

**Mémoire présenté devant l'Université Paris Dauphine
pour l'obtention du diplôme du Master Actuariat
et l'admission à l'Institut des Actuaire**

le 21 novembre 2016

Par : Clément Lion

Titre: Le cadre HJM appliqué à la génération de scénarios

Confidentialité : ☒ NON ☐ OUI (Durée : ☒ 1 an ☐ 2 ans)

Les signataires s'engagent à respecter la confidentialité indiquée ci-dessus

*Membre présent du jury de l'Institut
des Actuaire :*

Signature : Entreprise : Predica

Nom :

Signature :

Directeur de mémoire en entreprise :

Nom : Mme Shaojuan Huang

Signature :

*Membres présents du jury du Master
Actuariat de Dauphine :*

Autorisation de publication et de mise en ligne sur un site de diffusion de documents actuariels (après expiration de l'éventuel délai de confidentialité)

Signature du responsable entreprise :

Secrétariat :

Bibliothèque :

Signature du candidat :

Remerciements

Je tiens d'abord à remercier Florence Bailly pour m'avoir accueilli dans son équipe. Le sujet proposé fut étonnamment riche, et m'a permis d'affiner mes connaissances sur la modélisation des taux, ainsi qu'élargir ma culture des méthodes numériques utilisées dans le domaine.

Mes remerciements vont également à Paul Gassiat pour son suivi et sa relecture très attentionnée de ce mémoire, et à Nicole Burrus pour ses conseils avisés.

Merci à toute l'équipe EDM et aux voisins pour leur collaboration et la bonne ambiance à laquelle ils contribuèrent : Abdelkader, Cédric, Dominique, Francis, Lopez, Michael, Mohamed, Nafia, Othman, Pierre-Hugues, Thami, Yaya.

Enfin, je souhaite remercier tout particulièrement Shaojuan Huang, pour son encadrement durant six mois. Son soutien, sa disponibilité à toute épreuve, son excellence technique et son ouverture d'esprit furent autant de qualités sans lesquelles ce mémoire aurait été bien différent.

Table des matières

I	Les générateurs de scénarios économiques	2
1	Contexte	2
1.1	Solvabilité II	2
1.2	La gestion actif-passif	3
1.3	Indicateurs	4
2	Les taux	6
2.1	Différents taux	6
2.2	Produits de taux	6
2.3	Pricing des dérivés	7
2.4	Modèles classiques de taux	8
II	Un modèle HJM avec volatilité stochastique	13
1	Présentation	13
1.1	Le cadre HJM	13
1.2	Forme générale	13
1.3	Taux forward	13
1.4	Prix des zéro-coupons	14
2	Options sur zéro-coupons	16
2.1	Fonction caractéristique Ψ	16
2.2	Prix de l'option	17
3	Swaptions	19
3.1	Dynamiques sous la mesure LVL	19
3.2	Première approximation	20
3.3	Duration stochastique	20
4	Implémentation	23
4.1	Résolution de l'EDO de Riccati	23
4.2	Intégration de Ψ	25
4.3	Optimisation	27
4.4	Simulation	30
III	Analyse et comparaison des résultats	31
1	Calibration du modèle	31
1.1	Données	31
1.2	Résultats de calibration	33
1.3	Prix et volatilités implicites	35
2	Simulation	40
3	Intégration et comparaison	45
3.1	Résultats ALM	45
3.2	Comparaison des modèles	45

Annexes	I
A Quelques méthodes numériques	I
1 Algorithme de Nelder-Mead	I
2 Méthodes de Runge-Kutta : tabulation de coefficients	I
3 Quadratures de Gauss	II
4 Simulation de variables aléatoires	III
5 Comparaison des performances	IV
B Rappels de calcul stochastique	V
1 Formule d'Itô	V
2 Changement de numéraire	V
3 Formule de Bachelier	V
4 Formule de Black	VI
5 Volatilités implicites	VI
6 Convergence des schémas d'approximation des EDS	VI
C Modèles classiques : détails de calcul	VIII
1 Modèle de Vašíček	VIII
2 Modèle CIR	VIII
D Modèle HJM de Schwartz et Trolle : détails de calculs	X
1 Taux forward	X
2 Zéro-coupons	XI
E Données de l'étude	XIII
1 Prix des swaptions	XIII
2 Courbes de taux	XIV

Résumé

Ce mémoire a pour objet l'implémentation d'un modèle de taux décrit par Schwartz et Trolle dans un récent article (2009). La modélisation des taux est devenue un enjeu crucial pour les compagnies d'assurance-vie. D'une part, l'environnement actuel de taux bas rend leur gouvernance délicate (les assureurs limitent autant que possible la collecte de produits d'épargne, dont ils ont du mal à assurer les rendements), anticiper leurs variations et adopter des stratégies de placements adéquates est devenu une priorité. D'autre part, la réglementation actuelle crée une forte demande de modélisation financière et aujourd'hui, la quasi-totalité des études quantitatives passent par des simulations stochastiques, donc par la modélisation de l'environnement de marché et de la compagnie.

Le mémoire se décompose en trois parties : une première remet ce mémoire dans son contexte légal, et rappelle les modèles de taux couramment utilisés. Une seconde traite du modèle lui-même : les équations qui le décrivent, la méthode de valorisation des options, et les aspects plus pratiques avec les méthodes numériques qui ont été choisies. Enfin, les résultats de la calibration puis de la simulation du modèle seront commentés.

Assurance vie, gestion actif-passif, générateur de scénarios économiques, calcul stochastique, modèle de taux, cadre HJM

Abstract

This thesis is about the implementation of an interest rate model. It was drawn by Schwartz and Trolle in a 2009 paper. Fixed income modelling has become a key issue for life insurers. Firstly, their governance is made difficult by the current low interest rates environment (more and more insurers try to avoid as much as possible to open new savings accounts). The anticipation of the future interest rates and the adoption of appropriate investment strategies is now a top priority. Secondly, the new legislation created a strong need for financial modelling and now, a majority of quantitative studies are based on stochastic simulations. That means all those studies are taken care of thanks to the stochastic modelling of the company and its environment.

The thesis is composed of three parts : the first one puts it in its legal context and describe the classical rate models. The model of Schwartz and Trolle fills the second part, with its theoretical analysis (from the first equations to the option pricing formulae) and practical use (the numerical methods). Finally, we comment in a third part the calibration of that model and its simulation through an ALM tool.

Life insurance, asset-liability management, economic scenarios generator, stochastic calculus, interest rate model, HJM framework

Notations

Symbole	Description
P	Mesure de probabilité historique
Q	Mesure de probabilité risque-neutre
W^M	Processus de Wiener sous la mesure M (si c'est une mesure) ou sous la mesure associée à M
$\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$	Filtration naturelle associée aux processus de Wiener
$\mathbb{E}_t^M, \mathbb{V}_t^M$	Espérance et variance conditionnellement à la tribu \mathcal{F}_t selon la mesure M (si c'est une mesure) ou selon la mesure associée à M
$\langle X_t \rangle, \langle X_t, Y_t \rangle$	Variation quadratique de X , covariation quadratique de X et Y
$\mathcal{N}(\mu, \sigma^2)$	Loi normale de moyenne μ et d'écart-type σ
$F_{\mathcal{N}}$	Fonction de répartition de la loi normale centrée, réduite
$(X)_+$	$= \max(X, 0)$: partie positive de X
$X \wedge Y$	$= \inf\{X, Y\}$
$\Re[X], \Im[X]$	Partie réelle de X , partie imaginaire de X
$P(t, T)$	Prix à l'instant t du zéro-coupon d'échéance $T \geq t$
$f(t, T)$	Taux forward instantané à T vu à $t \leq T$
r_t	Taux court à l'instant t
$\mathcal{P}(t, T, S, K)$	Prix à t d'une option de vente de maturité T et de strike K sur un zéro-coupon d'échéance S
$\text{Caplet}(t, T, S, K)$	Prix à t du caplet de maturité T livré à S et de strike K
$S_{\alpha, \beta}(t)$	Taux swap à la date t associé aux dates $T_\alpha, T_{\alpha+1}, \dots, T_\beta$
$\text{Swaption}_{\alpha, \beta}(t)$	Prix à t de la swaption d'échéance T_α et de ténor $T_\beta - T_\alpha$
$LVL(t)$	$= \sum_{i=\alpha+1}^{\beta} (T_i - T_{i-1})P(t, T_i)$: prix du point de base, ou <i>level</i>
$\text{Black}(X, K, s)$	$= XF_{\mathcal{N}}\left(\frac{1}{s} \ln\left(\frac{X}{K}\right) + \frac{s}{2}\right) - KF_{\mathcal{N}}\left(\frac{1}{s} \ln\left(\frac{X}{K}\right) - \frac{s}{2}\right)$ (formule de Black)

Introduction

La rédaction de ce mémoire a été faite au sein de Predica, une compagnie d'assurance vie filiale de Crédit Agricole Assurances. Avec une large gamme de produits allant de l'épargne à la dépendance, et un encours de 228 milliards d'euros, l'entreprise est le deuxième assureur vie en France¹.

Plus particulièrement, les travaux ont été réalisés dans l'équipe Étude et Développement de Modèles, en charge d'implémenter des outils de simulation pour la gestion actif-passif. Son objectif actuel est de développer un générateur de scénarios économiques entier et indépendant des outils déjà existants du commerce. Le modèle de taux qui sera présenté dans ce mémoire sera intégré à cet outil.

La nouvelle réglementation Solvabilité II, appliquée depuis le début de l'année 2016, demande aux assureurs plus d'études quantitatives qu'auparavant. Notamment, chaque compagnie doit calculer annuellement un ratio de solvabilité, qui agrège l'ensemble des risques de la compagnie sous divers scénarios catastrophes. Avec ces études naissent de forts besoins de modélisation et de simulation, aussi bien du passif (comportement des assurés, lois de longévité...) que de l'actif (le marché).

En tant que gestionnaire d'épargne, Predica est particulièrement sensible à l'environnement de taux d'intérêts. En effet c'est sur les taux du marché, via des obligations (60% du portefeuille de la compagnie), que l'assureur va puiser ses rendements pour d'une part se rémunérer, et d'autre part rémunérer les contrats des assurés. L'environnement de taux bas actuel rend sa gestion plus délicate ; nombre d'assureurs limitent autant que possible la collecte sur les produits d'épargnes, qu'ils ont du mal à rémunérer. Aussi la modélisation des taux d'intérêts, au cœur des études quantitatives de la compagnie, est-elle particulièrement importante. C'est ce à quoi nous allons nous intéresser dans ce mémoire.

Une première partie rappellera en quoi l'environnement réglementaire impacte le travail de l'ALM, et comment l'ALM répond aux besoins réglementaires. Nous y reverrons aussi la base de la modélisation des taux d'intérêts, avec les principaux modèles pour leur importance historique, mais aussi car ils constituent souvent la base de modèles plus raffinés. Nous étudierons dans un second temps un modèle particulier, tiré d'un article d'Eduardo Schwartz et Anders Trolle [1] : d'abord d'un point de vue théorique, en reprenant largement l'article, puis d'un point de vue plus pratique en étudiant son implémentation. La troisième partie sera consacrée à l'étude du modèle dans son environnement naturel, c'est-à-dire au sein d'un générateur de scénarios économiques.

1. Chiffres de 2014

Première partie

Les générateurs de scénarios économiques

1 Contexte

1.1 Solvabilité II

Solvabilité II est une réforme européenne de la réglementation des assurances. Elle fait suite à la réforme Solvency exigeant des assureurs une marge de solvabilité, c'est-à-dire une quantité de fonds propres minimale, calculée en fonction du portefeuille, afin de faire face aux risques futurs. Cette dernière réforme s'est montrée peu adaptée et insuffisante après la crise de 2008. Acceptée en mars 2014 par le Parlement européen pour une application dès janvier 2016, Solvabilité II s'inspire largement des accords de Bâle III avec sa composition en trois piliers :

- Pilier I – Exigences quantitatives : La directive prévoit le calcul régulier d'un MCR (*Minimum Capital Requirement*) et d'un SCR (*Solvency Capital Requirement*). Le premier se calcule avec une formule linéaire simple fonction des primes et provisions techniques. Un ajustement est éventuellement opéré s'il est trop faible, ou trop élevé, par rapport au SCR. Celui-ci se calcule suivant un algorithme déterminé par la directive, ou par un modèle interne partiel ou complet validé par le régulateur. le principe est d'évaluer le quantile d'ordre 99.5% des fonds propres à horizon un an.
Les assureurs ne respectant pas le SCR doivent soumettre un plan de redressement au régulateur pour se remettre en conformité en six mois. Ceux ne respectant pas le MCR ont trois mois pour retrouver des fonds propres au moins égaux au MCR. Dans le cas contraire, le régulateur peut suspendre leur autorisation d'exercer.
- Pilier II – Exigences qualitatives : Ce pilier émet des exigences en terme de gouvernance et de pilotage de l'entreprise. Elle décrit notamment quatre fonctions devant être intégrées à l'entreprise (fonctions actuarielle, d'audit interne, de gestion des risques et de conformité). En particulier, la direction se doit d'être au fait des risques encourus par la compagnie et des stratégies pour y faire face.
Enfin, un ensemble de processus appelé ORSA (*Own Risk and Solvency Assessment*) doivent être mis en place afin d'identifier et quantifier les risques sur le long terme, afin d'assurer que les exigences de fonds propres seront respectées année après année, et d'adapter en fonction les stratégies financières et commerciales.
- Pilier III – Exigences de transparence : Ce dernier pilier demande aux assureurs de publier des informations détaillées sur leurs conditions financières, selon un format spécifié.

1.2 La gestion actif-passif

La gestion actif-passif, ou ALM (pour *Asset-Liability Management*) est un métier très présent en banque et en assurance, qui consiste à étudier les investissements au regard du portefeuille de clients, et de s'assurer de leur alignement en terme de duration et de sensibilités à certains facteurs exogènes. La vue globale qu'ont ces équipes sur le bilan de l'entreprise et les outils auxquels elles ont accès les prédisposent naturellement à diverses études dans le cadre de Solvabilité II.

L'ALM étudie notamment :

- Le risque de liquidité : le risque que l'entreprise ne parviennent pas à investir ou à céder des actifs, l'empêchant ainsi de réaliser sa stratégie d'allocation
- Le risque de change : les risques liés aux variations des taux de change, dans le cas où l'entreprise a des activités ou des placements en différentes devises
- Le risque de taux : ce sont les risques liés aux incertitudes des taux de prêt et d'emprunt sur les marchés. Ce risque est particulièrement important pour les compagnies d'assurance vie car ces taux définissent les prix des obligations, constituant environ deux tiers des placements de ces entreprises.

Aux premiers temps de l'ALM, il s'agissait par exemple de choisir les obligations à acheter de sorte que chaque année, en l'absence de changement des investissements et du portefeuille, les durations du passif et de l'actif restent égales. Aujourd'hui les tâches se sont complexifiées : les investissements se sont diversifiés, les contrats laissent plus de liberté aux clients, le régulateur est plus contraignant, et les objectifs dépassent le simple cadre des durations : idéalement, on aimerait rendre la valeur de notre entreprise insensible à son environnement. Or la complexité de chaque partie du bilan et de leurs interactions rend impossible toute étude analytique : le passif comme l'actif ont des comportements très peu linéaires, et sont sujets à des effets de seuil (déclenchement de rachats, changement de stratégie, etc.) rendant le tout imprévisible et peu intuitif. On a alors recours à la modélisation stochastique.

Dans un premier temps, on va chercher à modéliser l'environnement de l'entreprise. Pour le passif, il s'agit essentiellement des clients : lois de rachat, longévité, catastrophes... Le comportement de l'actif pourrait se décomposer en trois parties : actions, taux, et crédit, avec d'éventuelles corrélations. Par-dessus intervient alors la réaction de l'entreprise : changements dans les investissements, participation aux bénéfices. Enfin, il s'agit de calculer les indicateurs permettant le pilotage de l'entreprise : ratio de solvabilité, VIF...

La modélisation étant effectuée, on peut alors simuler. Typiquement, plusieurs milliers de trajectoires sont générées via un outil unique, et ce dans différents scénarios (baisse des taux, hausse des taux, des actions, rachats massifs...). Ces scénarios permettent de rendre compte de la sensibilité de la valeur de l'entreprise à différents paramètres, et du besoin en fonds propres dans chaque cas.

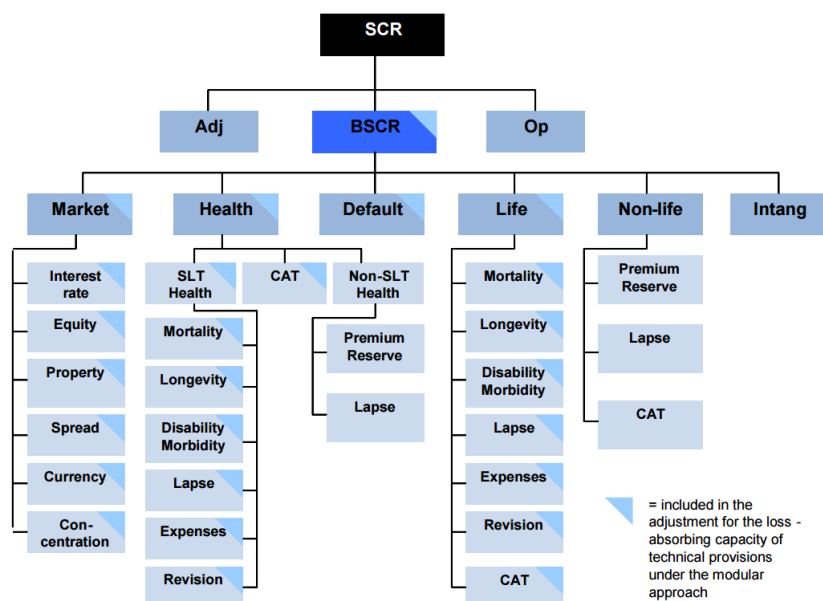


FIGURE 1 – Schéma de la décomposition du calcul du SCR.
Source : *eiopa.europa.eu*

1.3 Indicateurs

1.3.1 VIF, BEL

La VIF (*Value of In-Force business*) est la valeur actuelle probable des profits futurs nets d'impôts, calculée sur la base des contrats déjà souscrits. Il peut se calculer comme la différence entre les cash-flows générés par ces contrats, et le coût du capital qui sera immobilisé à ce titre.

