathbf{P}(\tau_{\epsilon}^x < \tau_M^x) (s(\epsilon) - s(M)) + s(M) \quad \square \end{aligned}$$

Si $a \geq \frac{\sigma^2}{2}$ alors $\lim_{\epsilon \rightarrow 0} s(\epsilon) = -\infty$. En effet

$$|s(\epsilon)| = \left| \int_{\epsilon}^1 e^{y \frac{2b}{\sigma^2}} y^{-\frac{2a}{\sigma^2}} dy \right| \geq \int_{\epsilon}^1 y^{-\frac{2a}{\sigma^2}} dy$$

Le terme de droite est une intégrale de Riemann qui converge vers la quantité suivante lorsque ϵ tend vers 0:

$$\int_0^1 y^{-\frac{2a}{\sigma^2}} dy = \begin{cases} \frac{1}{1-\frac{2a}{\sigma^2}}, & \text{if } \frac{2a}{\sigma^2} < 1. \\ +\infty, & \text{otherwise.} \end{cases}$$

En prenant la limite lorsque ϵ tend vers 0 dans \square , on a nécessairement $\mathbf{P}(\tau_0^x < \tau_M^x) = 0$ (\star)

On a montré que $\forall M > 0, X$ visits M avant 0.

Finalement, la diffusion X^x a des trajectoires continues, donc $(\tau_M)_M$ est une suite strictement croissante de temps d'arrêt qui diverges en $+\infty$. En effet, pour chaque événement ω de l'univers, $t \rightarrow X_t^x(\omega)$ est une fonction continue. Cela implique que cette fonction est bornée sur tout horizon de temps fini $[0, \tau]$ de telle sorte que τ_∞ ne peut pas être fini presque sûrement, sinon le processus atteindrait toutes les valeurs réelles positives avant τ_∞ .

En prenant la limite quand $M \rightarrow +\infty$ dans (\star) ,

$$\mathbf{P}(\tau_0^x < \infty) = 0$$

9. Prix de Call dans un modèle displaced log-normal à coefficients constants

Soit $\lambda > 0, b \in [0, 1]$ On considère la dynamique de taux swap sous l'*annuity measure* \mathbb{Q}^A suivante:

$$\begin{aligned} dS(t) &= \sigma(bS(t) + (1-b)S(0)) dW^A(t) \\ &= \sigma(bS(t) + \beta) dW^A(t) \end{aligned}$$

Si $U(t) = bS(t) + \beta$, on a

$$dU(t) = b dS(t) = b\sigma U(t) dW^A(t)$$

d'où

$$U(T) = U(t)e^{b\sigma(W_T^A - W_t^A) - \frac{b^2\sigma^2}{2}(T-t)}$$

De manière équivalente,

$$S(T) = -\frac{(1-b)}{b}S(0) + \frac{S(0)}{b}e^{b\sigma W_T^A - \frac{b^2\sigma^2}{2}T}$$

Dans ce cadre,

$$V_{\text{swaption}}(t) = A(t)\mathbb{E}\left[\left(S(T_0) - K\right)^+ \middle| \mathcal{F}_t\right]$$

Par changement de variable et la formule de Black Scholes, si $S(0) + \frac{1-b}{b}S(0) > 0$ et $K + \frac{1-b}{b}S(0) > 0$ il vient:

$$\boxed{\begin{aligned} V_{swaption}(t) &= A(t) \left(\left(S(t) + \frac{1-b}{b} S(0) \right) \Phi(d^+) - \left(K + \frac{1-b}{b} S(0) \right) \Phi(d^-) \right) \\ d^{+/-} &= \frac{\ln \left(\frac{bS(t) + (1-b)S(0)}{bK + (1-b)S(0)} \right) + /- \frac{\sigma^2 b^2}{2} (T_0 - t)}{\sigma b \sqrt{T_0 - t}} \end{aligned}}$$

Cas:

