

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 is here 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 de la swaption	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	Construction de la courbe Zero Coupon	16
3.5	Courbe des prix et choix de simulation	21
3.6	Approximation par un modèle "displaced log-normal"	25
3.6.1	La swaption comme Call sur taux swap	25
3.6.2	La projection Markovienne pour la volatilité locale	26
3.6.3	La formule de Pricing	32
4	Courbe de volatilité implicite	34
4.1	Volatilité implicite log-normale	34
4.2	Volatilité implicite normale	34
4.3	Observations et propriétés sur le skew	34
5	Calibration de volatilité locale en dimension 1	35
5.1	Paramétrisation du problème	35
5.2	Algorithme et résultats	35
6	Conclusion	36
	Références	37
	Annexe	38

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 européens du fait de leur importante liquidité sur le marché. De plus, le payoff de ces options fait intervenir plusieurs sous-jacents 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 du swaption européen 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 stochastique.

Notations

- Probabilité historique \mathbb{P} / probabilité risque neutre \mathbb{Q}
- $\mathbb{E}_{\mathbb{Q}}$: 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 : $B(t) = e^{-\int_0^t r(s) ds}$
- 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 (On notera $P(t, T)$ dans l'ensemble du document pour désigner le prix de l'obligation zéro coupon entre les dates t et T) 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 $T - N$
- $S > T$: 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

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

Cela 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**" (modèles de taux court et 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 $t \rightarrow P(t, T)$ de manière continue pour tous les horizons 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 sont issus 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, le prix du zéro coupon est borné et prend ses valeurs dans l'intervalle $]0, 1]$. Puisque la fonction $x \rightarrow \ln(x) \in \mathcal{C}^2(\mathbb{R}^{+*}, \mathbb{R})$, elle est bornée sur $]0, 1]$ et ses deux premières dérivées également. On peut donc 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 théorème de Schwarz, 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) 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. 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 exclue 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,

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

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 $g : \mathbb{R} \rightarrow \mathbb{R}^d$.

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 $\lambda * a$ et $\lambda * b$. 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 de la swaption

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} , $i = 0, N - 1$. Mathématiquement, on peut écrire le payoff du swap de la manière suivante:

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

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:

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

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:

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

3.1 Valorisation sous la probabilité risque neutre

Afin de valoriser les swaptions, nous allons exploiter leur expression sous la forme d'espérance afin de l'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:

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

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, 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:

$$d\left(\frac{P(t, T_N)}{B(t)}\right) = -\frac{P(t, T_N)}{B(t)} \sigma_P(t, T_N)^\top dW_t^\mathbb{Q}.$$

$$\implies \frac{P(t, T_N)}{B(t)} = P(0, T_N) e^{-\int_0^t \sigma_P(s, T_N)^\top dW_s^\mathbb{Q} - \frac{1}{2} \int_0^t \sigma_P(s, T_N)^\top \sigma_P(s, T_N) ds}$$

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

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

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 4.

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

$$\chi(t) = \chi \in \mathbb{R}$$

et

$$\sigma_r(t) = \sigma_r(t, x(t), y(t)) = \lambda(a + bx(t))$$

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:

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

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. 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.

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,

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

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 que l'on va borné artificiellement.