Le BEL (*Best Estimate Liability*) est la valeur actuelle probable du portefeuille actuel de contrats.

Le capital disponible est la différence entre l'actif de l'entreprise, en valeur de marché, retiré de la valeur du passif (BEL + une marge pour risques non couvrables).

1.3.2 Ratio de solvabilité

Le SCR est la somme du BSCR, du SCR opérationnel, et d'un ajustement. Le BSCR se calcule en décomposant le risque de la compagnie en *modules* (vie, non-vie, marché, santé, actifs immatériels), eux-mêmes se déclinant en *sous-modules* : taux, actions, spread, liquidité, longévité, etc. (cf. figure 1). Le SCR pour chaque sous-module se calcule soit par formule fermée, soit via une approche par scénario.

Dans l'approche par scénario, on applique un choc (déterministe) correspondant au sous-module (baisse des taux, hausse des actions, hausse de la mortalité, etc.), et à une probabilité de 0.5%. Le SCR de ce sous-module est alors la différence en valeur absolue du capital disponible avec le scénario central (sans choc).

Les différents SCR de sous-modules sont agrégés par des matrices de corrélations en SCR de modules, puis en BSCR :

$$SCR_{TOT} = \sqrt{\sum_{i,j} \rho_{i,j} SCR_i SCR_j}$$

L'agrégation de cette sorte permet de rendre compte de la diversification du risque.

2 Les taux

Cette sous-partie donne un aperçu de la modélisation des taux. On y rappelle d'abord le formalisme qui permet un traitement mathématique, et quelques produits dérivés de taux courants. Viennent ensuite quelques méthodes de valorisation de ces produits, et des modèles classiques de taux d'intérêt.

2.1 Différents taux

Il convient de distinguer différents types de taux d'intérêt. Tous s'appuient sur la notion de zéro-coupon : une obligation zéro-coupon de maturité T est un titre assurant la livraison de 1€ à son détenteur à la date T . Son prix à l'instant t , donné par le marché, est noté $P(t, T)$. On définit alors :

- Le taux linéaire, ou taux LIBOR : $L(t, T) = \frac{1}{T-t} \frac{1-P(t, T)}{P(t, T)}$
- Le taux exponentiel, ou taux continu : $R(t, T) = -\frac{1}{T-t} \ln P(t, T)$
- Le taux actuariel : $Y(t, T) = P(t, T)^{-\frac{1}{T-t}} - 1$

Enfin, le taux court r_t est leur limite commune lorsque T tend vers t . C'est avec les taux $L(t, T)$ et r_t que nous travaillerons dans ce mémoire.

2.2 Produits de taux

Outre les obligations zéro-coupons, il existe aussi des obligations à coupons. Si une obligation délivre aux temps $T_1 < T_2 < \dots < T_N$ les coupons C_1, \dots, C_N , alors elle vaut en t : $\sum_{i=1}^N C_i P(t, T_i) \mathbb{I}\{t < T_i\}$.

Le *Forward Rate Agreement* ou FRA est un contrat par lequel deux parties s'accordent à échanger un taux LIBOR contre un taux fixe K entre deux dates T et $T + \theta$. Le flux, échangé à $T + \theta$, est connu en T et vaut $(L(T, T + \theta) - K)\theta$.

On définit alors le taux LIBOR à terme (ou *forward*) $L(t, T, T + \theta)$ comme le strike annulant le prix de ce contrat à t : $L(t, T, T + \theta) = \frac{1}{T-t} \frac{P(t, T) - P(t, T + \theta)}{P(t, T + \theta)}$. En découle le taux forward instantané $f(t, T) = \lim_{\theta \rightarrow 0} L(t, T, T + \theta)$.

Un swap est une somme de FRA de même strike K , sur chaque intervalle $[T_{i-1}, T_i]$ ($\alpha < i \leq \beta$). Le strike qui donne son juste prix au contrat (i.e. tel que sa valeur lors de la signature est nulle) est appelé le taux swap. Il vaut :

$$S_{\alpha, \beta}(t) = \frac{P(t, T_\alpha) - P(t, T_\beta)}{\sum_{i=\alpha+1}^{\beta} (T_i - T_{i-1}) P(t, T_i)}$$

Il est appelé swap payeur pour celui qui reçoit la jambe variable et paie la jambe fixe, et receveur pour l'autre entité.

Un cap est un swap payeur dont on ne retient que les flux positifs. Un caplet est un cap ne comportant qu'un flux. Ainsi un cap peut s'écrire comme une somme de caplets.

Les floors et floorlets se définissent de la même manière, à partir des swaps receveurs.

Une swaption est un contrat donnant la possibilité à son acheteur d'entrer dans un swap d'échéancier et de strike fixés.

2.3 Pricing des dérivés

2.3.1 Relation fondamentale

Tout contrat délivrant un flux Φ à la date T possède un unique juste prix en $t < T$ qui vaut : $\mathbb{E}_t^Q \left[e^{-\int_t^T r_s ds} \Phi \right]$ où Q désigne la mesure martingale ou risque-neutre. Par définition, c'est une mesure telle que pour tout actif échangeable S , le processus $\left(S_t e^{-\int_0^t r_s ds} \right)_{t \geq 0}$ soit une martingale. Elle existe dès lors qu'il n'y a pas d'opportunité d'arbitrage, et est unique si le marché est complet.

On en déduit par exemple le prix du zéro-coupon : $P(t, T) = \mathbb{E}_t^Q \left[e^{-\int_t^T r_s ds} \right]$.

De même, les caplets :

$$\begin{aligned} C(t, T, T + \theta, K) &= \mathbb{E}_t^Q \left[e^{-\int_t^T r_s ds} P(T, T + \theta) \theta (L(T, T + \theta) - K)_+ \right] \\ &= \mathbb{E}_t^Q \left[e^{-\int_t^T r_s ds} (1 - P(T, T + \theta) - K \theta P(T, T + \theta))_+ \right] \end{aligned}$$

2.3.2 Différentes mesures

Pour évaluer plus simplement les produits dérivés de taux, on sera amené à changer de mesure de probabilité. On cherchera à simplifier l'expression à évaluer, et à faire apparaître des variables dont les lois sont connues (typiquement, des martingales gaussiennes ou log-normales).

Dans ce mémoire, nous utiliserons donc la mesure risque-neutre Q , associée au numéraire $\left(e^{\int_0^t r_s ds} \right)_{t \geq 0}$, la mesure dite T -forward associée au numéraire $(P(t, T))_{0 \leq t \leq T}$, et, un échéancier $T_\alpha < \dots < T_\beta$ étant préalablement défini, la mesure LVL associée au numéraire $LVL(t) = \sum_{i=\alpha+1}^\beta (T_i - T_{i-1}) P(t, T_i)$.

Par exemple, sous la mesure $T + \theta$ -forward, le prix du caplet s'écrit alors :

$$\begin{aligned} C(t, T, T + \theta, K) &= \mathbb{E}_t^{T+\theta} \left[\frac{P(t, T + \theta) e^{\int_0^T r_s ds}}{P(T, T + \theta) e^{\int_0^T r_s ds}} e^{-\int_t^T r_s ds} (1 - P(T, T + \theta) - K \theta P(T, T + \theta))_+ \right] \\ &= P(t, T + \theta) \mathbb{E}_t^{T+\theta} \left[\left(\frac{P(t, T)}{P(T, T + \theta)} - (1 + K \theta) \right)_+ \right] \end{aligned}$$

et sous cette même mesure, $\left(\frac{P(t, T)}{P(T, T + \theta)} \right)_{0 \leq t \leq T}$ est une martingale. Il nous suffit alors de connaître sa volatilité pour calculer le prix du caplet.

De même le prix de la swaption se calcule en se plaçant sous la mesure LVL . En supposant les dates régulières (pour tout i , $T_i - T_{i-1} = \theta$) :

$$\begin{aligned} \text{Swaption}_{\alpha, \beta}(t) &= \mathbb{E}_t^Q \left[e^{-\int_t^{T_\alpha} r_s ds} \left(\sum_{i=\alpha+1}^\beta \theta P(T_\alpha, T_i) (L(T_\alpha, T_{i-1}, T_i) - K) \right)_+ \right] \\ &= LVL(t) \mathbb{E}_t^{LVL} \left[(S_{\alpha, \beta}(T_\alpha) - K)_+ \right] \end{aligned}$$

et $S_{\alpha, \beta}$ est bien une martingale sous la mesure LVL .

2.4 Modèles classiques de taux

Afin de pouvoir calculer les espérances écrites plus haut, il faudra faire des hypothèses sur la dynamique des taux. C'est là tout l'objet des modèles de taux, que l'on peut classer en trois grandes catégories :

- Les modèles de taux court, qui supposent simplement une dynamique du processus $(r_t)_{t \geq 0}$.
- Les modèles de marché (*market models*), plus raffinés et qui diffusent une série de taux donnés en en partant de leurs observations à $t = 0$
- Les modèles HJM, qui peuvent se voir comme une généralisation des deux autres classes de modèles, qui diffusent la courbe des taux entière

On peut aussi citer d'autres types de modèles comme les séries chronologiques, le SABR, les modèles *shadow rate*... mais ceux-là sont, dans le contexte des compagnies d'assurances, peu ou pas utilisés.

Nous présentons ici deux modèles fondamentaux de taux courts, dont le CIR qui est réutilisé par la suite pour modéliser la volatilité. Nous introduisons ensuite le modèle LMM, et son extension LMM+ qui est un modèle de référence dans le contexte assurantiel, et qui est actuellement utilisé par Predica. Ce dernier servira d'élément de comparaison pour notre nouveau modèle.

2.4.1 Modèle de Vašíček

Le modèle de Vašíček, publié en 1977, est un des premiers modèles stochastiques de taux d'intérêt. Il s'agit d'un modèle de taux court à retour à la moyenne. Il suppose la diffusion suivante :

$$dr_t = \kappa(\theta - r_t)dt + \sigma dW_t^Q$$

où $(W_t^Q)_{t \geq 0}$ est un processus de Wiener sous la probabilité risque-neutre Q , $\theta \in \mathbb{R}$, $r(0) = r_0 \in \mathbb{R}$ et $\sigma, \kappa > 0$. L'expression précédente s'intègre aisément en développant $d(r_t e^{\kappa t})$:

$$r_t = \theta + (r_s - \theta)e^{-\kappa(t-s)} + \sigma \int_s^t e^{-\kappa(t-u)} dW_u^Q$$

Le taux court r_t , conditionnellement à \mathcal{F}_s , est donc gaussien. Ses deux premiers moments valent :

$$\begin{aligned} \mathbb{E}_s^Q [r_t] &= \theta + (r_s - \theta)e^{-\kappa(t-s)} \\ &\xrightarrow[t \rightarrow +\infty]{} \theta \\ \mathbb{V}_s^Q [r_t] &= \frac{\sigma^2}{2\kappa} \left(1 - e^{-2\kappa(t-s)}\right) \\ &\xrightarrow[t \rightarrow +\infty]{} \frac{\sigma^2}{2\kappa} \end{aligned}$$

L'intégrale de r_t est elle-même gaussienne et vérifie :

$$\begin{aligned} \mathbb{E}_t^Q \left[\int_t^T r_s ds \right] &= \theta(T-t) + (r_t - \theta) \frac{1 - e^{-\kappa(T-t)}}{\kappa} \\ \mathbb{V}_t^Q \left[\int_t^T r_s ds \right] &= \frac{\sigma^2}{\kappa^3} \left[\kappa(T-t) - 2 \left(1 - e^{-\kappa(T-t)}\right) + \frac{1}{2} \left(1 - e^{-2\kappa(T-t)}\right) \right] \end{aligned}$$

On en déduit alors le prix du zéro-coupon :

$$\begin{aligned} P(t, T) &= \mathbb{E}_t^Q \left[e^{-\int_t^T r_s ds} \right] \\ &= \exp \left(-\mathbb{E}_t^Q \left[\int_t^T r_s ds \right] + \frac{1}{2} \mathbb{V}_t^Q \left[\int_t^T r_s ds \right] \right) \\ &= e^{A(t, T) - r_t B(t, T)} \end{aligned}$$

où les fonctions A et B suivent les notations usuelles de la littérature :

$$\begin{aligned} B(t, T) &= \frac{1 - e^{-\kappa(T-t)}}{\kappa} \\ A(t, T) &= \left(\theta - \frac{\sigma^2}{2\kappa^2} \right) (B(t, T) - (T - t)) - \frac{\sigma^2}{4\kappa} B(t, T)^2 \end{aligned}$$

Et :

$$\frac{dP(t, T)}{P(t, T)} = r_t dt - \sigma B(t, T) dW_t^Q$$

$$\text{Donc } d\frac{P(t, T)}{P(t, S)} = \frac{P(t, T)}{P(t, S)} \sigma (B(t, S) - B(t, T)) dW_t^S.$$

Un caplet vaut ainsi

$$\begin{aligned} C(t, T, S, K) &= P(t, S) \mathbb{E}^S \left[\left(\frac{P(T, T)}{P(T, S)} - (1 + (S - T)K) \right)_+ \middle| \mathcal{F}_t \right] \\ &= P(t, S) \text{Black} \left(\frac{P(t, T)}{P(t, S)}, 1 + (S - T)K, \sigma_{t, T, S} \right) \\ \sigma_{t, T, S}^2 &= \int_t^T (B(u, S) - B(u, T))^2 du \\ &= \frac{\sigma^2}{2\kappa} B(0, \theta)^2 \left(1 - e^{-2\kappa(T-t)} \right) \end{aligned}$$

2.4.2 Modèle de Hull et White

Les trois paramètres constants du modèle précédent ne permettent pas de reproduire correctement la courbe des taux. Pour pallier ce problème, Hull et White proposent de remplacer la moyenne à long terme θ par une fonction déterministe $t \mapsto \theta_t$ qui serait calibrée sur la courbe des taux du marché.

En procédant de la même manière :

$$r_t = r_s e^{-\kappa(t-s)} + \kappa \int_s^t e^{-\kappa(t-u)} \theta_u du + \sigma \int_s^t e^{-\kappa(t-u)} dW_u^Q$$

puis

$$\begin{aligned}\mathbb{E}_t^Q \left[\int_t^T r_s ds \right] &= r_t \frac{1 - e^{-\kappa(T-t)}}{\kappa} + \int_t^T (1 - e^{-\kappa(T-s)}) \theta_s ds \\ \mathbb{V}_t^Q \left[\int_t^T r_s ds \right] &= \frac{\sigma^2}{\kappa^3} \left[\kappa(T-t) - 2 \left(1 - e^{-\kappa(T-t)} \right) + \frac{1}{2} \left(1 - e^{-2\kappa(T-t)} \right) \right]\end{aligned}$$

On en déduit le prix des zéro-coupons, avec lesquels on peut se calibrer sur le marché.

2.4.3 Modèle de Cox, Ingersoll et Ross

Le modèle CIR est un autre modèle de taux court, introduit pour pallier ce qui était vu comme un inconvénient dans le modèle de Vašíček : la possible négativité de r_t . On s'en sert plutôt aujourd'hui pour modéliser des volatilités stochastiques (et c'est exactement ce qui sera fait dans la deuxième partie).

En imposant cette fois θ et r_0 strictement positifs, on définit r par l'équation différentielle stochastique suivante :

$$dr_t = \kappa(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t^Q$$

Le processus présente donc un retour à la moyenne, et le terme $\sqrt{r_t}$ devant le bruit lui permet de rester positif, s'il vérifie la condition de Feller $\frac{2\kappa\theta}{\sigma^2} \geq 1$ (voir annexes).

On ne connaît pas de formule fermée exprimant r_t , mais on connaît sa loi et la transformée de Laplace du couple $\left(r_t, \int_0^t r_s ds\right)_{t \geq 0}$. En effet la forme de l'équation différentielle nous assure que la transformée peut se mettre sous la forme

$$F_{\lambda,\mu}(t, r_0) = \mathbb{E} \left[e^{-\lambda r_t} e^{-\mu \int_0^t r_s ds} \right] = e^{-\kappa\theta\varphi_t(\lambda,\mu) - r_0\psi_t(\lambda,\mu)}$$

Le processus $\left(e^{-\mu \int_0^t r_s ds} F_{\lambda,\mu}(T-t, r_t)\right)_{0 \leq t \leq T}$ est une martingale à la condition que $F_{\lambda,\mu}$ soit solution de :

$$\begin{aligned}\frac{\partial F_{\lambda,\mu}}{\partial t} &= -\mu r F_{\lambda,\mu} + \kappa(\theta - r) \frac{\partial F_{\lambda,\mu}}{\partial r} + \frac{\sigma^2 r}{2} \frac{\partial^2 F_{\lambda,\mu}}{\partial r^2} \\ F_{\lambda,\mu}(0, r) &= e^{-\lambda r}\end{aligned}$$

c'est-à-dire :

$$\begin{aligned}-\psi' &= \frac{\sigma^2}{2} \psi^2 + \kappa\psi - \mu \\ \varphi' &= \psi\end{aligned}$$

La résolution de ces équations en φ et ψ donne l'expression de $F_{\lambda,\mu}$:

$$F_{\lambda,\mu} = \left[\frac{2\gamma e^{t\frac{\gamma+\kappa}{2}}}{\sigma^2 \lambda (e^{\gamma t} - 1) + \gamma - \kappa + e^{\gamma t}(\gamma + \kappa)} \right]^{2\frac{\kappa\theta}{\sigma^2}} \\ \times \exp \left[-r_0 \frac{\lambda(\gamma + \kappa + e^{\gamma t}(\gamma - \kappa)) + 2\mu(e^{\gamma t} - 1)}{\sigma^2 \lambda (e^{\gamma t} - 1) + \gamma - \kappa + e^{\gamma t}(\gamma + \kappa)} \right] \\ \gamma = \sqrt{\kappa^2 + 2\mu\sigma^2}$$

En particulier, avec $\mu = 0$, on reconnaît la transformée de Laplace d'une loi du χ^2 décentrée, de paramètre de décentrage $\frac{4r_0\kappa}{\sigma^2(e^{\kappa t}-1)}$ et avec $\frac{4\kappa\theta}{\sigma^2}$ degrés de liberté. C'est donc la loi de r_t .