- Si $S(0) - K < \frac{S(0)}{b} < 0$: $V_0 = 0$
- Si $0 < \frac{S(0)}{b} < S(0) - K$: $V_0 = A(0)(S(0) - K)$
- Si $\frac{S(0)}{b} < 0$ et $K + \frac{1-b}{b} S(0) < 0$: $V_0 = A(0) \frac{S(0)}{b} \Phi(-d^+) - A(0) \left(K + \frac{1-b}{b} S(0) \right) \Phi(-d^-)$
- Si $S(0) = 0$: $V_0 = -KA(0)\mathbf{1}_{K < 0}$

10. Prix de Call dans un modèle displaced log-normal à coefficients constants par morceaux

Dans cette sous-section, nous allons développer de manière succincte les calculs auxquels nous aurions dû nous confronter si nous n'avions pas utilisé les méthodes de moyennage des paramètres dépendant du temps dans la dynamique du taux swap. Nous n'effectuerons pas l'intégralité des calculs, nous nous donnons seulement un aperçu d'une méthode de calcul.

On considère la dynamique suivante pour le taux swap:

$$dS_t = \lambda(t) \left(b(t)S_t + (1 - b(t))S_0 \right) dW_t$$

Où

$$\begin{aligned} \lambda(t) &= \lambda_1 \mathbf{1}_{t \in [0, T_1[} + \lambda_2 \mathbf{1}_{t \in [T_1, T_2[} \\ b(t) &= b_1 \mathbf{1}_{t \in [0, T_1[} + b_2 \mathbf{1}_{t \in [T_1, T_2[} \end{aligned}$$

Donc pour $t \geq T_1$,

$$\begin{aligned} dS_t &= \lambda_2 \left(b_2 S_t + (1 - b_2)S_0 \right) dW_t \\ \iff S_{T_2} &= -S_0 \frac{1-b_2}{b_2} + \left(S_{T_1} + S_0 \frac{1-b_2}{b_2} \right) e^{\lambda_2 b_2 (W_{T_2} - W_{T_1}) - \frac{1}{2} \lambda_2^2 b_2^2 (T_2 - T_1)} \end{aligned}$$

De même, sur $t \in [0, T_1[$,

$$S_{T_1} = -S_0 \frac{1-b_1}{b_1} + \frac{S_0}{b_1} e^{\lambda_1 b_1 W_{T_1} - \frac{1}{2} \lambda_1^2 b_1^2 T_1}$$

Enfin, en injectant dans l'expression de S_{T_2} , il vient:

$$\begin{aligned} S_{t_2} &= -S_0 \frac{1-b_2}{b_2} + S_0 \left(\frac{1-b_2}{b_2} - \frac{1-b_1}{b_1} \right) e^{\lambda_2 b_2 (W_{T_2} - W_{T_1}) - \frac{1}{2} \lambda_2^2 b_2^2 (T_2 - T_1)} \\ &\quad + \frac{S_0}{b_1} e^{\lambda_2 b_2 (W_{T_2} - W_{T_1}) + \lambda_1 b_1 W_{T_1} - \frac{1}{2} (\lambda_2^2 b_2^2 (T_2 - T_1) + \lambda_1^2 b_1^2 T_1)} \\ &= \tilde{S}_{t_2} + \hat{S}_{T_2} \end{aligned}$$

Où \tilde{S}_{t_2} et \hat{S}_{T_2} sont log-normaux.

On a finalement:

$$\begin{aligned} & \mathbb{E} \left[\left(S_{T_2} - K \right)^+ \right] \\ &= \mathbb{E} \left[\left(\tilde{S}_{t_2} + \hat{S}_{T_2} - \left(K + S_0 \frac{1-b_2}{b_2} \right) \right)^+ \right] \\ &= \mathbb{E} \left[\tilde{S}_{T_2} \mathbf{1}_{\tilde{S}_{t_2} + \hat{S}_{T_2} > K + S_0 \frac{1-b_2}{b_2}} \right] - \left(K + S_0 \frac{1-b_2}{b_2} \right) \mathbb{P} \left[\tilde{S}_{t_2} + \hat{S}_{T_2} > K + S_0 \frac{1-b_2}{b_2} \right] \\ &+ \mathbb{E} \left[\hat{S}_{T_2} \mathbf{1}_{\tilde{S}_{t_2} + \hat{S}_{T_2} > K + S_0 \frac{1-b_2}{b_2}} \right] \end{aligned}$$

A ce stade, les calculs deviennent compliqués. On va tenter d'exprimer le terme en probabilité.