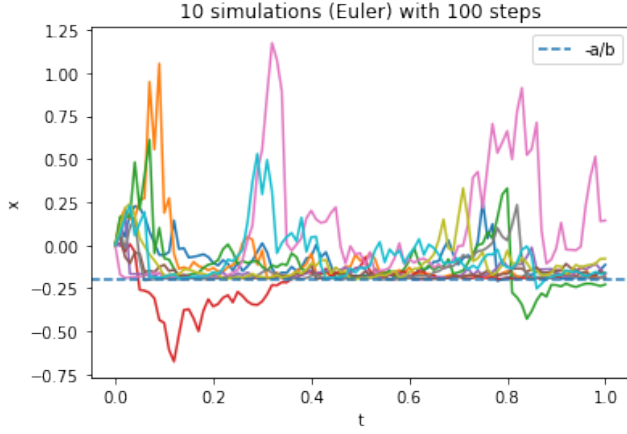


Figure 1: Schéma d'Euler

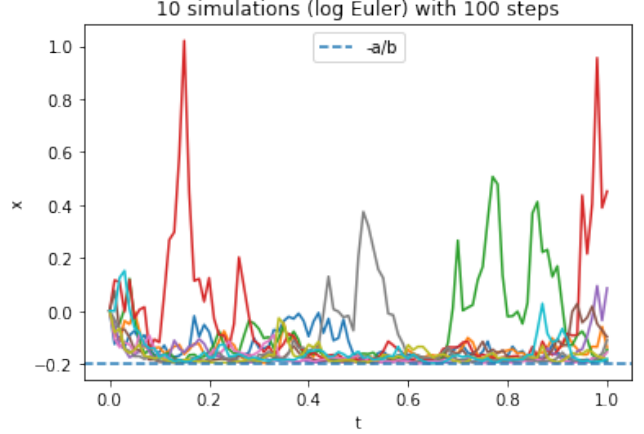


Figure 2: Schéma log-Euler

3.4 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 prix des premiers zero coupons $P(0, t) = P(t)$, $t \leq T_1$:

On a dans notre cas $P(0, 0.5) = P(0, 1)$

Donc

$$S(1) = \frac{1 - P(1)}{\tau(P(1) + P(0.5))} \iff P(1) = P(0.5) = \frac{1}{2\tau S(1) + 1}$$

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:

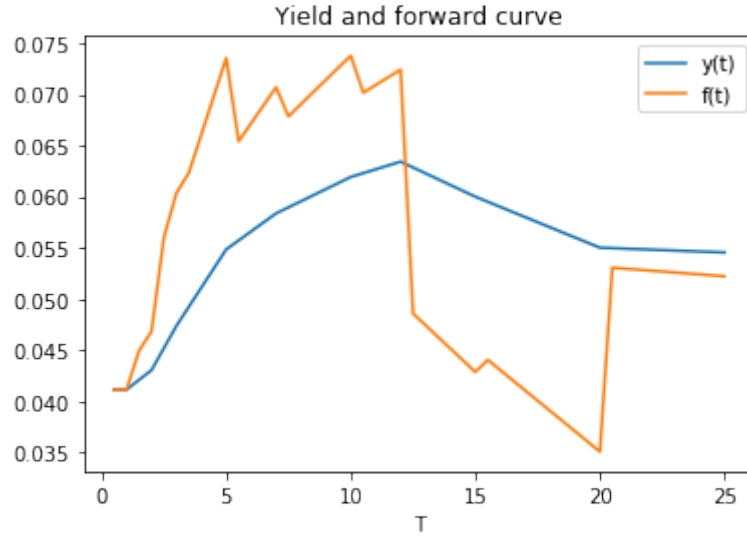


Figure 3: Courbe de taux linéaire par morceaux

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.

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)

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

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 $P(0.5) = P(1)$ 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

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

Voici le rendu ci-dessous:

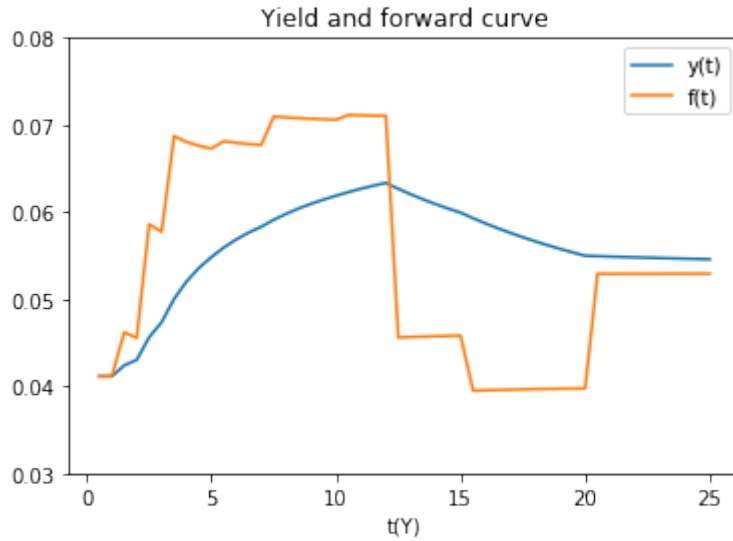


Figure 4: Courbe de taux forward constant par morceaux

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 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 usant 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.