On déduit de $F_{\lambda,\mu}$ le prix du zéro-coupon dans ce modèle :

$$P(t, T) = \mathbb{E}_t \left[e^{-\int_t^T r_s ds} \right] \\ = F_{\lambda=0, \mu=1}(T-t, r_t) \\ = \left[\frac{2\gamma e^{\frac{1}{2}(T-t)(\gamma+\kappa)}}{\gamma - \kappa + e^{\gamma(T-t)}(\gamma + \kappa)} \right]^{2\frac{\kappa\theta}{\sigma^2}} \exp \left[-2r_t \frac{e^{\gamma(T-t)} - 1}{\gamma - \kappa + e^{\gamma(T-t)}(\gamma + \kappa)} \right] \\ \gamma = \sqrt{\kappa^2 + 2\sigma^2}$$

ainsi que sa dynamique :

$$\frac{dP(t, T)}{P(t, T)} = r_t dt - 2 \frac{e^{\gamma(T-t)} - 1}{\gamma - \kappa + e^{\gamma(T-t)}(\gamma + \kappa)} \sigma \sqrt{r_t} dW_t^Q$$

Les modèles de taux court ont ce défaut que de la simple dynamique de r_t se déduit la courbe des taux entière. Si la différence entre la courbe marché et la courbe modèle est trop grande, on se heurte à un problème d'arbitrabilité du modèle, que l'on peut pallier en remplaçant certains paramètres constants (typiquement la moyenne à long terme θ) par une fonction du temps. Les modèles de marché diffusent une série de taux donnée, ce qui comble aussi ce défaut.

2.4.4 Le LIBOR Market Model

Plutôt que de faire des hypothèses sur le taux court, qui contient assez peu d'informations, on se donne dans le cadre du LMM une famille de n taux LIBOR forward $L_j(t) = L(t, T_{j-1}, T_j)$ ($T_0 < \dots < T_n$) et on fait l'hypothèse suivante sur leur dynamique :

$$\frac{dL_j(t)}{L_j(t)} = \sigma_t^j dW_t^{j, T_j}$$

le processus $(W_t^{T_j})_{t \geq 0}$ étant un mouvement brownien sous la probabilité T_j -forward, et $d\langle W_t^i, W_t^j \rangle = \rho^{i,j} dt$.

On a donc naturellement les prix des caplets :

$$C(t, T_{i-1}, T_i, K) = P(0, T_i) \text{Black} \left(L_i(0), K, \sqrt{\int_t^{T_{i-1}} \sigma_s^i ds} \right)$$

où la formule de Black est définie par :

$$\begin{aligned}\text{Black}(S_0, K, \sigma) &= S_0 \ln(d_1) - K \ln(d_2) \\ d_1 &= \frac{1}{\sigma} \left[\ln \left(\frac{S_0}{K} \right) + \frac{\sigma^2}{2} \right] \\ d_2 &= \frac{1}{\sigma} \left[\ln \left(\frac{S_0}{K} \right) - \frac{\sigma^2}{2} \right]\end{aligned}$$

2.4.5 LMM+

Ce modèle est une extension du LMM. Il présente un décalage rendant possible des taux négatifs (mais minorés), et dont les volatilités des taux suivent un processus CIR :

$$\begin{aligned}\frac{dL_j(t)}{\delta + L_j(t)} &= \mu_j(t)dt + \vec{\sigma}_{t,T} \cdot d\vec{W}_t \\ \vec{\sigma}_{t,T} \cdot d\vec{W}_t &= \sqrt{V_t} g_{j-m(t)} \sum_{i=1}^N \beta_{j-m(t)}^i dW_t^i \\ g_k &= (a + bT_k) e^{-cT_k} \\ dV_t &= \kappa(\theta - V_t)dt + \varepsilon \sqrt{V_t} dZ_t \\ d \left\langle \frac{1}{\sqrt{N}} \sum_{i=1}^N W_t^i, Z_t \right\rangle &= \rho dt\end{aligned}$$

où $[t] = \min\{T_1, T_2, \dots\} \cap [t, +\infty[$, les processus W^i et Z sont indépendants.

Le décalage le rend adapté au contexte de taux négatifs actuel. Et, comme dans les modèle actions, il a été prouvé que la volatilité des taux étaient variable, et de manière aléatoire. Il semble donc naturel de considérer des modèles à volatilité stochastique comme celui-ci.

Aujourd'hui, il constitue un modèle de référence pour la modélisation des taux au sein des compagnies d'assurance. Il est notamment au cœur du générateur de scénarios économique distribué par l'agence Moody's, largement présent sur le marché français. Predica, suite aux récentes baisses des taux, a voulu remettre en question ce modèle, utilisé pour toutes ses analyses quantitatives dont son calcul de ratio de solvabilité. Il a semblé naturel de se pencher vers une nouvelle classe de modèle. Ceux de taux court étant trop limités, nous nous sommes tournés vers un modèle type HJM.

Deuxième partie

Un modèle HJM avec volatilité stochastique

1 Présentation

1.1 Le cadre HJM

Heath, Jarrow et Morton publièrent en 1992 un article décrivant un vaste cadre d'étude des taux d'intérêt [2]. Il part de l'hypothèse suivante sur les taux à terme :

$$df(t, T) = \alpha^P(t, T)dt + \sum_{i=1}^N \sigma_i(t, T)dW_i^P(t)$$

où α^P et σ peuvent être stochastiques, et vérifient certaines conditions d'intégrabilité. Si la mesure martingale Q existe de manière unique (ce qui est le cas des marchés complets et sans arbitrage), il est démontré que, sous cette mesure de probabilité :

$$df(t, T) = \sum_{i=1}^N \sigma_i(t, T) \int_t^T \sigma_i(t, u)du + \sum_{i=1}^N \sigma_i(t, T)dW_i^Q(t)$$

De ces larges hypothèses découlent une grande variété de modèles de taux, dont le modèle présent, introduit par Eduardo Schwartz et Anders Trolle.

1.2 Forme générale

Schwartz et Trolle proposent dans leur article [1] la diffusion suivante :

$$df(t, T) = \mu(t, T)dt + \sum_{i=1}^N \sigma_{f,i}(t, T)\sqrt{v_i(t)}dW_i^Q(t)$$

$$dv_i(t) = \kappa_i(\theta_i - v_i(t))dt + \sigma_i\sqrt{v_i(t)}\left(\rho_i dW_i^Q(t) + \sqrt{1 - \rho_i^2}dZ_i^Q(t)\right)$$

avec pour structure de volatilité :

$$\sigma_{f,i}(t, T) = (\alpha_0^i + \alpha_1^i(T - t))e^{-\gamma_i(T-t)}$$

tous les processus de Wiener W_i^Q et Z_i^Q étant indépendants.

1.3 Taux forward

La condition de non-arbitrage du cadre HJM impose :

$$\begin{aligned}
\mu(t, T) &= \sum_{i=1}^N \sqrt{v_i(t)} \sigma_{f,i}(t, T) \int_t^T \sqrt{v_i(u)} \sigma_{f,i}(t, u) du \\
&= \sum_{i=1}^N v_i(t) (\alpha_0^i + \alpha_1^i (T-t)) e^{-\gamma_i(T-t)} \left[\left(\frac{\alpha_0^i}{\gamma_i} + \frac{\alpha_1^i}{\gamma_i^2} \right) (1 - e^{-\gamma_i(T-t)}) - \frac{\alpha_1^i}{\gamma_i} (T-t) e^{-\gamma_i(T-t)} \right]
\end{aligned}$$

On en déduit alors l'expression de $f(t, T)$: sachant que

$$f(t, T) - f(0, T) = \int_0^t df(u, T)$$

on aboutit à l'expression suivante :

$$f(t, T) = f(0, T) + \sum_{i=1}^N B_i(T-t) x_i(t) + \sum_{i=1}^N \sum_{j=1}^6 D_i^j(T-t) \varphi_{i,j}(t)$$

Les fonctions B_i et D_i^j étant déterministes, les processus $\varphi_{i,j}$ étant à variations bornées, et x_i étant des semi-martingales continues. Ces fonctions x_i et $\varphi_{i,j}$ sont caractérisées par leur valeur initiale nulle et leurs équations de diffusion (voir annexes).

1.4 Prix des zéro-coupons

Pour obtenir le prix des zéro-coupons, on intègre d'abord l'expression de $f(t, T)$ selon la maturité :

$$\begin{aligned}
\int_t^T f(t, s) ds &= \int_t^T f(0, s) ds - \sum_{i=1}^N \mathcal{B}_i(T-t) x_i(t) - \sum_{i=1}^N \sum_{j=1}^6 \mathcal{D}_i^j(T-t) \varphi_{i,j}(t) \\
\mathcal{B}_i(\tau) &= - \int_0^\tau B_i(s) ds \\
\mathcal{D}_i^j(\tau) &= - \int_0^\tau D_i^j(s) ds
\end{aligned}$$

Les expressions des \mathcal{B}_i et \mathcal{D}_i^j sont précisées en annexe.

On peut alors exprimer les zéro-coupons par :

$$P(t, T) = \frac{P(0, T)}{P(0, t)} \exp \left[\sum_{i=1}^N \mathcal{B}_i(T-t) x_i(t) + \sum_{i=1}^N \sum_{j=1}^6 \mathcal{D}_i^j(T-t) \varphi_{i,j}(t) \right]$$

On en déduit leur dynamique :

$$\frac{dP(t, T)}{P(t, T)} = r_t dt + \sum_{i=1}^N \mathcal{B}_i(T-t) \sqrt{v_i(t)} dW_t^i$$

Ainsi les paramètres $\alpha_0^i, \alpha_1^i, \kappa_i, \theta_i, \gamma_i, \sigma_i, \rho_i$ sont fixés par le modèle, et permettent de diffuser les fonctions d'état $x_i, v_i, \varphi_{i,j}$ et les facteurs de risques (indépendants) W_i et Z_i , ce qui permet par la suite de calculer les taux forward et les prix des zéro-coupons.

2 Options sur zéro-coupons

La calibration risque-neutre du modèle consiste en l'ajustement aux prix de produits dérivés, ce qui inclut les options d'achat et de vente sur les zéro-coupons. On détaille ici le calcul du prix d'un put (option de vente) sur zéro-coupon, le call (option d'achat) se déduisant par parité. Comme nous le verrons ensuite, on pourra calculer le prix d'une swaption grâce aux puts sur zéro-coupons.

2.1 Fonction caractéristique Ψ

On cherche à estimer le prix d'un put de maturité T sur un zéro-coupon d'échéance S :

$$\mathcal{P}(t, T, S, K) = \mathbb{E}_t^Q \left[e^{-\int_t^T r_s ds} (P(T, S) - K)_+ \right]$$

Introduisons la fonction caractéristique de $\ln P(T, S)$ avec actualisation :

$$\Psi(u, t, T, S) = \mathbb{E}_t^Q \left[e^{-\int_t^T r_s ds} e^{u \ln P(T, S)} \right]$$

Sachant que $\Psi(u, T, T, S) = \mathbb{E}_T^Q \left[e^{-\int_T^T r_s ds} e^{u \ln P(T, S)} \right] = e^{u \ln P(T, S)}$, on peut alors écrire que $e^{-\int_0^t r_s ds} \Psi(u, t, T, S) = \mathbb{E}_t^Q \left[e^{-\int_0^T r_s ds} \Psi(u, T, T, S) \right]$. Autrement dit, $(\tilde{\Psi}(u, t, T, S))_{0 \leq t \leq T} = \left(e^{-\int_0^t r_s ds} \Psi(u, t, T, S) \right)_{0 \leq t \leq T}$ est une martingale.

Supposons que Ψ puisse s'écrire sous la forme suivante :

$$\Psi(u, t, T, S) = \exp \left[M(T-t) + \sum_{i=1}^N N_i(T-t) v_i(t) + u \ln P(t, S) + (1-u) \ln P(t, T) \right]$$

Appliquons à $\tilde{\Psi}$ la formule d'Itô :

$$\begin{aligned} \frac{d\tilde{\Psi}}{\tilde{\Psi}} &= -r_t dt - \frac{\partial M}{\partial \tau}(T-t)dt + \sum_{i=1}^N \left(-\frac{\partial N_i}{\partial \tau}(T-t)v_i(t)dt + N_i(T-t)dv_i(t) \right) + u \frac{dP(t, S)}{P(t, S)} \\ &\quad + (1-u) \frac{dP(t, T)}{P(t, T)} + \frac{u(u-1)}{2} \frac{d\langle P(t, S) \rangle}{P(t, S)^2} + \frac{u(u-1)}{2} \frac{d\langle P(t, T) \rangle}{P(t, T)^2} + u(1-u) \frac{d\langle P(t, S), P(t, T) \rangle}{P(t, S)P(t, T)} \\ &\quad + \sum_{i=1}^N \frac{1}{2} N_i^2(T-t) d\langle v_i(t) \rangle + \frac{u N_i(T-t)}{P(t, S)} d\langle v_i(t), P(t, S) \rangle + \frac{(1-u) N_i(T-t)}{P(t, T)} d\langle v_i(t), P(t, T) \rangle \\ &\quad + \sum_{i=1}^N (\dots) dW_t^i + \sum_{i=1}^N (\dots) dZ_t^i \\ &= -r_t dt - \frac{\partial M}{\partial \tau}(T-t)dt + \sum_{i=1}^N \left(-\frac{\partial N_i}{\partial \tau}(T-t)v_i(t)dt + N_i(T-t)\kappa_i(\theta_i - v_i(t))dt \right) \\ &\quad + r_t dt + \sum_{i=1}^N \left[\frac{u(u-1)}{2} \mathcal{B}_i^2(S-t) + \frac{u(u-1)}{2} \mathcal{B}_i^2(T-t) + u(1-u)\mathcal{B}_i(S-t)\mathcal{B}_i(T-t) \right] v_i(t)dt \end{aligned}$$

$$\begin{aligned}
& + \sum_{i=1}^N \left[\frac{1}{2} N_i^2(T-t) \sigma_i^2 + u N_i(T-t) \rho_i \sigma_i \mathcal{B}_i(S-t) + (1-u) N_i(T-t) \rho_i \sigma_i \mathcal{B}_i(T-t) \right] v_i(t) dt \\
& + \sum_{i=1}^N (\dots) dW_t^i + \sum_{i=1}^N (\dots) dZ_t^i
\end{aligned}$$

Sachant que $\tilde{\Psi}$ est une martingale, son drift est nul. En regroupant les terme en dt et $v_i(t)dt$ qui sont tous nuls (car leur somme vaut zéro quelles que soient les valeurs des différents $v_i(t)$), on obtient $N+1$ équations de Riccati :

$$\begin{aligned}
0 &= -\frac{\partial M}{\partial \tau}(T-t) + \sum_{i=1}^N N_i(T-t) \kappa_i \theta_i \\
0 &= -\frac{\partial N_i}{\partial \tau}(T-t) - N_i(T-t) \kappa_i + \frac{u(u-1)}{2} \mathcal{B}_i^2(S-t) + \frac{u(u-1)}{2} \mathcal{B}_i^2(T-t) + u(1-u) \mathcal{B}_i(S-t) \mathcal{B}_i(T-t) \\
&+ \frac{1}{2} N_i^2(T-t) \sigma_i^2 + u N_i(T-t) \rho_i \sigma_i \mathcal{B}_i(S-t) + (1-u) N_i(T-t) \rho_i \sigma_i \mathcal{B}_i(T-t)
\end{aligned}$$

Ce qui revient à :

$$\begin{aligned}
\frac{\partial M}{\partial \tau}(\tau) &= \sum_{i=1}^N \kappa_i \theta_i N_i(\tau) \\
\frac{\partial N_i}{\partial \tau}(\tau) &= N_i(\tau) [\rho_i \sigma_i (u \mathcal{B}_i(\Theta + \tau) + (1-u) \mathcal{B}_i(\tau)) - \kappa_i] \\
&+ \frac{1}{2} N_i^2(\tau) \sigma_i^2 + \frac{u(u-1)}{2} [\mathcal{B}_i(\Theta + \tau) - \mathcal{B}_i(\tau)]^2
\end{aligned}$$

où $\tau = T-t$ et $\Theta = S-T$.

De plus $\Psi(u, T, T, S) = \exp \left[M(0) + \sum_{i=1}^N N_i(0) v_i(T) + u \ln P(T, S) \right] = e^{u \ln P(T, S)}$ et ce, quels que soient les valeurs des $v_i(T)$. On a alors la condition initiale $M(0) = N_i(0) = 0$.

2.2 Prix de l'option

Revenons à notre put de strike K :

$$\begin{aligned}
\mathcal{P}(t, T, S, K) &= \mathbb{E}_t^Q \left[e^{-\int_t^T r_s ds} (K - P(T, S))_+ \right] \\
&= \mathbb{E}_t^Q \left[e^{-\int_t^T r_s ds} K \mathbb{I}_{P(T, S) \leq K} \right] - \mathbb{E}_t^Q \left[e^{-\int_t^T r_s ds} P(T, S) \mathbb{I}_{P(T, S) \leq K} \right] \\
&= K \mathbb{E}_t^Q \left[e^{-\int_t^T r_s ds} \mathbb{I}_{X \leq k} \right] - \mathbb{E}_t^Q \left[e^{-\int_t^T r_s ds} e^X \mathbb{I}_{X \leq k} \right] \\
&= G_0(k) - G_1(k)
\end{aligned}$$

avec $X = \ln P(T, S)$, $k = \ln K$, et $G_\varepsilon(k) = \mathbb{E} \left[e^{-\int_t^T r_s ds} e^{\varepsilon X} \mathbb{I}_{X \leq k} \right]$.