On note $X = W_{T_2} - W_{T_1}$ et $Y = W_{T_1}$.

Soit $x \in \mathbb{R}$.

On a ci dessus une forme comme celle ci:

$$\begin{aligned} & \tilde{S}_{t_2} + \hat{S}_{T_2} > C \\ \iff & A e^{\lambda_2 b_2 x - \frac{1}{2} \lambda_2^2 b_2^2 (T_2 - T_1)} + B e^{\lambda_2 b_2 x + \lambda_1 b_1 Y - \frac{1}{2} (\lambda_2^2 b_2^2 (T_2 - T_1) + \lambda_1^2 b_1^2 T_1)} > C \\ \iff & e^{\lambda_2 b_2 x - \frac{1}{2} \lambda_2^2 b_2^2 (T_2 - T_1)} \left[A + B e^{\lambda_1 b_1 Y - \frac{1}{2} \lambda_1^2 b_1^2 T_1} \right] > C \\ \iff & Y > \frac{1}{2} \lambda_1 b_1 T_1 + \frac{1}{\lambda_1 b_1} \left(\ln \left(C e^{\lambda_2 b_2 x - \frac{1}{2} \lambda_2^2 b_2^2 (T_2 - T_1)} - A \right) - \ln(B) \right) = D(x) \end{aligned}$$

Comme $X \perp Y$, $X \sim \mathcal{N}(0, T_2 - T_1)$, $Y \sim \mathcal{N}(0, T_1)$

$$\begin{aligned} & \mathbb{P} \left[\tilde{S}_{t_2} + \hat{S}_{T_2} > C \right] \\ &= \int_{x \in \mathbb{R}} \mathbb{P} \left(Y > D(x), X = x \right) dx \\ &= \int_{x \in \mathbb{R}} f_X(x) \int_{y > D(x)} f_Y(y) dy dx \\ &\dots \\ &= \int_{x \in \mathbb{R}} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \Phi \left(- \frac{D(\sqrt{T_2 - T_1} x)}{\sqrt{T_1}} \right) dx \end{aligned}$$

A partir de ce stade, nous ne parvenons plus à développer. Ces calculs nous montrent que les méthodes de moyennage des paramètres dépendants du temps dans la dynamique du taux swap peuvent grandement simplifier la tâche de pricing d'une swaption.

11. Le modèle quasi-Gaussien multi-facteurs

La motivation principale pour s'orienter vers un modèle multi-facteurs est le comportement des taux forward sur le marché. En effet, on peut observer des taux forward à différentes maturités qui sont plus ou moins corrélés et il est possible de dégager une décorrélation entre le taux forward de maturité 10 ans et le taux forward de maturité 30 ans par exemple. Il s'agit bien d'instruments financiers différents et rien ne garantit une corrélation parfaite entre eux. Toutefois, c'est ce que l'on modélise lorsque l'on utilise un modèle mono factoriel. En effet, si la seule source d'aléa provient d'un mouvement brownien, on remarque que la corrélation instantanée de taux forward est toujours de 1, indépendamment des maturités.

Pour pallier cela, il suffit d'ajouter des sources d'aléa en considérant des taux forward dont les dynamiques suivent un mouvement Brownien multidimensionnel.

La majeure partie des résultats apportés dans le cas unidimensionnel sont généralisables de manière assez naturelle en écriture matricielle. Nous allons introduire rapidement les notations. D'abord, on considère une structure de volatilité à variables séparables

$$\sigma_f(t, T, \omega) = g(t, \omega)h(T)$$

où h prend ses valeurs dans \mathbb{R}^d et g dans $\mathbb{R}^{d \times d}$.