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

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

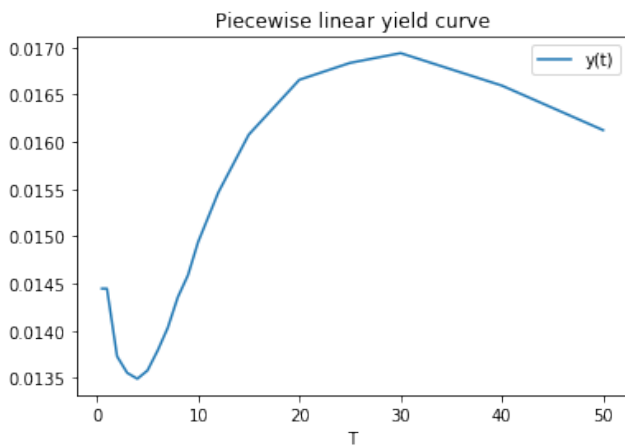


Figure 5: Yield linéaire par morc.

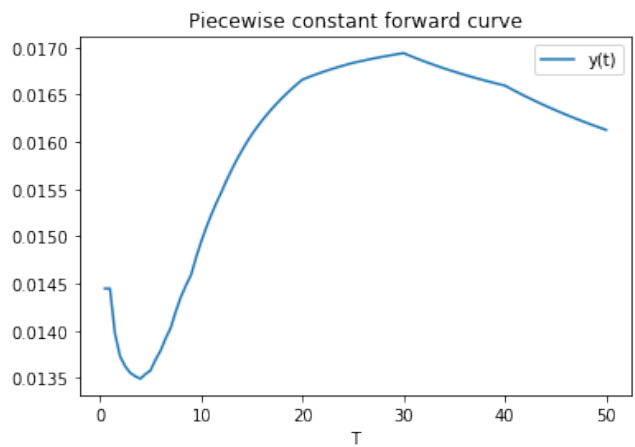
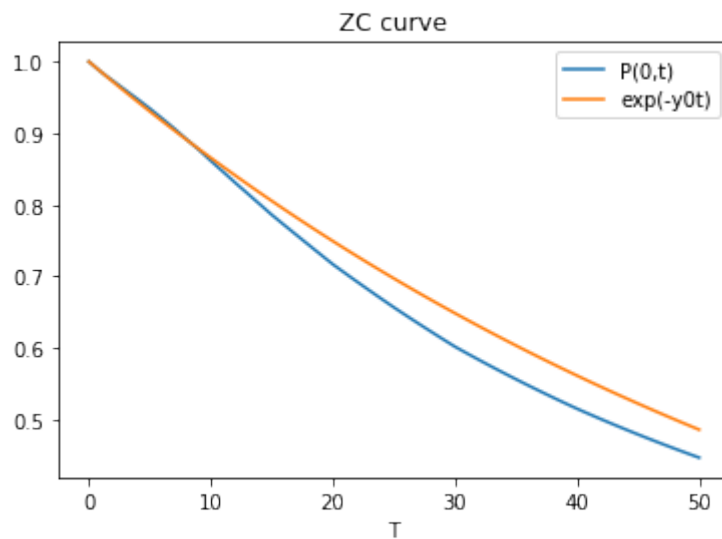


Figure 6: Forward constant par morc.

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.



3.5 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.

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

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

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 $2 \cdot 10^{-4}$. Enfin, on observe 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 on calcule le nombre de simulations à réaliser afin d'atteindre cet objectif d'erreur, il vient:

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

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) $(\lambda, a, b, \chi) = (1, 1, 1, 1)$.