Les fonctions G_ε étant croissantes, bornées et continues à droite, elles définissent une mesure de Stieltjes dont la transformée de Fourier vaut :

$$\mathcal{G}_\varepsilon(u) = \int_{-\infty}^{+\infty} e^{iuy} dG_\varepsilon(y)$$

En admettant l'existence de $f_{(X,R)}$, la densité de $(X, \int_t^T r_s ds)$, on peut exprimer la dérivée de G_ε par :

$$\begin{aligned} \frac{\partial G_\varepsilon}{\partial y} &= \frac{\partial}{\partial y} \int_{r \in \mathbb{R}} \int_{x=-\infty}^y e^{-r} e^{\varepsilon x} f_{X,R}(x, r) dx dr \\ &= \int_{r \in \mathbb{R}} e^{-r} e^{\varepsilon y} f_{X,R}(y, r) dr \end{aligned}$$

ce qui donne ensuite :

$$\begin{aligned} \mathcal{G}_\varepsilon(y) &= \mathbb{E}_t^Q \left[e^{-\int_t^T r_s ds} e^{(iy+\varepsilon)X} \right] \\ &= \Psi(iy + \varepsilon, t, T, S) \end{aligned}$$

Inverser cette transformée nous donne la densité associée à la mesure de Stieltjes définie par G . Et $\lim_{-\infty} G_\varepsilon = 0$, donc :

$$\begin{aligned} G_\varepsilon(y) &= \int_{x=-\infty}^y (\mathcal{F}^{-1} \mathcal{G}_\varepsilon)(x) dx \\ &= \frac{1}{2\pi} \int_{x < y} \int_{u \in \mathbb{R}} e^{-iux} \mathcal{G}_\varepsilon(u) du dx \\ &= \frac{1}{2\pi} \lim_{A \rightarrow +\infty} \int_{u \in \mathbb{R}} \mathcal{G}_\varepsilon(u) \int_{-A}^y e^{-iux} dx du \\ &= \frac{1}{2\pi} \int_{u \in \mathbb{R}} \mathcal{G}_\varepsilon(u) \frac{e^{-iuy}}{-iu} du - \frac{1}{2\pi} \lim_{A \rightarrow +\infty} \int_{u \in \mathbb{R}} \mathcal{G}_\varepsilon(u) \frac{e^{-iuA}}{-iu} du \\ &= \frac{i}{2\pi} \left[\int_0^{+\infty} \mathcal{G}_\varepsilon(u) \frac{e^{-iuy}}{u} du + \int_0^{+\infty} \mathcal{G}_\varepsilon(-u) \frac{e^{iuy}}{-u} du \right] + \frac{1}{2\pi} \lim_{A \rightarrow +\infty} \int_{u \in \mathbb{R}} \mathcal{G}_\varepsilon\left(\frac{u}{A}\right) \frac{e^{-iu}}{iu} du \\ &= \frac{i}{2\pi} \left[\int_0^{+\infty} \mathcal{G}_\varepsilon(u) \frac{e^{-iuy}}{u} du - \int_0^{+\infty} \overline{\mathcal{G}_\varepsilon(u)} \frac{e^{-iuy}}{u} du \right] + \frac{1}{2\pi} \int_{u \in \mathbb{R}} \lim_{A \rightarrow +\infty} \mathcal{G}_\varepsilon\left(\frac{u}{A}\right) \frac{e^{-iu}}{iu} du \\ &= \frac{i}{2\pi} 2i\Im \left[\int_0^{+\infty} \mathcal{G}_\varepsilon(u) \frac{e^{-iuy}}{u} du \right] + \frac{1}{2\pi} \mathcal{G}_\varepsilon(0) \int_{u \in \mathbb{R}} \frac{e^{-iu}}{iu} du \\ &= -\frac{1}{\pi} \int_0^{+\infty} \Im \left[\mathcal{G}_\varepsilon(u) \frac{e^{-iuy}}{u} \right] du + \frac{1}{2} \mathcal{G}_\varepsilon(0) \end{aligned}$$

Pour calculer le prix de l'option sur zéro-coupon, Schwartz et Trolle proposent donc de calculer la fonction Ψ en résolvant l'équation différentielle portant sur M et N_i avec un schéma de Runge-Kutta d'ordre 4, puis de l'intégrer via une quadrature de Gauss-Legendre (en tronquant à 8000, 20 points sur $[0, 1000]$ et 20 autres sur $[1000, 8000]$) pour en déduire G_0 , G_1 , et enfin \mathcal{P} .

3 Swaptions

3.1 Dynamiques sous la mesure LVL

En différenciant le taux swap, exprimé en fonction des zéro-coupons, et en se plaçant dans la mesure level :

$$\begin{aligned}
dS_{\alpha,\beta}(t) &= d \frac{P(t, T_\alpha) - P(t, T_\beta)}{LVL(t)} \\
&= \frac{dP(t, T_\alpha) - dP(t, T_\beta)}{LVL(t)} - \frac{P(t, T_\alpha) - P(t, T_\beta)}{LVL(t)^2} dLVL(t) + (\dots)dt \\
&= \frac{1}{LVL(t)} \sum_{i=1}^N [\mathcal{B}_i(T_\alpha - t)P(t, T_\alpha) - \mathcal{B}_i(T_\beta - t)P(t, T_\beta)] \sqrt{v_i(t)} dW_i^{LVL}(t) \\
&\quad - \frac{P(t, T_\alpha) - P(t, T_\beta)}{LVL(t)^2} \sum_{j=\alpha+1}^\beta (T_j - T_{j-1})P(t, T_j) \sum_{i=1}^N \mathcal{B}_i(T_j - t) \sqrt{v_i(t)} dW_i^{LVL}(t) \\
&= \sum_{i=1}^N \left[\sum_{j=\alpha+1}^\beta \zeta_j(t) \mathcal{B}_i(T_j - t) \right] \frac{\sqrt{v_i(t)}}{LVL(t)} dW_i^{LVL}(t) \\
\zeta_\alpha(t) &= P(t, T_\alpha) \\
\zeta_\beta(t) &= -P(t, T_\beta) - S_{\alpha,\beta}(t)(T_\beta - T_{\beta-1})P(t, T_\beta) \\
\zeta_j(t) &= -S_{\alpha,\beta}(t)(T_j - T_{j-1})P(t, T_j)
\end{aligned}$$

Plus généralement, si $(X_t)_{t>0}$ est une martingale sous la mesure risque-neutre, on se ramène à une martingale sous la mesure LVL en lui retirant la quantité

$$\left\langle X_t, \sum_{1 \leq i \leq N, \alpha < j \leq \beta} (T_j - T_{j-1}) \mathcal{B}_i(T_j - t) P(t, T_j) \frac{\sqrt{v_i(t)}}{LVL(t)} W_i^Q(t) \right\rangle$$

On peut par exemple exprimer la dynamique du processus v_i sous cette mesure :

$$\begin{aligned}
dv_i(t) &= \kappa_i(\theta_i - v_i(t))dt + \sigma_i \sqrt{v_i(t)} \left(\rho_i dW_i^Q(t) + \sqrt{1 - \rho_i^2} dZ_i^Q(t) \right) \\
&= \kappa_i(\theta_i - v_i(t))dt + \sigma_i \sqrt{v_i(t)} \left(\rho_i dW_i^{LVL}(t) + \sqrt{1 - \rho_i^2} dZ_i^{LVL}(t) \right) \\
&\quad + \sum_{1 \leq k \leq N, \alpha < j \leq \beta} (T_j - T_{j-1}) \mathcal{B}_k(T_j - t) P(t, T_j) \frac{\sqrt{v_k(t)}}{LVL(t)} \left\langle \rho_i \sigma_i \sqrt{v_i(t)} dW_i^Q(t), W_k^Q(t) \right\rangle \\
&= \left[\kappa_i(\theta_i - v_i(t)) + \sum_{\alpha < j \leq \beta} (T_j - T_{j-1}) \mathcal{B}_i(T_j - t) P(t, T_j) \frac{\rho_i \sigma_i v_i(t)}{LVL(t)} \right] dt \\
&\quad + \sigma_i \sqrt{v_i(t)} \left(\rho_i dW_i^{LVL}(t) + \sqrt{1 - \rho_i^2} dZ_i^{LVL}(t) \right)
\end{aligned}$$

3.2 Première approximation

Chaque terme $\frac{P(t, T_j)}{LVL(t)}$ est, sous la mesure LVL, une martingale. La meilleure approximation constante qu'on puisse lui faire est donc sa valeur actuelle, à $t = t_0$. Ainsi :

$$dS_{\alpha, \beta}(t) \approx \sum_{i=1}^N \left[\sum_{j=\alpha+1}^{\beta} \frac{\zeta_j(t_0)}{LVL(t_0)} \mathcal{B}_i(T_j - t) \right] \sqrt{v_i(t)} dW_i^{LVL}(t)$$

Sous cette approximation, le taux swap est alors gaussien conditionnellement aux v_i que l'on peut remplacer par leurs valeurs présentes.

On peut aller plus loin et remplacer la volatilité de $S_{\alpha, \beta}$ par son espérance moyennée sur l'intervalle de temps considéré :

$$\begin{aligned} dS_{\alpha, \beta} &\approx \sum_{i=1}^N \sqrt{T_{\alpha} - t_0} \sigma_i^{\alpha, \beta} dW_i^{LVL}(t) \\ \sigma_i^{\alpha, \beta^2} &= \frac{1}{T_{\alpha} - t_0} \int_{t_0}^{T_{\alpha}} \mathbb{E}_{t_0}^{LVL} \left[\left(\sum_{j=\alpha+1}^{\beta} \frac{\zeta_j(t_0)}{LVL(t_0)} \mathcal{B}_i(T_j - t) \sqrt{v_i(t)} \right)^2 \right] dt \\ &= \frac{1}{T_{\alpha} - t_0} \int_{t_0}^{T_{\alpha}} \left(\sum_{j=\alpha+1}^{\beta} \frac{\zeta_j(t_0)}{LVL(t_0)} \mathcal{B}_i(T_j - t) \right)^2 \mathbb{E}_{t_0}^{LVL}[v_i(t)] dt \end{aligned}$$

Sous cette approximation, $S_{\alpha, \beta}(T_{\alpha})$ suit une loi normale conditionnellement à \mathcal{F}_t ($t \geq t_0$). On peut alors calculer le prix de la swaption par une formule de Black.

Or, l'expression de l'espérance des v_i est assez complexe. On va préférer une approche allégée en calculs (donc moins sujette à erreurs), et assez efficace.

3.3 Duration stochastique

Cette approche pour valoriser les swaptions, nous venant de Claus Munk [3], permet de se ramener au pricing d'une option sur zéro-coupon via une simple recherche dichotomique.

3.3.1 Définition

Une swaption peut être vue comme une option de vente de strike 1, sur obligation délivrant des coupons ($K(T_i - T_{i-1})$ à chaque instant T_i , $\alpha < i < \beta$, et $K(T_{\beta} - T_{\beta-1}) + 1$ à T_{β}) :

$$\text{Swaption}_{\alpha, \beta}(t) = \mathbb{E}_t^Q \left[e^{-\int_t^{T_{\alpha}} r_s ds} \left(1 - P(T_{\alpha}, T_{\beta}) - \sum_{i=\alpha+1}^{\beta} K(T_i - T_{i-1}) P(T_{\alpha}, T_i) \right)_+ \right]$$

Notons P_c la valeur de l'obligation et C_i les différents coupons (qui valent soit $K(T_i - T_{i-1})$ soit $1 + K(T_\beta - T_{\beta-1})$) :

$$P_c(t) = \sum_{i=\alpha+1}^{\beta} C_i P(t, T_i)$$

$$\text{Swaption}_{\alpha, \beta}(t) = \mathbb{E}_t^Q \left[e^{-\int_t^{T_\alpha} r_s ds} (1 - P_c(T_\alpha))_+ \right]$$

On définit sa duration stochastique $D(t)$ comme étant la maturité du zéro-coupon qui a la même volatilité log-normale que l'obligation. Elle vérifie donc :

$$\sum_{i=1}^N \mathcal{B}_i(D(t))^2 v_i(t) = \frac{1}{P_c(t)} \sum_{i=1}^N v_i(t) \left(\sum_{k=\alpha+1}^{\beta} C_k P(t, T_k) \mathcal{B}_i(T_k - t) \right)^2$$

Ce $D(t)$, que l'on obtient par dichotomie, existe toujours et est unique si les \mathcal{B}_i ont tous la même monotonie.

3.3.2 Approximation par une option sur zéro-coupon

Comparons une option de maturité T sur notre obligation P_c de strike K , et $\frac{P_c(t)}{P(t, S)}$ options sur un zéro-coupon, de strike $K P(t, S) / P_c(t)$, de même maturité ($S > T$) :

$$\begin{aligned} \mathcal{C}(P_c) - \mathcal{C}(P) &= P(t, S) \mathbb{E}_t^S \left[\frac{1}{P(T, S)} (P_c(T) - K)_+ \right] \\ &\quad - P(t, S) \frac{P_c(t)}{P(t, S)} \mathbb{E}_t^S \left[\frac{1}{P(T, S)} \left(P(T, S) - K \frac{P(t, S)}{P_c(t)} \right)_+ \right] \\ &= P(t, S) \mathbb{E}_t^S \left[\left(\frac{P_c(T)}{P(T, S)} - \frac{K}{P(T, S)} \right)_+ - \left(\frac{P_c(t)}{P(t, S)} - \frac{K}{P(t, S)} \right)_+ \right] \end{aligned}$$

Remarquons que $\mathbb{E}_t^S \left[\frac{P_c(T)}{P(T, S)} \right] = \frac{P_c(t)}{P(t, S)}$. Si l'option est très *in the money*, la probabilité de ne pas exercer l'option est très faible et alors :

$$\begin{aligned} \mathcal{C}(P_c) - \mathcal{C}(P) &\approx P(t, S) \mathbb{E}_t^S \left[\frac{P_c(T)}{P(T, S)} - \frac{K}{P(T, S)} - \frac{P_c(t)}{P(t, S)} + \frac{K}{P(t, S)} \right] \\ &\approx P(t, S) \left(\mathbb{E}_t^S \left[\frac{P_c(T)}{P(T, S)} \right] - \frac{P_c(t)}{P(t, S)} \right) = 0 \end{aligned}$$

Inversement, si l'option est très *out of the money*, la probabilité de l'exercer est négligeable et l'expression résultante est proche de zéro. L'erreur effectuée se situe dans le cas *at the money*.

On souhaite que les termes $\frac{P_c(T)}{P(T, S)}$ et $\frac{P_c(t)}{P(t, S)}$ restent le plus proche possible l'un de l'autre, ce qui revient à minimiser la variance de $\frac{P_c(T)}{P_c(t)} - \frac{P(T, S)}{P(t, S)}$. Une manière de procéder est de choisir S de telle sorte que les dynamiques de $P_c(t)$

et $P(t, S)$ soient identiques à l'instant t . Or les deux actifs offrent le même rendement, et leurs volatilités initiales sont égales pour $S = t + D(t)$, par définition de $D(t)$. On va donc prendre $S = D(t)$.

Le résultat ne change pas que ce soit deux calls, ou deux puts.

Ainsi peut-on faire l'approximation suivante :

$$\text{Swaption}_{\alpha, \beta}(t) \approx \frac{P_c(t)}{P(t, t + D(t))} \mathcal{P} \left(t, T_\alpha, t + D(t), \frac{P(t, t + D(t))}{P_c(t)} \right)$$

4 Implémentation

La plupart des formules présentées s'implémentent sans soucis dans n'importe quel langage. En revanche, le pricing des options est particulièrement délicat. On distinguera trois étapes : la résolution de l'EDO sur les fonctions M et N_i , et l'intégration de la fonction Ψ pour obtenir G_0 ou G_1 . Enfin, l'optimisation permettra de trouver les paramètres du modèle qui lui feront répliquer les prix d'options.

La formulation présentée ne permet guère d'optimisation a priori : la fonction Ψ dépend à la fois de u (la variable d'intégration), de la maturité et du ténor de l'option, et ce de manière non triviale. De plus, même des ténors non entiers peuvent se présenter à cause des swaptions. Enfin, u peut prendre des valeurs très diverses, sans qu'aucune méthode ne puisse prédire les bornes optimales d'intégration. Il faudra donc intégrer Ψ pour chaque option, et l'équation de Riccati pour chaque point de Ψ nécessaire à son intégration.

4.1 Résolution de l'EDO de Riccati

4.1.1 Changement de variable

Une équation de Riccati est une équation différentielle de la forme

$$y' = P(t)y + Q(t)y^2 + R(t)$$

Ces équation n'ont pas de solution analytique dans le cas général. Un des cas dont on connaît des solutions est celui où les coefficients sont constants : $P(t) = p$, $Q(t) = q$, $R(t) = r$. On peut alors se ramener à une équation différentielle plus simple via le changement de variable $y = \frac{1}{z} + \alpha$, $\alpha \in \mathbb{R}$:

$$\begin{aligned} y' &= \frac{-z'}{z^2} \\ &= p \left(\frac{1}{z} + \alpha \right) + q \left(\frac{1}{z} + \alpha \right)^2 + r \\ z' &= -z(p + 2\alpha q) - z^2(p\alpha + q\alpha^2 + r) - q \end{aligned}$$

On peut alors choisir α solution de $q\alpha^2 + p\alpha + r = 0$, et alors :

$$z(t) = e^{-(p+2\alpha q)(t-t_0)} \left(z(t_0) - \frac{q}{2\alpha q + p} \right) + \frac{q}{2\alpha q + p}$$

et ce pour tous t_0 , $z(t_0)$ et sur tout intervalle intégrant un voisinage de t_0 tel que l'expression soit définie.