On pose

$$\chi(t) = -\frac{dH(t)}{dt}H(t)^{-1}$$

Et $H(t) = \text{diag}(h(t)) = \text{diag}(h_1(t), \dots, h_d(t))$

Donc

$$\chi(t) = \begin{pmatrix} -h'_1(t) & 0 \\ \dots & \dots \\ 0.. & -h'_d(t) \end{pmatrix} \begin{pmatrix} h_1(t) & 0 \\ \dots & \dots \\ 0.. & h_d(t) \end{pmatrix}^{-1} \iff \chi(t) = \left(-\frac{h'_i(t)}{h_i(t)} \mathbf{1}_{i=j} \right)$$

Donc

$$\forall i = 1, d, \quad h_i(t) = e^{-\chi_i t}$$

On construit grâce à ce qui précède la volatilité du taux court, qui devient:

$$\sigma_r(t, \omega) = g(t, \omega)H(t)$$

Avec cela,

$$\begin{aligned}
H(t)H^f(t)^{-1} &= (H^f(t)H(t)^{-1})^{-1} = \left(\begin{bmatrix} e^{-\chi_1(t+\delta_1)} & \dots & e^{-\chi_d(t+\delta_1)} \\ & \dots & \\ e^{-\chi_1(t+\delta_d)} & \dots & e^{-\chi_d(t+\delta_d)} \end{bmatrix} \begin{bmatrix} e^{-\chi_1 t} & \dots & 0 \\ 0 & \dots & 0 \\ 0 & \dots & e^{-\chi_d t} \end{bmatrix} \right)^{-1} \\
&= \begin{bmatrix} e^{\chi_1 t} e^{-\chi_1(t+\delta_1)} & \dots & e^{\chi_d t} e^{-\chi_d(t+\delta_1)} \\ & \dots & \\ e^{\chi_1 t} e^{-\chi_1(t+\delta_d)} & \dots & e^{\chi_d t} e^{-\chi_d(t+\delta_d)} \end{bmatrix}^{-1} \\
&= \begin{bmatrix} e^{-\chi_1 \delta_1} & \dots & e^{-\chi_d \delta_1} \\ & \dots & \\ e^{-\chi_1 \delta_d} & \dots & e^{-\chi_d \delta_d} \end{bmatrix}^{-1}
\end{aligned}$$

Enfin on en déduit la volatilité locale:

$$\begin{aligned}
\sigma_r(t, x(t), y(t))^\top &= H(t)H^f(t)^{-1} \begin{pmatrix} \lambda_1(a_1 + b_1 f_1(t)) & 0 & 0.. \\ 0 & \lambda_2(a_2 + b_2 f_2(t)) & 0.. \\ 0 & .. & \lambda_d(a_d + b_d f_d(t)) \end{pmatrix} D^f(t)^\top \\
&= \tilde{H}(t) \sigma^f(t, f(t)) D^f(t)^\top
\end{aligned}$$

Où $D^f(t)^\top$ est telle que

$$\Gamma = D^f(t)^\top D^f(t)$$

avec Γ la matrice de corrélation des taux $f_i(t)$. Pour faire ce calcul, nous utiliserons une décomposition de Cholesky.

Enfin,

$$\forall i = 1, d, \quad f_i(t) = f(t, t + \delta_i) = f(0, t + \delta_i) + \mathbf{1}^\top H(t + \delta_i)H(t)^{-1}(x(t) + y(t) G(t, t + \delta_i))$$

$$G(t, T) = \int_t^T H(u)H(t)^{-1} \mathbf{1} \, du = \begin{bmatrix} \frac{1-e^{-\chi_1(T-t)}}{\chi_1} \\ \dots \\ \frac{1-e^{-\chi_d(T-t)}}{\chi_d} \end{bmatrix}$$

Avec toutes la construction que l'on a faite, on conserve les propriétés importantes du modèle unidimensionnel. Les voici sous leur forme matricielle:

Proposition 8 *Dans ce cadre, le prix des obligations zero coupon et le taux court satisfont:*

$$\begin{aligned}
P(t, T, x(T), y(T)) &= \frac{P(0, T)}{P(0, t)} \exp \left(-G(t, T)^\top x(T) - \frac{1}{2} G(t, T)^\top y(T) G(t, T) \right) \\
r(t) &= f(0, t) + \mathbf{1}^\top x(t)
\end{aligned}$$