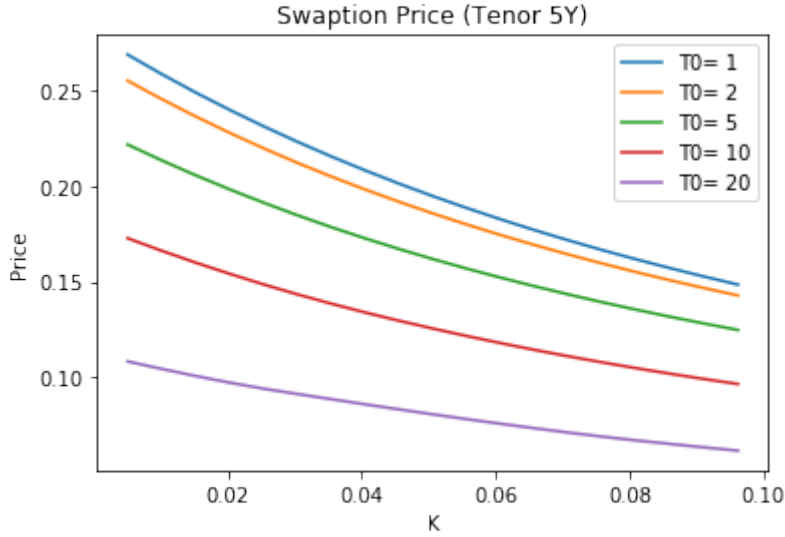
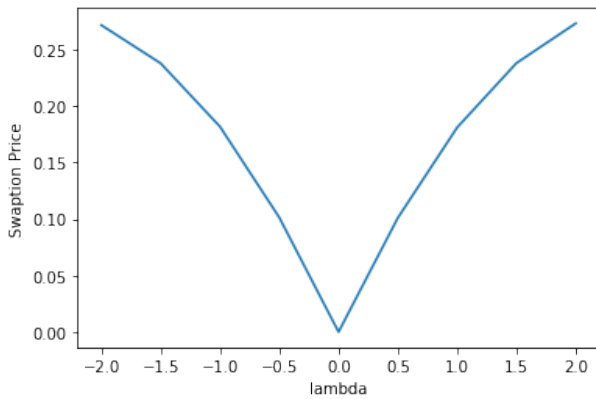


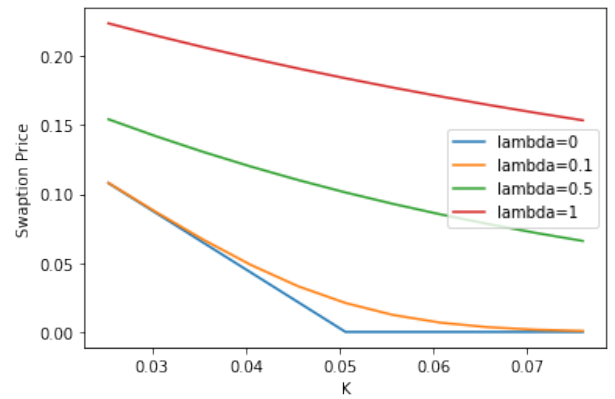
Figure 7: Courbe des prix de swaption générés