Bien évidemment, ce cas est très rare et n'apparaît pas dans notre problème. On peut néanmoins s'y ramener en approximant nos coefficients par des fonctions en escaliers. On subdivise notre intervalle $[0, T_0] = \cup_{i=1}^I [\frac{i-1}{I}T_0, \frac{i}{I}T_0]$, et on approxime sur chaque sous-intervalle les coefficients par leurs moyennes sur ces intervalles.

Cette méthode offre une précision médiocre, mais a le grand avantage d'être très stable. Quels que soient les paramètres, le résultat retourné est proche du vrai résultat.

4.1.2 Méthode de Runge-Kutta simple

Les méthodes de Runge-Kutta sont une famille de méthode de résolution d'ODE. On souhaite ici résoudre l'équation différentielle dans \mathbb{R}^n et sur $[0, T_0]$ $y' = f(t, y)$ avec condition initiale $y(0) = y_0$.

On se donne :

- une famille $(c^i)_{1 \leq i < q} \in [0, 1]^{q-1}$ qui définira les points intermédiaires de résolution
- pour tout $i \in \llbracket 1, q-1 \rrbracket$, une famille $(a_j^i)_{1 \leq j \leq i}$ pour intégrer y' par une quadrature au choix :

$$\int_0^{c_i} g(u) du \approx \sum_{j=1}^i a_j^i g(c_j)$$

- une dernière quadrature $(b_j)_{1 \leq j \leq q}$ qui donnera une valeur approchée de $y(T_0)$

On pose alors $p_1 = f(0, y_0)$ puis pour tout $k \in \llbracket 2, q \rrbracket$:

$$p_k = f(c_{k-1}T_0, y_0 + T_0(a_1^{k-1}p_1 + \dots + a_{k-1}^{k-1}p_{k-1}))$$

Finalement :

$$y(T_0) \approx \sum_{k=1}^q b_k p_k$$

En particulier, la méthode de Runge-Kutta appelé RK4 possède les paramètres suivants :

$$\begin{aligned} (c_1, c_2, c_3, c_4) &= \left(0, \frac{1}{2}, \frac{1}{2}, 1\right) \\ a_1^1 &= \frac{1}{2} \\ (a_1^2, a_2^2) &= \left(0, \frac{1}{2}\right) \\ (a_1^3, a_2^3, a_3^3) &= (0, 0, 1) \\ (b_1, b_2, b_3, b_4) &= \left(\frac{1}{6}, \frac{1}{3}, \frac{1}{3}, \frac{1}{6}\right) \end{aligned}$$

Autrement dit :

$$\begin{aligned} p_1 &= f(0, y_0) \\ p_2 &= f\left(\frac{T_0}{2}, y_0 + T_0 \frac{p_1}{2}\right) \\ p_3 &= f\left(\frac{T_0}{2}, y_0 + T_0 \frac{p_2}{2}\right) \\ y(T_0) &\approx y_0 + T_0 \frac{p_1 + 2p_2 + 2p_3 + p_4}{6} \end{aligned}$$

L'erreur totale est ici de l'ordre de T_0^4 . Bien entendu, pour obtenir un résultat correct, il s'agit de subdiviser notre intervalle de résolution et d'appliquer la méthode sur chaque sous-intervalle. La seule variable dans cette méthode est donc le pas d'intégration : trop grand, elle sera imprécise voire instable, trop petit elle sera trop lente.

Dans l'implémentation qui a été faite, 300 pas sont un bon compromis entre ces différentes contraintes. Les cas où la méthode diverge sont facilement repérables, car dès qu'elle diverge, on atteint en quelques pas des valeurs non numériques en $C\#$. De plus ces mêmes cas correspondent généralement aux endroits où Ψ prend des valeurs négligeables, il suffit alors d'appliquer un traitement très simple.

4.1.3 Méthode adaptative

Pour pallier le problème de stabilité, et dans certains cas gagner en vitesse, on peut adapter le pas de résolution h . Pour cela, il nous faut estimer l'erreur commise à chaque itération. Pour ce faire, on se munit d'une méthode d'ordre q et une méthode d'ordre $q - 1$ qui utilisent les mêmes q valeurs de f . Leur différence donne une estimation de l'erreur.

On a donc toujours une famille $(a_j^i)_{1 \leq j \leq i < q} \in \mathbb{R}^{\frac{q(q-1)}{2}}$, mais deux quadratures $(b_i)_{1 \leq i \leq q}$ et $(c_i)_{1 \leq i \leq q}$. La première donne un résultat $y^p = y(T_0) + \mathcal{O}(h^{p+1})$, et la seconde une autre résultat $y^{p-1} + \mathcal{O}(h^p)$. L'erreur est donc $\Delta = y^p - y^{p-1} = \mathcal{O}(h^p)$. Si on désire obtenir une erreur de l'ordre de ε , on peut alors réajuster h en le remplaçant par $h \left| \frac{\varepsilon}{\Delta} \right|^{\frac{1}{p}}$. Nous avons ici implémenté la méthode de Cash-Karp, tabulée en annexe.

Cette méthode permet de résoudre n'importe quelle ODE avec une précision correcte. En revanche, rien n'est dit quant à sa rapidité : dans certains cas, le pas h décroît très vite, ce qui rend l'algorithme particulièrement lent.

4.1.4 Choix d'algorithme

La stabilité et la rapidité des différents algorithmes dépend fortement des paramètres de l'équation et de façon difficilement prédictible. Or, ceux-ci sont nombreux et sont amenés à varier à chaque appel. Dans notre implémentation, c'est la méthode adaptative de Cash-Karp qui est donc utilisée en premier. Cependant, dans certains cas, le pas h de résolution décroît très vite, ce qui rend impossible en pratique l'utilisation de cette méthode. Lorsque $h < \frac{T_0}{1000}$, on délègue alors à la méthode par approximation et changement de variable.

En annexe est présenté un tableau esquissant les performances de différentes méthodes en termes de précision et de rapidité.

4.2 Intégration de Ψ

4.2.1 Quadrature de Gauss-Legendre

Pour intégrer la fonction Ψ , nous faisons appel à une quadrature de Gauss-Legendre. Les quadratures de Gauss sont une famille de méthodes d'intégration numérique de fonctions qui ont la particularité, pour tout ordre n donné (i.e. n calculs de la fonction sont demandés), d'être exactes sur les polynômes de degré au plus $2n - 1$. Elles sont caractérisées par :

- Un intervalle d'intégration I
 - Une fonction de pondération $\omega : I \rightarrow]0, +\infty[$
 - n nœuds $(x_i)_{1 \leq i \leq n} \in I^n$
 - n poids associés à ces nœuds $(\omega_i)_{1 \leq i \leq n} \in]0, +\infty[$
- Une fonction $f : I \rightarrow \mathbb{R}$ est alors intégrée ainsi :

$$\int_I \omega(t)f(t)dt \approx \sum_{i=1}^n \omega_i f(x_i)$$

Diverses quadratures de ce type existent, en fonction de la fonction de pondération. Nous avons comme intégrande $u \in]0, +\infty[\rightarrow \Im \left[\frac{\Psi(a+iu, T_0, T_1, y)e^{-iuy}}{u} \right]$ et comme fonction de pondération : 1. C'est donc la quadrature de Gauss-Legendre que nous utilisons.

Ses nœuds x_i sont les racines du n -ième polynôme de Legendre P_n , ses poids valent : $\omega_i = \frac{2}{(1-x_i^2)P_n'(x_i)^2}$, et elle s'utilise sur l'intervalle $[-1, 1]$. Ces polynômes peuvent être calculés par la récurrence suivante :

$$\begin{aligned} P_0 &= 1 \\ P_1 &= X \\ (k+1)P_{k+1} &= (2k+1)XP_k - kP_{k-1} \end{aligned}$$

4.2.2 Méthode de Simpson

La quadrature de Gauss-Legendre est très efficace sur une fonction régulière, définie sur un segment. Or notre intégrande présente par moment une singularité en zéro, qui ne peut être intégrée efficacement par cette méthode. On fait alors appel à une méthode de Simpson modifiée.

La méthode de Simpson consiste à interpoler l'intégrande f avec un polynôme d'ordre 2 aux points a , b , et $\frac{a+b}{2}$:

$$\int_{[a,b]} f(t)dt \approx \frac{b-a}{6} \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right]$$

En s'inspirant de cette méthode, on peut construire une méthode itérative pour intégrer une fonction au voisinage de zéro. On définit d'abord la suite des abscisses par $x_0 > 0$ et $x_{n+1} = \alpha x_n$, $\alpha \in]0, 1[$. Sur chaque segment $[x_{n+2}, x_n]$, on interpole l'intégrande par un polynôme de degré 2 :

$$\begin{aligned} f(t) &\approx f(x_{n+2}) \frac{(x - x_{n+1})(x - x_n)}{(x_{n+2} - x_{n+1})(x_{n+2} - x_n)} \\ &\quad + f(x_{n+1}) \frac{(x - x_{n+2})(x - x_n)}{(x_{n+1} - x_{n+2})(x_{n+1} - x_n)} + f(x_n) \frac{(x - x_{n+2})(x - x_{n+1})}{(x_n - x_{n+2})(x_n - x_{n+1})} \\ &\approx \frac{f(x_{n+2})}{(\alpha^2 - \alpha)(\alpha^2 - 1)x_n^2} (x - \alpha x_n)(x - x_n) \\ &\quad + \frac{f(x_{n+1})}{(\alpha - \alpha^2)(\alpha - 1)x_n^2} (x - \alpha^2 x_n)(x - x_n) + \frac{f(x_n)}{(1 - \alpha^2)(1 - \alpha)x_n^2} (x - \alpha^2 x_n)(x - \alpha x_n) \end{aligned}$$

On aboutit alors à

$$\begin{aligned} \int_{x_{n+2}}^{x_n} f \approx & f(x_{n+2}) \times x_n^3 \left[\frac{1-\alpha^6}{3} - \frac{\alpha+1}{2}(1-\alpha^4) + \alpha(1-\alpha^2) \right] \\ & + f(x_{n+1}) \times x_n^3 \left[\frac{1-\alpha^6}{3} - \frac{\alpha^2+1}{2}(1-\alpha^4) + \alpha^2(1-\alpha^2) \right] \\ & + f(x_n) \times x_n^3 \left[\frac{1-\alpha^6}{3} - \frac{\alpha(\alpha+1)}{2}(1-\alpha^4) + \alpha^3(1-\alpha^2) \right] \end{aligned}$$

On calcule récursivement l'intégrale de f sur les intervalles $[x_n, x_0]$, en s'arrêtant lorsque $\left| \int_{x_{n+2}}^{x_n} f(t)dt \right|$ passe sous un seuil fixé.

4.2.3 Choix d'algorithme

Pour déterminer l'intervalle d'intégration, on parie sur une décroissance exponentielle de l'intégrande. On suppose donc que $|\Psi(u, T, S, y)| \underset{u \rightarrow \infty}{\sim} \Psi_0 e^{-\lambda u}$, et on observe que $\Psi_0 \approx 1$. Alors on estime λ grâce à une évaluation de Ψ en $u = 200$, puis on pose l'extrémité droite H de notre intervalle d'intégration ainsi :

$$H = 200 \frac{\ln \varepsilon}{\ln \Psi(200, T, S, y)}$$

où ε est un seuil de précision fixé indépendamment, en l'occurrence $\frac{1}{500}$.

Sur l'intervalle $]0, 2\%H]$ est appliquée la méthode itérative inspirée de la méthode de Simpson. Sur $[2\%H, H]$ est appliquée une quadrature de Gauss-Legendre, à 20 points d'intégration.

Cette méthode est malheureusement insuffisante pour valoriser des options hors de la monnaie. En effet, plus le strike s'éloigne de l'ATM, plus des oscillations apparaissent (à cause du terme $iu \ln K$ dans l'exponentielle). Il sera nécessaire, pour aller plus loin, d'adapter ces quadrature à des fonctions oscillantes ou en trouver d'autres.

4.3 Optimisation

Pour finalement trouver les paramètres du modèle qui lui permettront de correspondre au marché, nous chercherons à minimiser la fonction suivante :

$$f(\Theta) = \sum_{i \in I} \left(\frac{\hat{p}_i(\Theta) - p_i^*}{p_i^*} \right)^2$$

où $\Theta \in \mathbb{R}^{7N}$ représente un jeu de paramètres, les $\hat{p}_i(\Theta)$ (resp. p_i^*) les prix du modèle pour Θ (resp. les prix du marché) sur un ensemble I de produits dérivés de taux.

On compte pourtant 8 paramètres par facteurs : $\kappa_i, \theta_i, \sigma_i, \gamma_i, \alpha_0^i, \alpha_1^i, \rho_i$ et $v_i(0)$. Mais on peut montrer que cette représentation est redondante ; pour s'en convaincre, on peut montrer qu'on peut laisser inchangé un terme de volatilité du taux forward $\sigma_{f,i}(t, T) \sqrt{v_i(t)}$ en remplaçant (α_0^i, α_1^i) par $(\lambda^{-\frac{1}{2}} \alpha_0^i, \lambda^{-\frac{1}{2}} \alpha_1^i)$

et en remplaçant $(v_i(0), \theta_i, \sigma_i)$ par $(\lambda v_i(0), \lambda \theta_i, \sqrt{\lambda} \sigma_i)$. La diffusion de $f(t, T)$ reste la même, et les prix de tous les produits aussi, mais les paramètres sont différents. On se débarrasse alors d'un paramètre, et conformément à l'article, on fixe $\theta_i = 1$.

4.3.1 Descente de gradient

Cette méthode part d'un constat simple : si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est continûment différentiable et $t \geq 0 \mapsto x_t$ suit l'équation différentielle suivante

$$x'_t = -\vec{\nabla}_{x_t} f$$

alors $\partial_t f(x_t) = -\|\vec{\nabla}_{x_t} f\|^2 < 0$. En une zone dans laquelle f est convexe, x_t convergera vers le minimum de cette zone. On en déduit l'algorithme suivant :

- $x_0 = x$
- $x_{n+1} = x_n - \lambda \vec{\nabla}_{x_n} f$
- Retourner x_n si on a atteint un critère de convergence

Deux paramètres sont à choisir. L'algorithme mènera à un minimum local dépendant du point initial x . Si la fonction présente de nombreux minima locaux, il sera peut-être nécessaire de relancer l'algorithme avec différentes valeurs de x , voire de faire appel à un autre algorithme pour trouver un bon candidat. Enfin, λ détermine la longueur des pas. Trop petit, l'algorithme sera très lent. Trop grand, l'algorithme pourrait être instable et ne pas converger.

4.3.2 BFGS

Dans la même famille des algorithmes à direction de descente, on retrouve BFGS (de Broyden, Fletcher, Goldfarb et Shanno). On part d'une estimation numérique B_0 de la hessienne de f en x . Puis, à chaque pas n :

- On cherche la direction de descente p_n en résolvant $B_n p_n = -\vec{\nabla}_{x_n} f$
- On optimise suivant cette direction : $\alpha_n = \underset{\alpha \in \mathbb{R}}{\operatorname{argmin}} f(x_n + \alpha p_n)$, et
- $x_{n+1} = x_n + \alpha_n p_n$
- On met à jour l'estimation de la hessienne :

$$B_{n+1} = B_n + \frac{(\vec{\nabla}_{x_{n+1}} f - \vec{\nabla}_{x_n} f)(\vec{\nabla}_{x_{n+1}} f - \vec{\nabla}_{x_n} f)^T}{(\vec{\nabla}_{x_{n+1}} f - \vec{\nabla}_{x_n} f)^T \alpha_n p_n} - \frac{(\alpha_n B_n p_n)(\alpha_n B_n p_n)^T}{(\alpha_n p_n)^T B_n (\alpha_n p_n)}$$

Cet algorithme est plus performant que la descente de gradient dans le cas général, notamment parce qu'il s'affranchit du choix λ , implicitement déterminé de manière optimale. Mais il tombe dans le même travers que tous les algorithmes à direction de descente, c'est-à-dire qu'il ne peut s'extraire de la cuvette choisie par la valeur initiale x .

L'implémentation utilisée ici est celle de la bibliothèque ALGLIB, disponible sur internet [4].

4.3.3 Algorithme génétique

Les algorithmes génétiques sont une famille de méthodes stochastiques d'optimisation s'inspirant de la nature, plus particulièrement de l'évolution. Une population est générée aléatoirement. Puis, génération après génération, elle

évolue, et les individus les mieux adaptés à leur environnement sont sélectionnés. La population globale elle-même s'adapte petit à petit : on converge vers une solution au problème d'optimisation.

Dans les faits, la population est un ensemble de taille fixe n de jeux de paramètres. On mesure leur capacité d'adaptation à la valeur qu'ils donnent à f , la fonction à minimiser. La population initiale P_0 est tirée aléatoirement de manière uniforme, des bornes étant indiquées pour chaque paramètre. À chaque génération, on applique un opérateur de mutation à la population. Les éventuels paramètres qui sortent des bornes fixées au préalable subissent un traitement pour revenir dans le bon intervalle.

- On tire aléatoirement deux permutations de $\llbracket 1, n \rrbracket$ σ_1 et σ_2
- Pour tout individu i et chacun de ses paramètres $p_{i,j}$, avec une probabilité C :
 - On remplace ce paramètre $p_{i,j}$ par le barycentre de coefficient F entre ce paramètre et le paramètre correspondant du meilleur individu :

$$p_{i,j} \leftarrow p_{i,j} + F \times (p_j^{best} - p_{i,j})$$
 - On ajoute un bruit : $p_{i,j} \leftarrow p_{i,j} + F \times (p_{i,\sigma_1(j)} - p_{i,\sigma_2(j)})$

FIGURE 2 – Opérateur de mutation

On calcule alors la fonction objective pour tous les nouveaux individus, et on garde le meilleur de chaque couple individu père - individu fils. On met de côté le meilleur membre de la population totale, et on réitère.

L'avantage de cet algorithme est qu'il s'appuie sur une large population, initialisée aléatoirement. Il est donc beaucoup moins dépendant de son initialisation, et est plus à même d'éviter les minima non globaux. En revanche, son caractère aléatoire le rend imprévisible, et il peut donner, avec les mêmes paramètres, des résultats différents. De plus il est en général plus lent à converger qu'un algorithme comme le BFGS.