Où

$$\begin{cases} dx(t) &= (y(t)\mathbf{1} - \chi(t)x(t)) dt + \sigma_r(t, \omega)^\top dW_t \\ dy(t) &= (\sigma_r(t, \omega)^\top \sigma_r(t, \omega) - \chi(t)y(t) - y(t)\chi(t)) dt \end{cases}$$

$$x(t) \in \mathbb{R}^d, y(t) \in \mathbb{R}^{d \times d}, x(0) = 0, y(0) = 0$$

Schéma d'Euler

Comme $r(t) = f(0, t) + \mathbf{1}^\top x(t)$

$\implies e^{-\int_0^{T_0} r(s) ds} = P(0, T_0) e^{-\sum_{i=1}^d \int_0^{T_0} x^{(i)}(s) ds} = P(0, T_0) e^{\sum_{i=1}^d I_i(T_0)}$ nous allons faire les mêmes simulations que pour la version $d = 1$ en ajoutant une boucle d'itération pour la dimension d . La fonction de simulation du schéma discret renvoie M triplets du type

$$\begin{bmatrix} x_1(T_0) & \dots & x_d(T_0) \end{bmatrix}, \begin{bmatrix} y_{1,1}(T_0) & \dots & y_{1,d}(T_0) \\ \dots & & \\ y_{d,1}(T_0) & \dots & y_{d,d}(T_0) \end{bmatrix}, \begin{bmatrix} I_1(T_0) & \dots & I_d(T_0) \end{bmatrix}$$

On voit donc que l'effort computationnel est plus important en multi-dimension.

On peut même écrire le schéma log-Euler dans ce cas. En dimension 1, on avait:

$$x(t) = e^{u(t)} - \frac{a}{b}$$

En dimension d , on a:

$$x(t) = \left[e^{u^{(i)}(t)} - \frac{a^{(i)}}{b^{(i)}} \right]_{i=1,d}$$

$$\text{Or } \forall i, \quad dx^{(i)}(t) = \left(y^{(i)}(t)\mathbf{1} - \chi_i x^{(i)} \right) dt + \sum_{j=1}^d \sigma_r^{(i,j)} dW_t^{(j)}$$

Donc

$$\forall i, \quad du^{(i)}(t) = \frac{1}{x^{(i)}(t) + \frac{a^{(i)}}{b^{(i)}}} \left(y^{(i)}(t)\mathbf{1} - \chi_i x^{(i)} - \frac{1}{2} \frac{1}{x^{(i)}(t) + \frac{a^{(i)}}{b^{(i)}}} \sum_{k,j=1}^d \sigma_r^{(i,k)} \sigma_r^{(i,j)} \rho_{k,j} \right) dt + \frac{1}{x^{(i)}(t) + \frac{a^{(i)}}{b^{(i)}}} \sigma_r^{(i)} dW_t$$

Ici $\rho_{k,j} = \mathbf{1}_{k=j}$

$$= \frac{1}{x^{(i)}(t) + \frac{a^{(i)}}{b^{(i)}}} \left(y^{(i)}(t)\mathbf{1} - \chi_i x^{(i)} - \frac{1}{2} \frac{1}{x^{(i)}(t) + \frac{a^{(i)}}{b^{(i)}}} \sigma_r^{(i)} (\sigma_r^{(i)})^\top \right) dt + \frac{1}{x^{(i)}(t) + \frac{a^{(i)}}{b^{(i)}}} \sigma_r^{(i)} dW_t$$

Nous observons que les diffusions dans le modèle multidimensionnel sont plus "sensibles" aux paramètres du modèle, en raison des produits matricielles qui interviennent dans le calcul de la volatilité σ_r . Nous remarquons notamment que les diffusions explosent très rapidement lorsque l'un des paramètres χ_i est proche de zéro.

Enfin, on peut noter que ce modèle fait intervenir des paramètres supplémentaires, en l'occurrence les tenor de références δ_i . De ce fait, le modèle multidimensionnel avec $d = 1$ et le modèle unidimensionnel traité dans le corps du mémoire ne sont pas exactement les mêmes modèles. En effet, aucun tenor n'apparaît dans la structure de volatilité du modèle unidimensionnel.