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



(a) Prix ATM



(b) Courbe de Prix

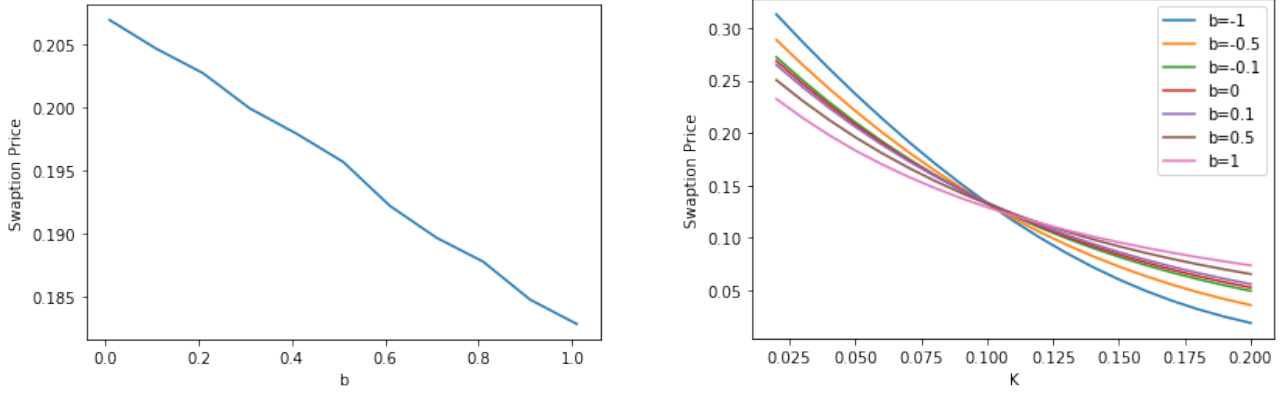
Figure 8: Variations du prix en fonction de λ

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



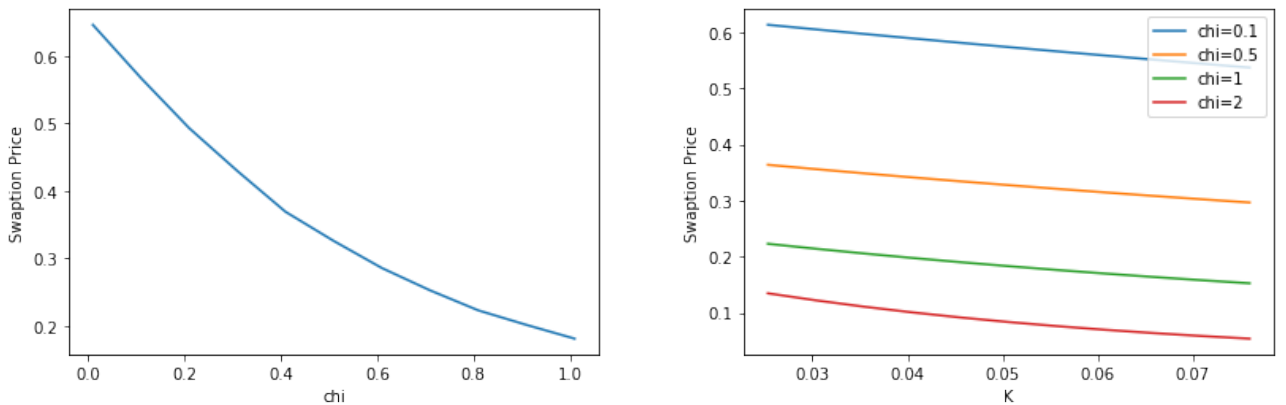
(a) Prix ATM

(b) Courbe de Prix

Figure 9: 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 10: 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.

3.6 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.

3.6.1 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.

3.6.2 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:

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

$$\begin{aligned} \implies A(t, x(t), y(t)) &= A(0) \exp \left\{ \int_0^t r(s) ds - \sum_{n=0}^{N-1} \tau_n \int_0^t P(s, T_{n+1}, x(s), y(s)) \omega_n(s) dW_s^{\mathbb{Q}} \right. \\ &\quad \left. - \frac{1}{2} \int_0^t \left(\sum_{n=0}^{N-1} \tau_n \int_0^t P(s, T_{n+1}, x(s), y(s)) \omega_n(s) \right)^2 ds \right\} \end{aligned}$$

Le théorème de Girsanov nous permet de 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:

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

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

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 vol locale.