Étant donné que nous travaillons sur une fonction comportant 14 paramètres pouvant comporter de nombreuses cuvettes, ce type d'algorithme est très intéressant, surtout pour réduire les bornes de recherches. On peut l'utiliser en premier lieu pour dégrossir, puis basculer vers un algorithme plus rapide pour optimiser plus finement.

4.3.4 Choix d'algorithme

L'algorithme génétique a l'avantage de ne pas se restreindre à une cuvette, mais de chercher dans tout l'espace des paramètres disponible. En revanche, il s'avère être extrêmement long dans notre cas (rappelons que le calcul de la fonction objective requiert la résolution de l'équation de Riccati à chaque point d'intégration, et qu'il y a deux intégrales à calculer par option à valoriser).

Il s'agit donc de l'utiliser afin de calibrer le modèle pour la première fois, ou après un événement économique majeur. Les réglages trimestriels pourront se faire via un algorithme à direction de descente (ici BFGS).

Par ailleurs, la fonction objective effectue à chaque calcul un test vérifiant si les paramètres sont contenus dans des bornes convenables afin que les deux fac-

teurs aient des structures de volatilités conformes à nos attentes, et qu'aucun ne « disparaisse ». La condition de Feller pour chaque volatilité est également vérifiée. Si ce test échoue, la fonction objective vaut 10^{50} , une valeur discriminante et autrement inatteignable.

4.4 Simulation

Dans notre implémentation, la simulation se fait via un schéma d'Euler, sauf pour les $v_i(t)$ auxquels on applique un schéma de Milstein. Pour une équation différentielle donnée $dX(t) = \mu(X(t))dt + \sigma(X(t))dW_t$, le schéma de Milstein consiste à subdiviser l'intervalle $[0, T]$ de résolution en N pas de temps $\frac{T}{N}$, et à calculer par récurrence :

$$X_{n+1} = X_n + \mu(X_n)\frac{T}{N} + \sigma(X_n)U_n + \frac{1}{2}\sigma(X_n)\frac{\partial\sigma(X_n)}{\partial X}\left(U^2 - \frac{T}{N}\right)$$

avec $X_0 = X(0)$ et $U_n \sim \mathcal{N}\left(0, \frac{T}{N}\right)$

Ce schéma converge fortement et faiblement ² en $\mathcal{O}(N^{-1})$, alors que le schéma d'Euler, qui n'a pas ce dernier terme, converge fortement en $\mathcal{O}(N^{-\frac{1}{2}})$.

2. Voir annexes

Troisième partie

Analyse et comparaison des résultats

1 Calibration du modèle

1.1 Données

1.1.1 Swaptions

Nous avons effectué plusieurs calibrations, sur des données espacées de trois mois (cette durée étant justifiée par le fait que les exercices sont trimestriels). Nous avons choisi une matrice de swaptions dont les maturités valent 1, 2, 3, 4, 5, 7, 10, 15 et 20 ans, et dont les ténors valent 1, 5 et 10 ans. Il est important d'avoir un large panel de produits dérivés : d'une part nos simulations ne se feront qu'à l'horizon de quelques années, mais les produits que nous détenons et notre portefeuille d'assurés ont plutôt de longues durations, il faut avoir des paramètres qui reproduisent bien à la fois le moyen terme (1 à 5 ans) et le long terme (plus de 10 ans). D'autre part, nous avons 14 paramètres à calibrer (7 par facteur de risque). Il faut donc au grand minimum 14 produits dérivés, si possible assez variés. Par simplicité nous nous contentons des swaptions, qui contiennent déjà beaucoup d'information dans leurs prix, il faut donc varier les produits par leurs paramètres de maturité et de ténor, afin de couvrir une large plage temporelle.

Le choix de ces couples de maturités et ténors est aussi dû à la liquidité de ces produits. Plus un produit est liquide (c'est-à-dire échangé fréquemment et en grands volumes sur le marché), plus le prix observé est représentatif des conditions de marché. Un produit peu liquide sera moins fréquemment coté et le prix, tel que donné par un fournisseur comme Bloomberg, peu faire l'objet d'une interpolation à partir d'autres produits financiers plus courants.

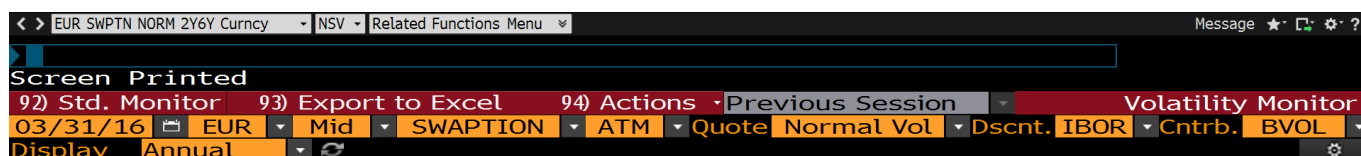


FIGURE 3 – Paramètres des extractions Bloomberg des volatilités

Nous avons donc deux jeux de 27 swaptions du marché Euro, l'un datant du 31 mars 2016, l'autre du 30 juin 2016. Les données ont été extraites de Bloomberg aux lendemains des cotations sous forme de volatilités implicites normales (l'environnement de taux négatifs empêchant le pricing de certains produits par Black-Scholes). Une donnée aberrante au mois de mars a été conservée dans la visualisation des résultats, mais était exclue de la calibration. L'extrait de capture d'écran en figure 3 donne les paramètres exacts de l'extraction des volatilités du mois de mars. Seule la date change au mois de juin. Les valeurs exactes

sont présentées en annexe.

Les figures 6 à 8 donnent un tracé des prix et volatilités des swaptions utilisées. On remarque d’abord que les prix sont croissants avec le ténor, presque proportionnels : en effet le ténor représente le nombre de flux sous-jacents. En quelque sorte, le nominal croît avec le ténor, ce qui fait augmenter le prix. Globalement ils sont aussi croissants avec la maturité, car l’incertitude sur le taux swap augmente avec sa date de maturité. Comme il s’agit d’une swaption payeuse, analogue à une option d’achat, le prix est croissant avec la variance de la variable aléatoire sous-jacente. Enfin, on pourrait interpréter la forme des nappes de volatilité de cette même façon, à savoir que l’incertitude augmente avec la maturité.

D’un trimestre à l’autre, on remarque que les volatilités ont diminué, de façon significative pour les courtes maturités, et de manière plus modérée pour les longues maturités. On peut y voir un apaisement des marchés, qui ne change pas les perspective à long terme.

1.1.2 Courbes de taux

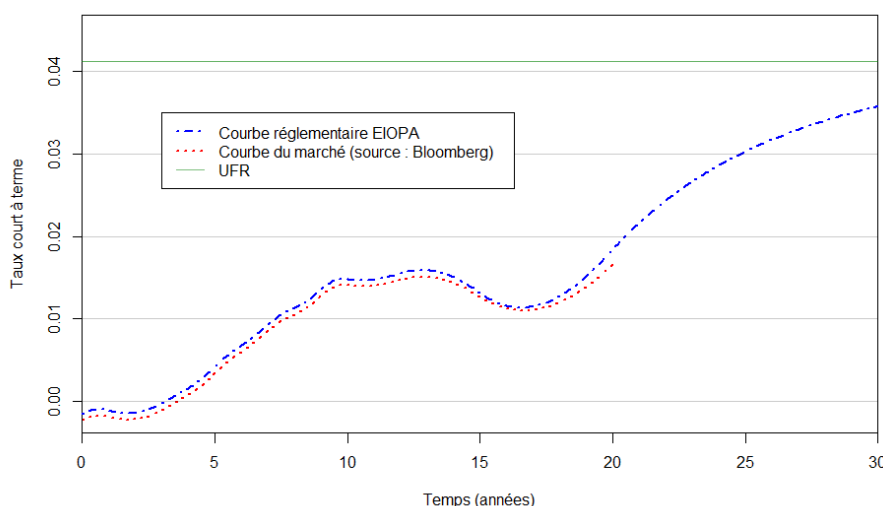


FIGURE 4 – Courbe de taux du marché (reconstruite à partir des taux swap), et courbe de taux réglementaire. Données du 30 juin

Les courbes des taux réglementaires sont fournies mensuellement par l’EIOPA (voir [9]) ; elles présentent un pas annuel. Pour obtenir les prix des zéro-coupons à toute date comme ci-dessous, nous utilisons les outils de stripping de la même suite de logiciels qui nous fournit le modèle LMM+. Ce sont ces prix de zéro-coupons qui sont ensuite utilisés dans les valorisations de produits dérivés et la calibration du modèle.

Ces courbes s’appuient sur des données de marché, mais intègrent des ajustements (CRA, *Credit Risk Adjustment*, et VA, *Volatility Adjustment*) et sont extrapolées jusqu’à 120 ans. Les ajustements sont à la source de la différence

observée avec la courbe des taux du marché (cf. figure 4). L'extrapolation est faite par l'algorithme de Smith-Wilson avec un UFR (*Ultimate Forward Rate*, i.e. $\lim_{t \rightarrow \infty} f(0, \cdot)$, le taux à terme limite) à 4.2%. Nous n'intervenons pas dans ce processus.

Comme souvent, la courbe présente une allure croissante. On l'explique souvent par l'appétence des investisseurs aux courtes maturités, moins risquées. Les premières valeurs sont négatives : ce phénomène, inédit en Europe, découle de la politique de la Banque Centrale Européenne qui pousse les banques à l'investissement. C'est ce qui rend la situation délicate pour toutes les compagnies d'assurances et fonds de pension, qui cherchent à se rémunérer sur les produits de taux. Enfin, la courbe du marché s'arrête à 20 ans, les maturités au-delà n'étant pas assez liquides pour être cotées précisément. La courbe est donc extrapolée pour nos besoins réglementaires à l'UFR, fixé par l'EIOPA et qui reste stable entre chaque exercice.

1.2 Résultats de calibration

1.2.1 Méthode

Le modèle a d'abord été calibré par un algorithme génétique sur les données de mars, afin de trouver un jeu de paramètres cohérent parmi l'ensemble des possibles. Ces paramètres ont ensuite été raffinés par BFGS. La calibration sur les données de juin a été faite uniquement par BFGS, en prenant les paramètres de mars comme point de départ. Le modèle retenu comprend deux facteurs, donc quatre browniens. Ce choix résulte d'un compromis entre facilité de calibration (plus il y a de facteurs, plus il y a de paramètres, et plus la calibration est lente et complexe) et adaptabilité du modèle (on peut déjà distinguer avec deux facteurs un premier qui agirait sur les courtes maturités et un autres sur les longues maturités).

De manière assez classique, on force la condition de Feller en rejetant tout jeu de paramètre ne la remplissant pas. Cette condition, portant sur la volatilité ($2\kappa_i\theta_i > \sigma_i^2$), assure que les processus CIR restent strictement positifs. Si le fait qu'il atteigne ponctuellement zéro n'est pas problématique d'un point de vue analytique, cela pourrait mener à des défauts de simulation assez importants. En effet, ce processus suit une loi du chi-deux décentrée (cf. Partie I) qui est difficile à simuler efficacement. On utilise alors un schéma d'Euler gaussien qui est très imprécis lorsque les conditions de Feller ne sont pas vérifiées.

1.2.2 Paramètres obtenus

Voici les paramètres obtenus successivement sur mars et juin :

31 mars	Facteur 1	Facteur 2
κ	1.75890039	0.31042644
σ	1.25214410	0.53966569
α_0	0.00022565	-0.0098245
α_1	0.00244782	0.02596648
γ	0.10682016	1.02980166
ρ	0.82854839	0.72156202
v_0	0.83371741	0.00118201

30 juin	Facteur 1	Facteur 2
κ	1.75665486	0.30960769
σ	1.25336971	0.53952690
α_0	0.00022557	-0.0098245
α_1	0.00237198	0.02575094
γ	0.10778579	1.02986481
ρ	0.82854899	0.72155919
v_0	0.82558767	0.00118198

On remarque que les paramètres sont très stables, la plus forte variation relative enregistrée étant de 3.1% (pour α_1^1). Cette stabilité est nécessaire, elle témoigne de la bonne cohérence du modèle. Si les paramètres se trouvaient être très dépendants du temps, le modèle ne pourrait pas être utilisé pour la simulation sur plusieurs années. On notera que dans l'intervalle de temps considéré, le Royaume-Uni a quitté l'Union Européenne. Le modèle a donc survécu au Brexit.

En revanche, c'est une stabilité quelque peu artificielle. Lors de la seconde calibration, on n'a pas cherché le minimum global de la fonction d'erreur mais on a utilisé un algorithme à direction de descente en supposant que ce minimum se trouvait dans la même cuvette qu'au mois de mars. Ça revient à faire l'hypothèse que les conditions de marché varient lentement avec le temps. Mais la calibration globale avec l'algorithme génétique se trouve être très lente (il a fallu cinq heures³), alors que vingt minutes suffisent au BFGS. De plus, le minimum atteint par algorithme génétique est *probablement* global, mais on ne peut pas être certain que ce ne soit pas un simple minimum local.

L'utilisation d'un BFGS à chaque trimestre permet un bon compromis entre d'une part rapidité des calculs et stabilité des paramètres, et d'autre part une reproduction précise des prix du marché.

Pour rappel, les paramètres κ et σ déterminent la dynamique de la volatilité de chaque facteur. Le premier est la vitesse de retour à la moyenne de la volatilité : une valeur de 1.75 (resp. 0.31) comme ci-dessus veut dire que la volatilité,

3. Avec un Intel Core i5-2400 4x3.10GHz

en l'absence de perturbation, retourne à son niveau moyen (fixé à 1) exponentiellement, avec un temps caractéristique de 7 mois (resp. 3 ans). Le second, σ , représente la variance du facteur de volatilité. À ratio $\frac{\kappa}{\sigma^2}$ fixé, un σ faible donnera donc lieu à de longues périodes de haute volatilité des taux suivies de longues périodes de faible volatilité des taux, alors qu'elles se succéderont plus rapidement si σ est élevé. La volatilité initiale est donnée par v_0 .

Les paramètres α_0 , α_1 et γ localisent les effets de chaque facteur de risque sur la courbe de taux via les fonction $\sigma_{f,i}$ qui sont représentées pour le mois de juin en figure 5. Conformément à nos désirs, un facteur permet d'expliquer les variations des taux de longues maturités (plus de cinq ans) et un autre se concentre sur les courtes maturités. On remarque que les deux deviennent négligeables aux très courtes maturités (moins d'un an) et tendent vers 0 au voisinage de l'infini. Cela pourrait paraître absurde, étant donné que les taux de maturités extrêmes connaissent aussi une certaine volatilité. La raison de ce phénomène est que ces dates-là n'apparaissent pas dans notre portefeuille de calibration. On ne peut donc pas quantifier ces volatilités-là et reproduire correctement les vraies $\sigma_{f,i}$. De plus, le fait qu'un facteur soit très concentré sur les courtes maturités (moins de cinq ans) alors que l'autre s'étale beaucoup plus reflète notre portefeuille de calibration, à moitié constitué de swaptions de courtes maturités.

On peut remarquer que les paramètres α_0 et α_1 diminuent et γ augmente, ce qui a pour effet de baisser les valeurs des fonctions $\sigma_{f,i}$ et diminuer globalement la volatilité des taux. De plus, les valeurs de v_0 diminuent aussi ce qui a le même effet sur le court terme (car il ne s'agit que de la volatilité à $t = 0$). Ces remarques sont à rapprocher de celles qui ont précédemment été faites sur les volatilités implicites des swaptions : d'un trimestre à l'autre, on observe qu'elles diminuent, en particulier aux courtes maturités. Il n'y a rien de surprenant, car même hors du modèle de Black-Scholes dans lequel elles sont égales, la volatilité des taux et la volatilité implicite des produits dérivés sont très liées.

Enfin, en lisant les valeurs de ρ , on observe partout une forte corrélation entre les taux et leurs volatilités : lorsque les taux augmentent, il en va de même de leur volatilité, et inversement.

1.3 Prix et volatilités implicites

La calibration est satisfaisante : au mois de mars, l'erreur relative quadratique moyenne sur l'ensemble des prix (hors aberration) est de -0.18%, avec un écart-type de 4.30%. Au mois de juin, l'erreur augmente avec une moyenne de -0.08% et un écart-type de 5.17%.