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,

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

Dans notre cas, 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:

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

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

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

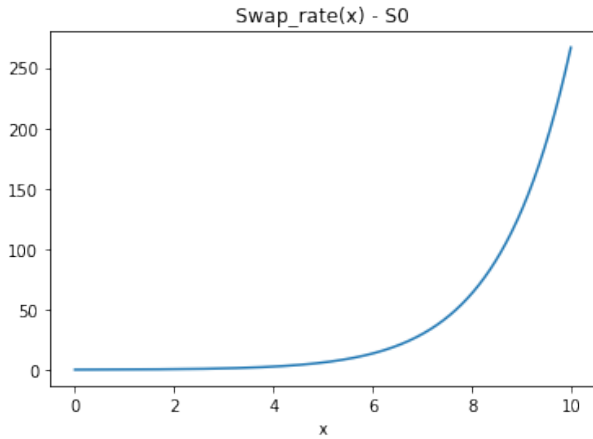
Remarque: La difficulté réside dans l'expression de la variance en raison de la stochasticité de la volatilité. On a donc

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

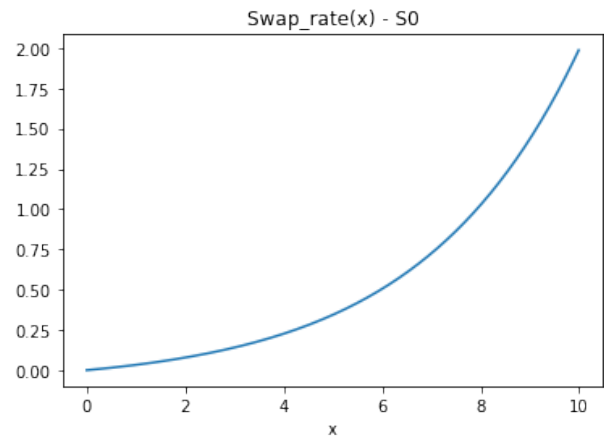
que l'on pourrait approcher par un algorithme de Monte Carlo par exemple. Mais une possibilité est de considérer une dynamique gaussienne pour le taux swap avec une volatilité déterministe $\frac{\partial S}{\partial x}(s, 0, 0)\sigma_r(s, 0)$ telle que $Var^A[S(t)] \approx \int_0^t \frac{\partial S}{\partial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

Discussion

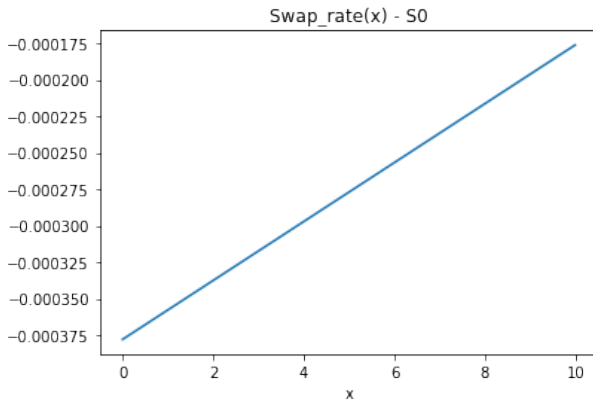
Nous avons souligné le fait que la fonction x_0 pouvait être évaluée à chaque instant t assez simplement grâce à une méthode de Newton-Raphson. Néanmoins, cette méthode présente une limite que nous allons mettre en lumière ici. En effet, on peut dégager une erreur d'approximation de la racine qui croît 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 $x \rightarrow Swaprate(x) - S_0$ 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 11) les graphiques observés (à l'instant t fixé à 0.95).



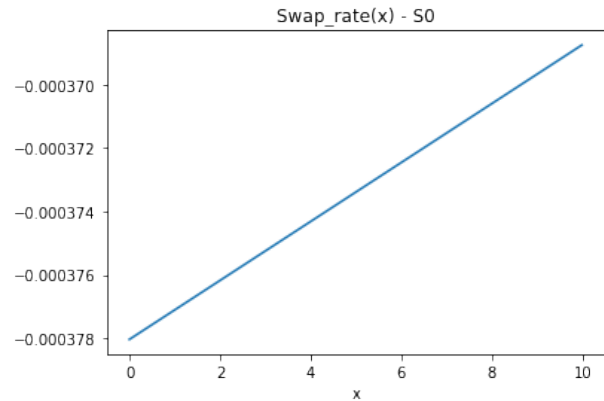
(a) $T_0 = 1Y$ $x_0 = 0.0020$



(b) $T_0 = 2Y$ $x_0 = 0.0084$



(c) $T_0 = 10Y$ $x_0 = 18.7018$



(d) $T_0 = 14Y$ $x_0 \approx 406$

Figure 11: 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 graphiques (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 12 nous montre ce que l'on observe si l'on trace la dérivée $\frac{\partial S}{\partial x}$.