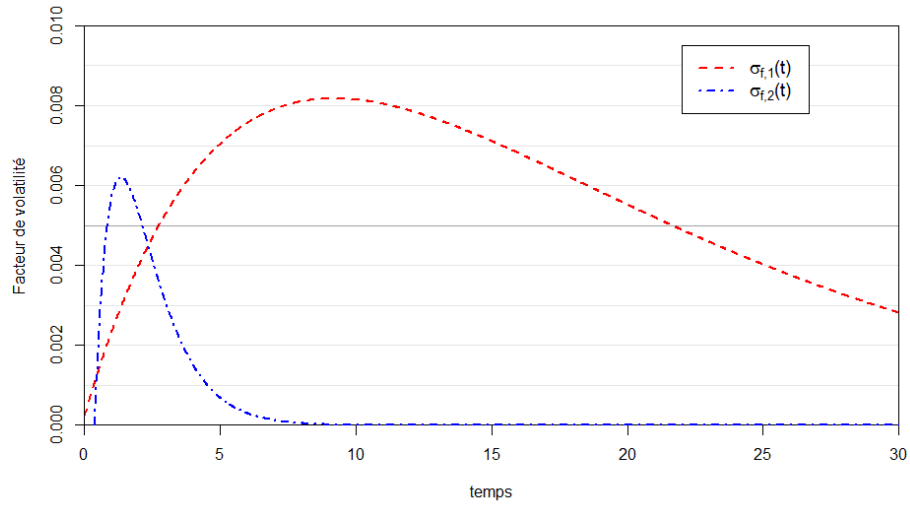


FIGURE 5 – Facteurs de volatilité $\sigma_{f,i}(t, T) = (\alpha_0^i + \alpha_1^i(T - t))e^{-\gamma_i(T-t)}$ tels que calibrés au 30 juin

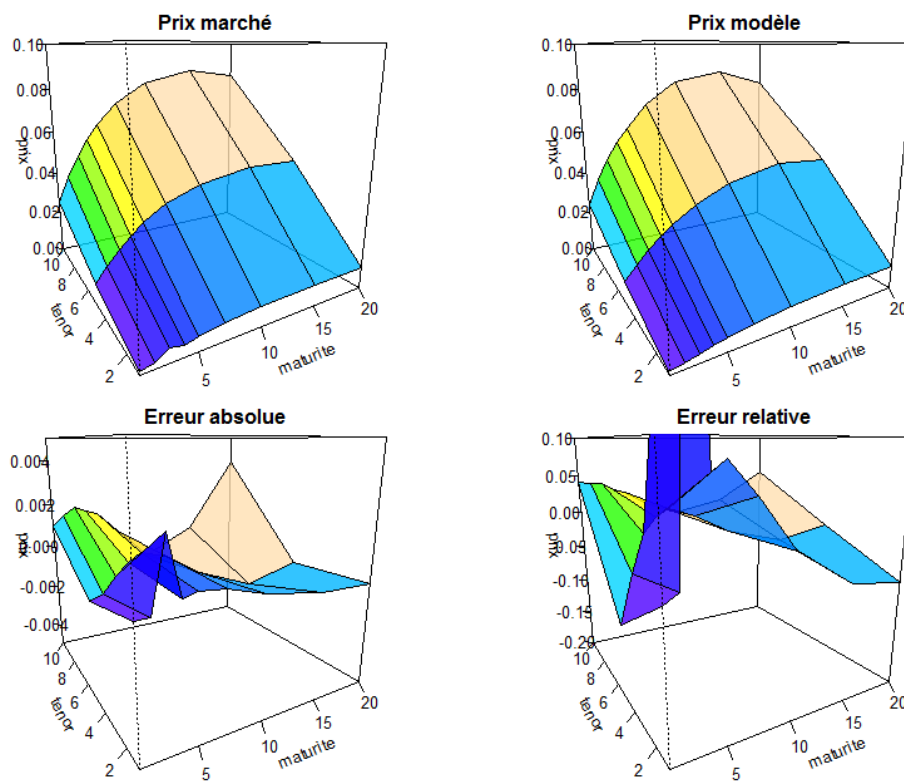


FIGURE 6 – En haut : surfaces des prix du marché et du modèle. En bas : erreur de calibration sur les prix. Données et calibration au 31 mars 2016

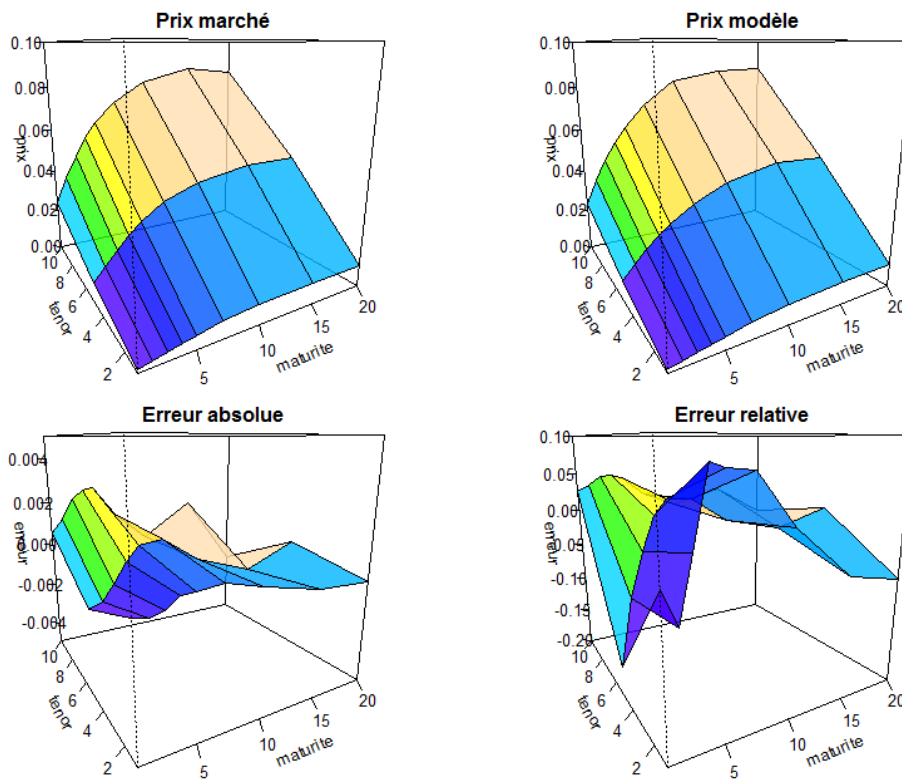


FIGURE 7 – En haut : surfaces des prix du marché et du modèle. En bas : erreur de calibration sur les prix. Données et calibration au 30 juin 2016

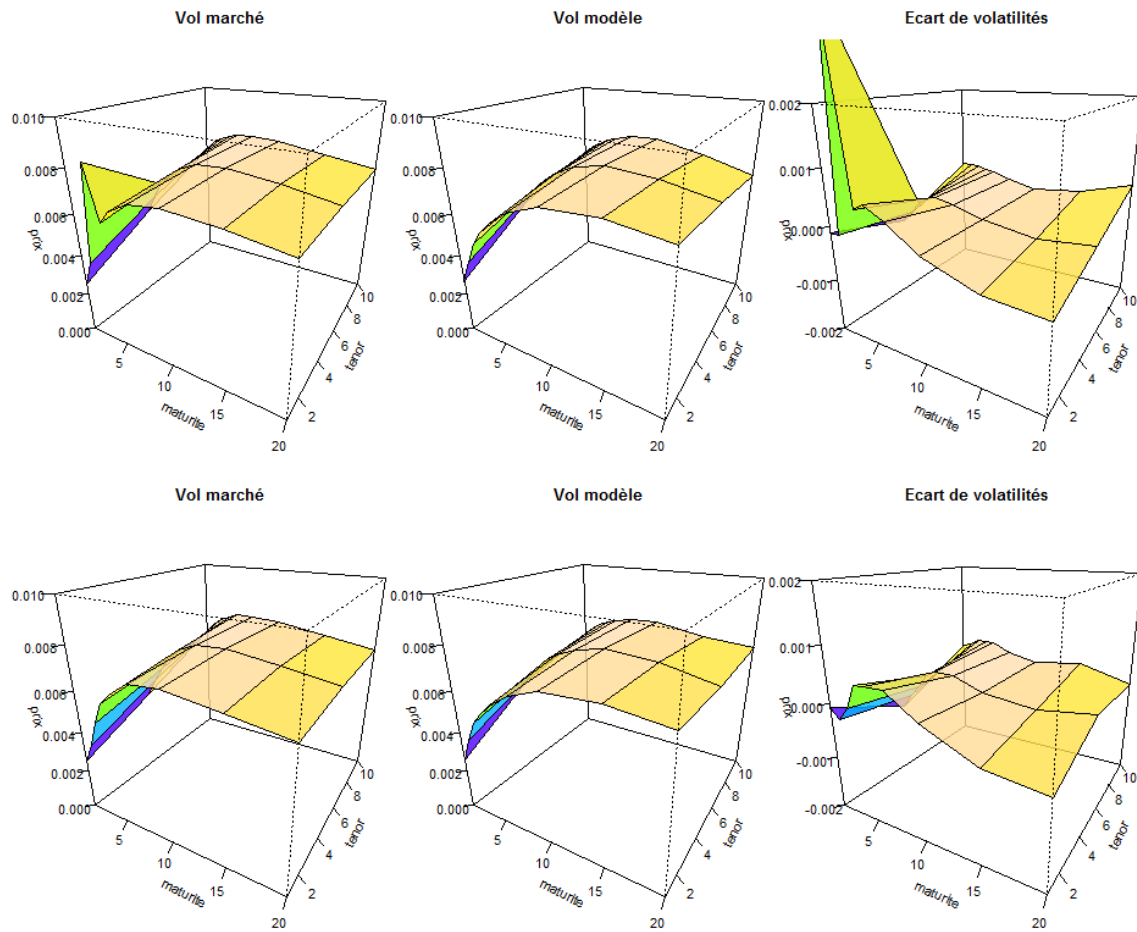


FIGURE 8 – Volatilités implicites au 31 mars 2016 (en haut) et au 30 juin 2016 (en bas), et différence entre les volatilités implicites du marché et du modèle

2 Simulation

En figures 10, 11 et 12 sont présentées les simulations de différents taux, et zéro-coupons. 5000 simulations sur trente ans ont été faites avec le modèle qui fait l'objet du présent mémoire, calibré sur les données de juin, ainsi qu'avec le LMM+ calibré à la même date. Ce dernier comporte trois facteurs (deux pour les taux et un pour la volatilité), et le terme de décalage est fixé à -45% . On peut déjà formuler quelques remarques quant à l'allure de certains graphes.

Les taux sont globalement croissants, et naturellement, les zéro-coupons décroissants avec le temps. On peut l'expliquer par le fait que dans nos modèles, la courbe de taux initiale constitue une anticipation du taux court. Dans les trajectoires des taux LIBOR et swap se retrouve donc la forme de la courbe EIOPA présentée précédemment.

À première vue, il n'y a guère de différence entre les simulations par le LMM+ et par le modèle HJM. En effet, les deux sont calibrés sur les mêmes données et, naturellement, les prix moyens de zéro-coupons vont converger vers les prix de zéro-coupons *forward*, ces derniers étant fixés par la même courbe des taux initiale. De plus, les prix des produits dérivés dépendent aussi des distributions de différentes variables aléatoires (zéro-coupons, taux LIBOR, taux swap...). Devoir répliquer les prix demande donc de répliquer des distributions implicitement imposées par le marché. En répliquant ces mêmes distributions, les différentes variables tracées dans les figures 10 à 12 présentent donc les mêmes moments (en l'occurrence mêmes moyennes et variances) entre le LMM+ et le HJM. Ces tracés nous donnent ainsi une bonne vision des anticipations du marché sur les taux, au travers des prix, mais ne permettent pas encore de distinguer nos deux modèles.

Simuler un modèle requiert en général des approximations, comme un schéma d'Euler ou de Milstein. Il est alors d'usage d'effectuer des test de martingalité de certains processus. L'hypothèse de martingalité des actifs actualisés, équivalente à l'hypothèse d'absence d'opportunité d'arbitrage, est fondamentale dans tout modèle financier et doit être vérifiée à chaque instant. Dans le cas présent, rappelons la relation suivante :

$$P(t, T + S) = \mathbb{E}_t^Q \left[e^{-\int_t^{T+S} r_s ds} \right] \quad (1)$$

$$= \mathbb{E}_t^Q \left[e^{-\int_t^T r_s ds} \mathbb{E}_T^Q \left[e^{-\int_T^{T+S} r_s ds} \right] \right] \quad (2)$$

$$= \mathbb{E}_t^Q [P(t, T)P(T, T + S)] \quad (3)$$

et de manière générale, avec $t < T_1 < \dots < T_m$:

$$P(t, T_m) = \mathbb{E}_t^Q [P(t, T_1) \times P(T_1, T_2) \times \dots \times P(T_{m-1}, T_m)]$$

Nous allons essayer de vérifier numériquement cette relation, pour $t = 0$ et $T_i = i$: on compare les quantités $P(0, T)$ (donné par la courbe de taux initiale) et $\frac{1}{N} \sum_{i=1}^N \prod_{k=1}^T P^i(k-1, k)$, où $P^i(u, v)$ est le prix du zéro-coupon $P(u, v)$ tel qu'obtenu lors de la simulation i . La figure 9 trace la différence de ces deux quantités en fonction de l'horizon T . On souhaite cette quantité la plus petite possible, et on la compare pour vérifier cela à l'écart-type estimé de la

moyenne calculée, et à l'écart-type multiplié par 1.96, c'est-à-dire aux bornes de l'intervalle de confiance à 95% de cette variable (en supposant qu'elle suive une loi normale, ce qui est justifié par le théorème central limite).

L'erreur dévie vers les négatifs, mais reste largement entre les bornes de l'intervalle de confiance, ce qui nous rassure quant à la qualité des simulations.

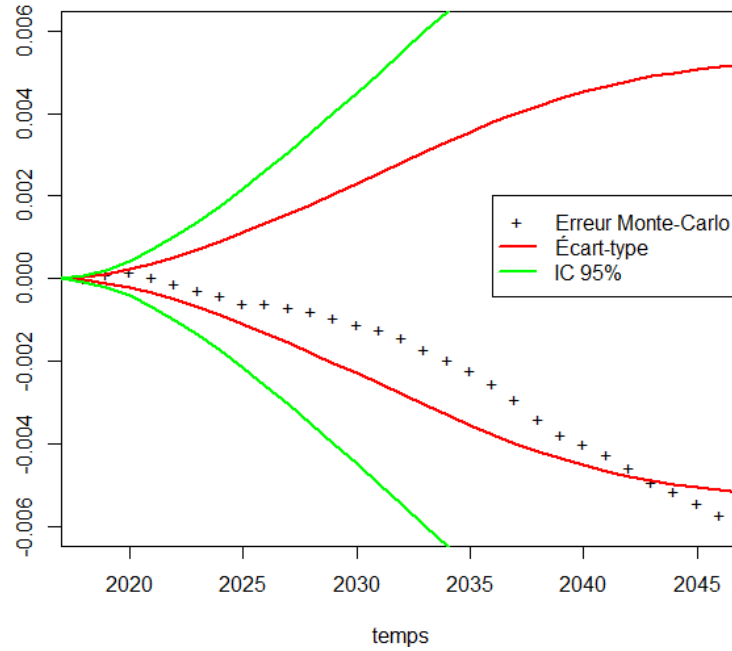


FIGURE 9 – Comparaison de l'erreur de simulation sur la moyenne des zéro-coupons, et leur écart-type. Intervalle de confiance gaussien à 95%. Calibration sur données de juin.

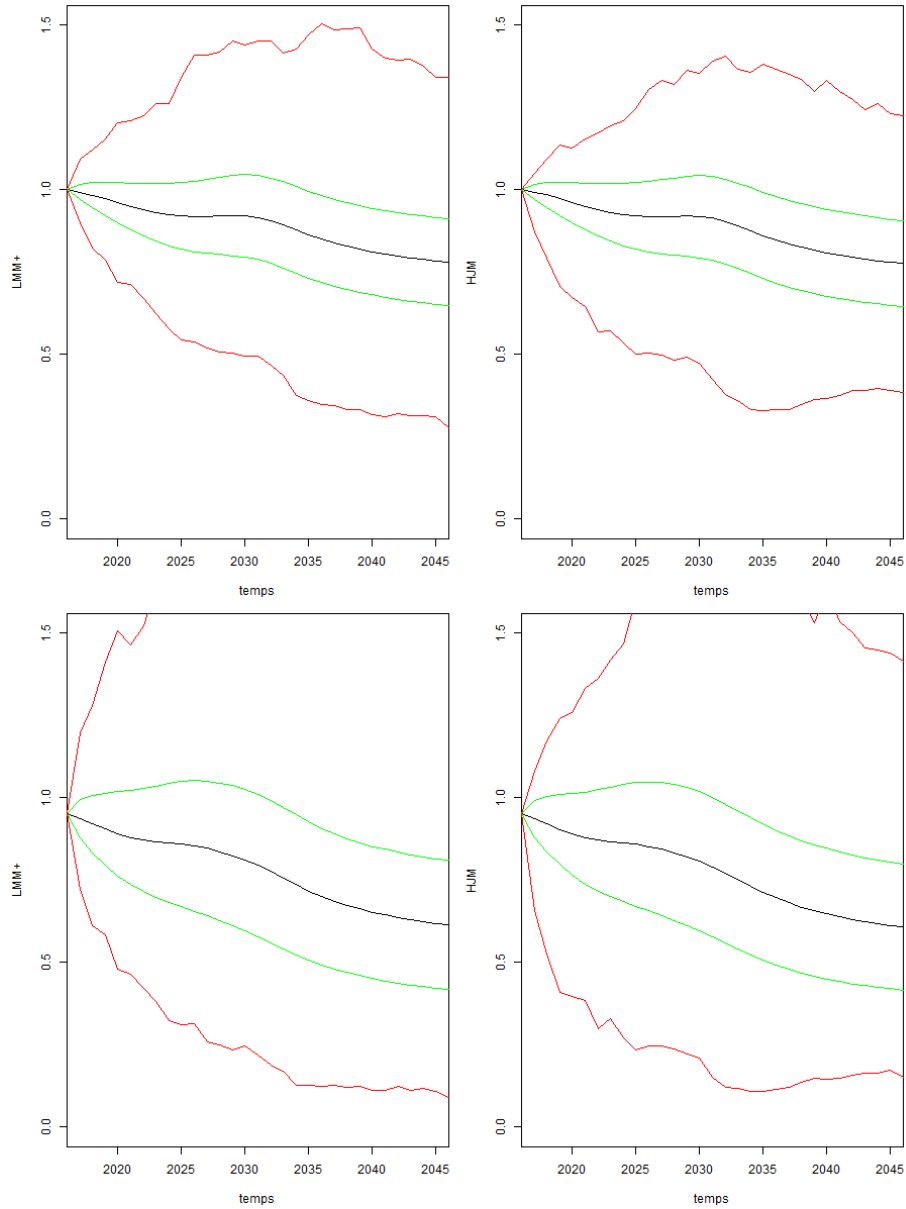


FIGURE 10 – En haut : moyenne (noir), écart-type (vert) et extrema (rouge) des prix des zéro-coupons, simulés 5000 fois sur 30 ans. En haut : zéro-coupon 5 ans. En bas : zéro-coupon 10 ans. À gauche : simulation par le modèle LMM+. À droite : simulation par le modèle type HJM de Schwartz et Trolle, deux facteurs. Calibration sur données de juin.