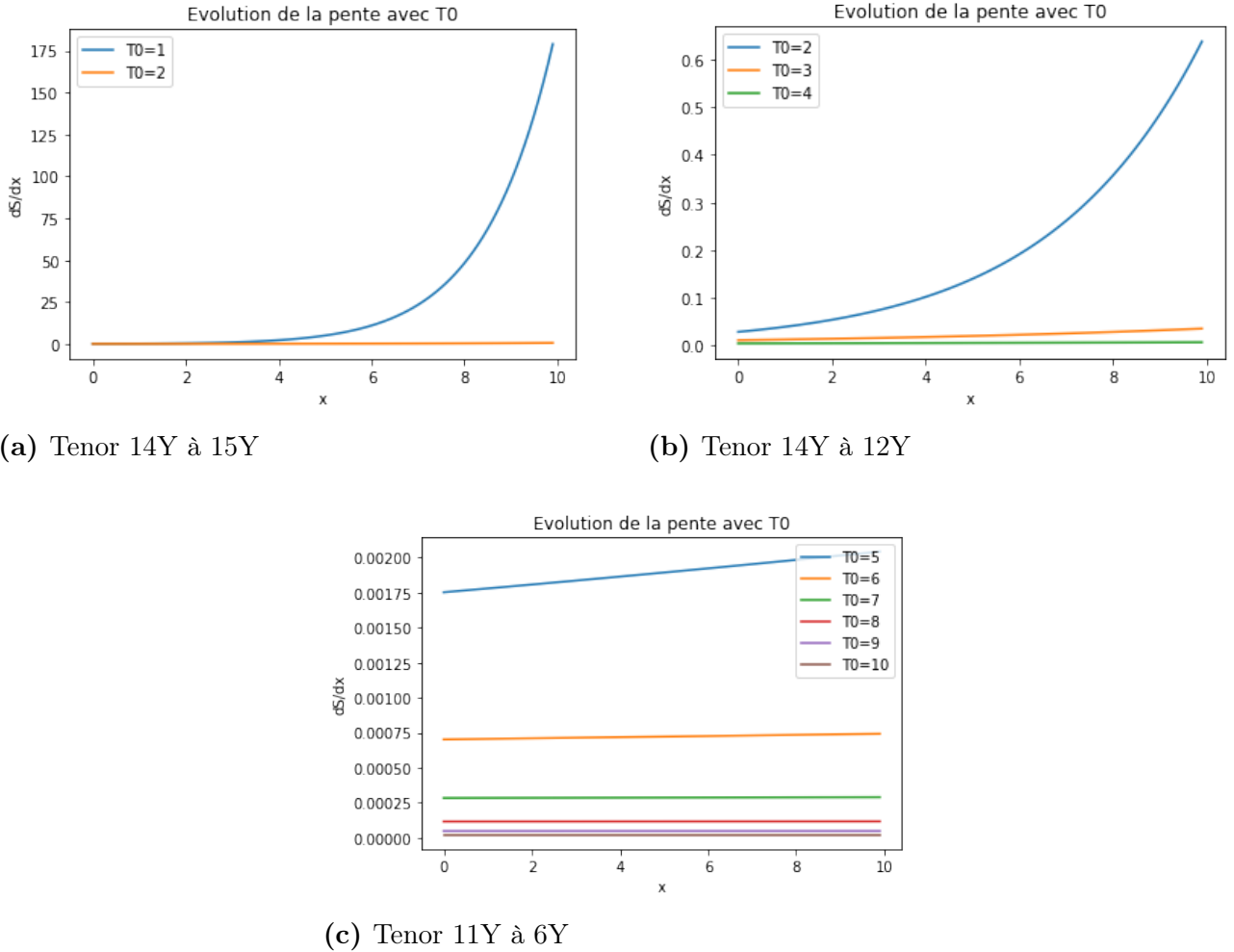


Figure 12: Evolution de la pente $\frac{\partial S}{\partial x}$

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.

3.6.3 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 défini en annexe.

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:

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

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

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, 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 6. 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, à la formule de pricing suivante

Proposition 3

$$V_0 = A(0) \left(\left(S(0) + \frac{1-\bar{b}_S}{\bar{b}_S} S(0) \right) \Phi(d^+) - \left(K + \frac{1-\bar{b}_S}{\bar{b}_S} S(0) \right) \Phi(d^-) \right)$$

$$d^{+/-} = \frac{\ln \left(\frac{S(0) + S(0) \frac{1-\bar{b}_S}{\bar{b}_S}}{K + S(0) \frac{1-\bar{b}_S}{\bar{b}_S}} \right) + / - \frac{\bar{\lambda}_S^2 \bar{b}_S^2}{2} T_0}{\bar{\lambda}_S \bar{b}_S \sqrt{T_0}}$$

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.

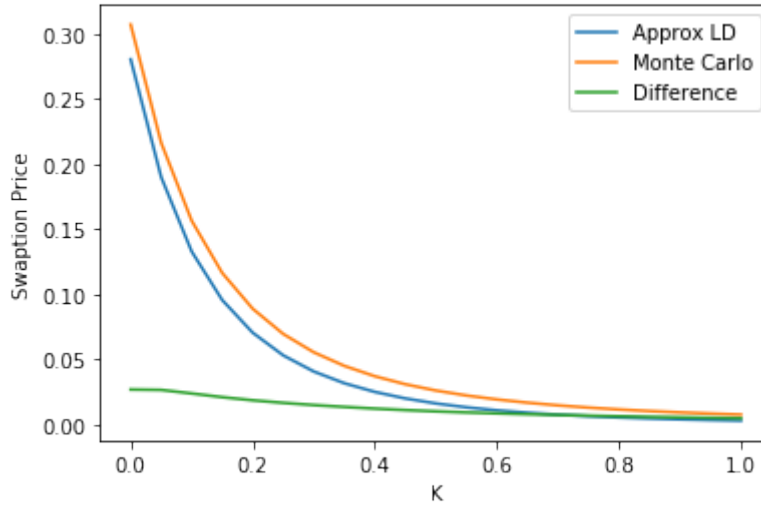


Figure 13: Comparaison des méthodes de pricing

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.

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

Utilité du calcul de la vol implicite pour la calibration: calibrage sur la courbe de vol. plutôt que sur le prix car le niveau de la vol implicite est plus élevé que celui des prix, ce qui facilite les étapes d'optimisation.

4.1 Volatilité implicite log-normale

4.2 Volatilité implicite normale

4.3 Observations et propriétés sur le skew

Limite de l'approx Log displaced : montrer forme IV LD (décroissante rigide) ne reproduit pas des smiles observés sur le marché et est 'trop rigide' par rapport à la courbe de vol. impli de Monte Carlo. (voir P. 456 Brigo mercurio)

5 Calibration de volatilité locale en dimension 1

5.1 Paramétrisation du problème

5.2 Algorithme et résultats

6 Conclusion

References

- [1] Piterbarg Andersen. *Interest rate modeling*. Vol. 1. 2010.
- [2] Piterbarg Andersen. *Interest rate modeling*. Vol. 2. 2010.
- [3] E.Chevalier.et.al. “Optimal exit strategies for investment projects”. In: *J. Math. Anal. Appl.* (2015).
- [4] J. Healy. “Equivalence between forward rate interpolations and discount factor interpolations for the yield curve construction”. In: *arXiv preprint arXiv:2005.13890* (2019).
- [5] O.Cheyette. “Markov Representation of the Heath-Jarrow-Morton Model”. In: 1 (1995).
- [6] Vladimir Piterbarg. “Markovian Projection Method for Volatility Calibration”. In: (2006).

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. Le Théorème de représentation des martingales

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

4. Différence de distribution entre probabilité risque neutre et terminale

5. Preuve de la condition de Feller

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