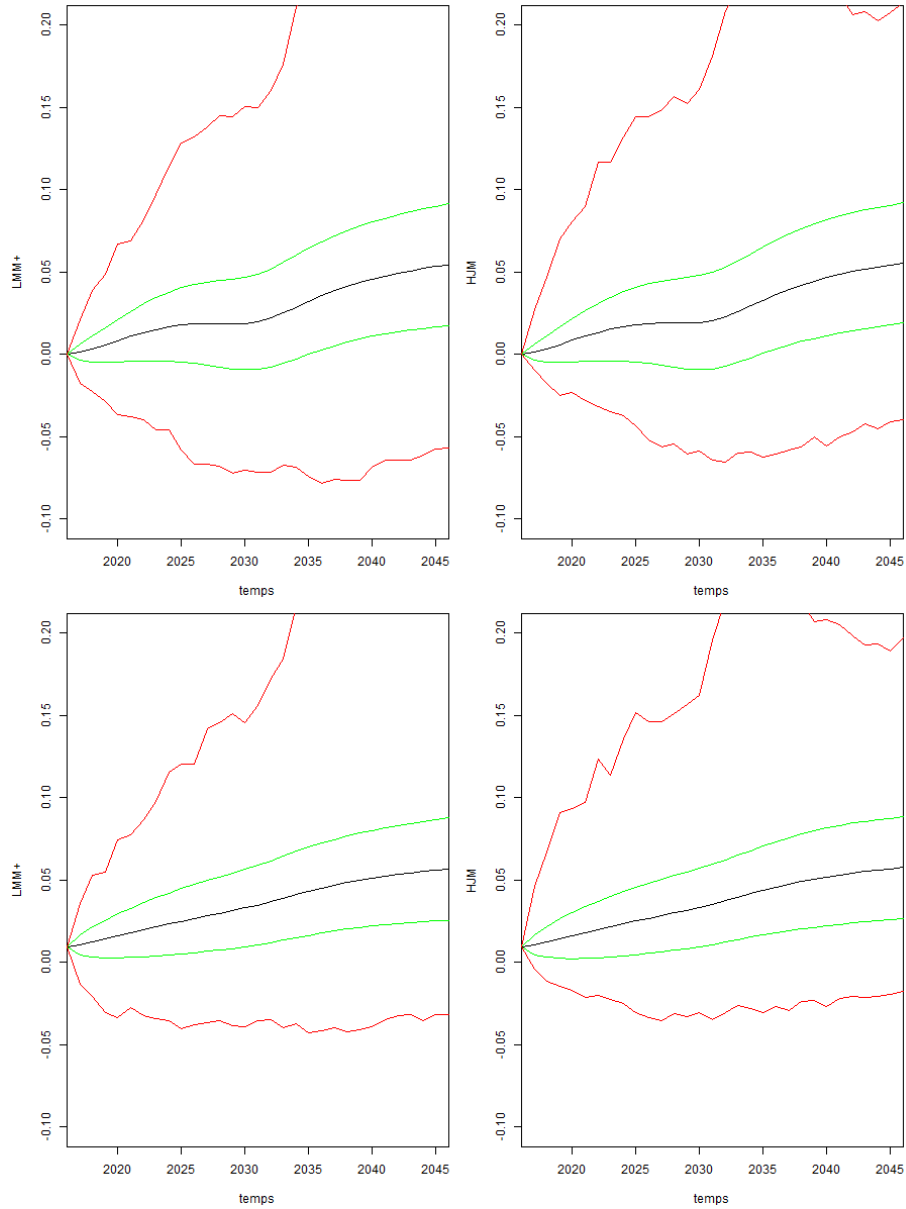


FIGURE 11 – En haut : moyenne (noir), écart-type (vert) et extrema (rouge) des taux swap, simulés 5000 fois sur 30 ans. En haut : taux swap 5 ans. En bas : taux swap 20 ans. À gauche : simulation par le modèle LMM+. À droite : simulation par le modèle type HJM de Schwartz et Trolle, deux facteurs. Calibration sur données de juin.

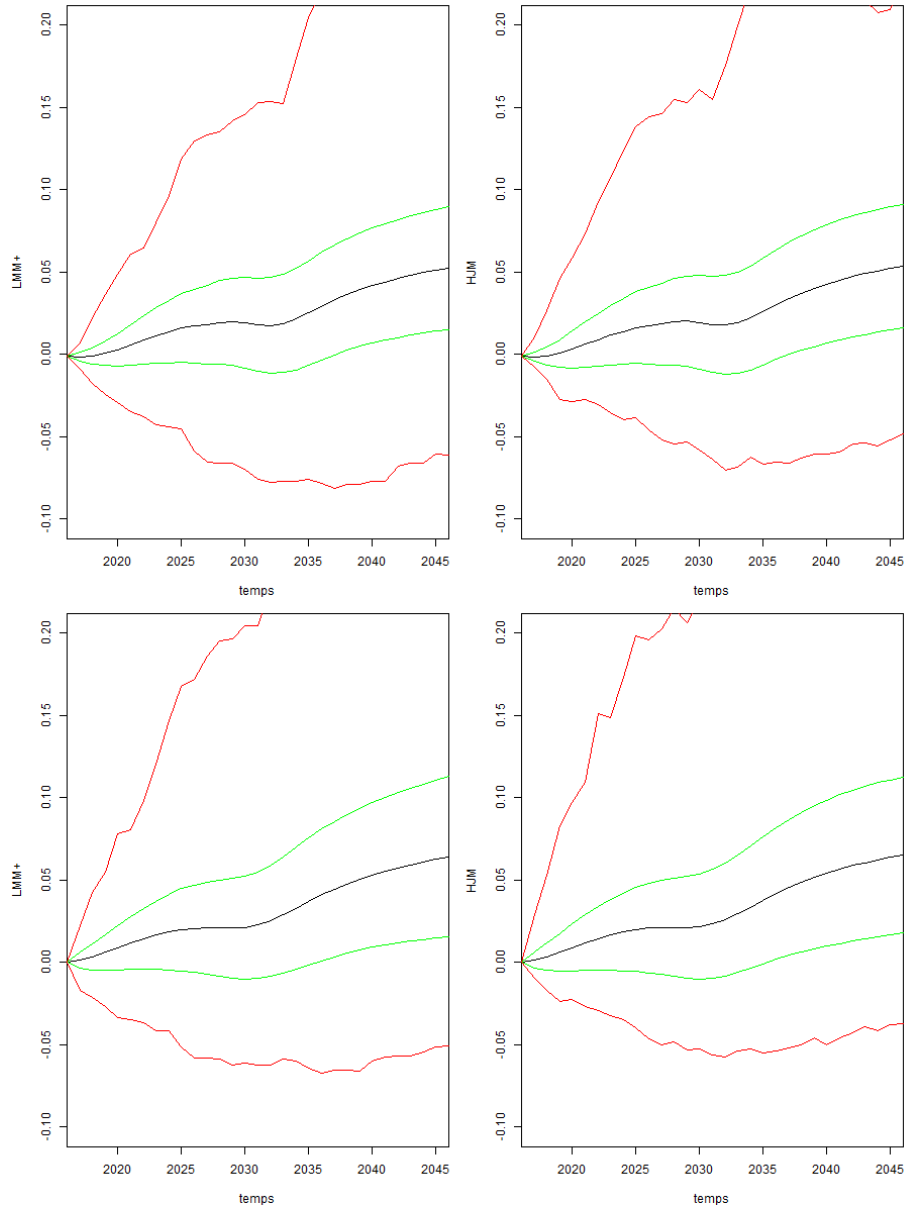


FIGURE 12 – En haut : moyenne (noir), écart-type (vert) et extrema (rouge) des taux LIBOR, simulés 5000 fois sur 30 ans. En haut : taux LIBOR 1 an. En bas : taux LIBOR 5 ans. À gauche : simulation par le modèle LMM+. À droite : simulation par le modèle type HJM de Schwartz et Trolle, deux facteurs. Calibration sur données de juin.

3 Intégration et comparaison

3.1 Résultats ALM

Voici la VIF et le BEL obtenus (relativement à ceux du modèle LMM+) :

	LMM+	HJM (1)	HJM (2)
BEL	100%	99.39%	99.83%
VIF	100%	143.86%	143.84%

La première colonne correspond aux simulations effectuées avec le modèle LMM+, en annulant toute corrélation entre les taux et les autres classes d'actifs. Dans les simulations de la seconde colonne ont été remplacés les prix des zéro-coupons du LMM+ par ceux générés avec le modèle étudié. Enfin, dans la dernière colonne, les zéro-coupons ont été remplacés, et les actions ont été revalorisées selon les prix des zéro-coupons. On voit que cette dernière opération a peu d'effets.

En revanche, le BEL est diminué de près d'environ 0.6% en passant au modèle HJM. La VIF, bénéficiant sûrement d'un effet ciseau, augmente de 44%. En effet la VIF résulte de la différence des flux entrants et sortant, et ne représente qu'une petite proportion de ceux-ci. Si l'ensemble des flux entrants, ou sortants, varie d'un faible pourcentage, la VIF variera de ce même pourcentage multiplié par un facteur qui lui, est élevé (50, si la VIF représente 2% des flux dans un sens).

On observe de plus un écart de prix moyen sur les zéro-coupons, croissant jusqu'à 0.1% sur les zéro-coupon 10 ans. Ce pourrait être simplement un écart de convergence : si les zéro-coupons ont un prix dont l'écart-type est σ , leur moyenne sur n échantillons a pour écart-type $\frac{\sigma}{\sqrt{n}}$. Les deux grandeurs (avec $n = 5000$) sont du même ordre. Cela pourrait expliquer en partie la variation de BEL.

Enfin, le modèle ALM étant extrêmement complexe, toute variation même faible des données en entrée peut avoir un fort impact sur les indicateurs via des effets de seuil, ou les stratégies d'allocation qui ne sont pas linéaires.

3.2 Comparaison des modèles

La comparaison des modèles ne peut s'arrêter à la simple comparaison des résultats de l'entreprise, pour au moins deux raisons :

- Choisir un modèle pour valoriser les résultats de l'entreprise serait faire preuve d'une honnêteté intellectuelle douteuse : un modèle donne un point de vue sur l'activité financière de l'entreprise. Aussi, s'il donne des chiffres valorisants, ce peut être parce qu'il néglige certains risques, mieux pris en compte par d'autres modélisations. Il faut donc s'assurer que les scénarios dangereux pour la compagnie sont quand-même bien reproduits.
- Selon le contexte économique, le modèle le plus avantageux peut varier d'un exercice à l'autre. Un changement de modélisation a un impact dans les chiffres présentés (l'« impact modèle »), et empêche ceux-ci d'être comparables d'un exercice à l'autre : si la VIF augmente, est-ce du fait des conditions de marché, de la stratégie de l'entreprise, ou juste du modèle ? On ne peut alors pas répondre à cette question.

Il faut donc faire des choix quant à ce que l'on souhaite modéliser. Voici un comparatif des deux modèles, LMM+ et le modèle HJM de Schwartz et Trolle. Pour rappel, le LMM+ est une extension d'un modèle classique (LMM), très présent dans le marché de l'assurance, utilisé actuellement par Predica, et que cette dernière souhaite remettre en cause. Quant au modèle HJM, il s'agit du candidat qui fait l'objet de ce mémoire.

	LMM+	HJM
Taux considéré	LIBOR <i>forward</i> : $f_j(t) = \frac{P(t, T_j) - P(t, T_{j+1})}{(T_{j+1} - T_j)P(t, T_{j+1})}$	Taux à terme instantané : $f(t, T) = -\partial_T \ln P(t, T)$
Diffusion	$df_j(t) = (f_j(t) + \delta) \left(\mu_j(t)dt + \overrightarrow{\sigma_{t,T}} \cdot d\overrightarrow{W}_t \right)$	$df(t, T) = \mu_{t,T}dt + \overrightarrow{\sigma_{t,T}} \cdot d\overrightarrow{W}_t$
Domaine de définition	$f_j(t) \in]-\delta, +\infty[$	$f(t, T) \in]-\infty, +\infty[$
Volatilité : $\overrightarrow{\sigma_{t,T}} \cdot \overrightarrow{W}_t =$	$\sqrt{V_t} g_{j-m(t)} \left(\beta_{j-m(t)}^1 dW_t^1 + \beta_{j-m(t)}^2 dW_t^2 \right)$ $m(t) = \inf[t, +\infty[\cap\{T_1, \dots, T_n\}]$	$\sqrt{v_t^1} \sigma_{f,1}(T-t) dW_t^1 + \sqrt{v_t^2} \sigma_{f,2}(T-t) dW_t^2$
Structure de volatilité	$g_k = (a + bT_k) e^{-cT_k}$	$\sigma_{f,i}(\tau) = (\alpha_0^i + \alpha_1^i \tau) e^{-\gamma_i \tau}$
Dynamique de la volatilité	$dV_t = \kappa(\theta - V_t)dt + \varepsilon \sqrt{V_t} dZ_t$	$dv_t^i = \kappa_i(\theta_i - v_t^i)dt + \sigma_i \sqrt{v_t^i} dZ_t^i$
Corrélation taux/volatilité	$d \left\langle \frac{W_t^1 + W_t^2}{\sqrt{2}}, Z_t \right\rangle = \rho dt$	$d \langle W_t^i, Z_t^i \rangle = \rho_i dt$
Données sources	$f_j(0) \ (j \in \{1, \dots, j_{\max}\})$ reconstruits à partir de taux swap observés sur le marché $\beta_j^i \ (j \in \{1, \dots, j_{\max}\}, i = 1, 2)$ d'après étude de Moody's sur des données historiques	$f(0, T) \ (T \in [0, T_{\max}])$ reconstruite à partir de taux swap observés sur le marché
Données de calibration	Diverses swaptions Structure de volatilité à terme, d'après étude de Moody's sur des données historiques	Diverses swaptions
Nombre de facteurs de risque	2 (taux) + 1 (volatilité)	2 (taux) + 2 (volatilité)
Nombre de paramètres	9	14

Commentons point par point ce tableau :

Le LMM+ modélise des taux LIBOR. La modélisation peut paraître plus grossière par rapport au HJM qui diffuse continûment la courbe des taux, mais est aussi plus pragmatique – la notion de taux instantané restant très théorique.

Le LMM+ suit une diffusion log-normale, avec un décalage rendant possibles les taux négatifs. À l'inverse, le HJM suit une diffusion dont la forme est plutôt gaussienne. Les taux négatifs sont donc aussi représentés, mais l'absence de borne inférieure et l'indépendance de la volatilité vis-à-vis du niveau des taux rend la diffusion symétrique par rapport à la courbe initiale. Ce n'est pas ce qu'on observe sur les données historiques : la volatilité des taux est en général proportionnelle aux taux quand ceux-ci sont bas, ce qui est caractéristique des modèles log-normaux. Le LMM+ devrait être plus à même de rendre compte de ce comportement, mais le décalage assez élevé (45%, à comparer aux taux qui ne dépassent guère -5% dans les simulations) rend ce phénomène très discret dans le modèle.

De plus, ce paramètre δ est compliqué à calibrer. Le changer d'un trimestre à l'autre n'aurait aucun sens financier (les taux devraient garder le même domaine de définition), mais il est compliqué de lui donner une valeur pérenne et fiable. Moody's, d'après des études de sensibilité et de calibration, suggère cette valeur de 45% : c'est ce qui est utilisé par Predica, sans subir de calibration supplémentaire. Il est à noter qu'une valeur plus faible du *shift* peu engendrer des simulations irréalistes dans le cas de scénarios de chocs à la baisse. Le HJM, en relevant la contrainte de la borne inférieure, apporte aussi une certaine stabilité au modèle.

Les deux modèles ont des volatilités bien similaires : des diffusions CIR, et une structure de type Rebonato ($\sigma(\tau) = (a + b\tau)e^{-c\tau}$). Ce qui les différencie sur ce plan-là, est l'utilisation des différents facteurs de risque. Dans le LMM+, leurs effets sont répartis sur la courbe des taux selon les fonctions β_j^1 et β_j^2 , et une unique fonction V_t assure la stochasticité de leur variance.

Dans le HJM, les fonctions $\sigma_{f,i}$ assurent de manière indépendante la répartition des effets de chaque facteur sur la courbe et chaque facteur de risque possède sa propre volatilité stochastique. Il n'est en effet pas évident que les variances des taux de longs termes et de ceux de courts termes varient de la même manière : le modèle HJM seulement peut capter ces différences (au prix de paramètres supplémentaires à calibrer).

Le modèle LMM+ demande plus de données de calibration que le modèle HJM : en plus de la courbe des taux, il nécessite les fonctions β_j^i (données par Moody's, en l'occurrence), et la structure de la volatilité à long terme, aussi donnée par le cabinet. Ces données ont fait l'objet d'études historiques, ce qui a pour effet de rendre l'utilisateur dépendant du fournisseur de ces données (à moins de refaire ces mêmes travaux) et le modèle moins flexible (en cas de changement radical de contexte économique, ces données pourraient devenir caduques).

Au vu de ce comparatif, on pourra retenir de ce modèle HJM sa flexibilité, sa bonne calibration, et un traitement prudent des taux négatifs par l'absence de borne inférieure. De plus, il nécessite moins de données historiques pour sa calibration, mais cela se fait au prix de paramètres supplémentaires à calibrer. Enfin, il permet de distinguer les volatilités de chaque facteur de risque.

Conclusion

Nous avons étudié dans ce mémoire un modèle de taux multifacteur, à volatilité stochastique, et conforme au cadre HJM.

Une étude théorique en a été faite, jusqu'à la valorisation exacte d'options vanilles et une approximation du prix des swaptions. L'approche par fonction caractéristique, similaire au modèle d'Heston, a un coût non négligeable, et d'autres pistes comme la résolution de l'équation de Kolmogorov pourraient donc être explorées.

Nous avons également vu différents algorithmes permettant de résoudre numériquement les équations apparaissant dans les prix d'options, et leurs limites. Notamment, ceux présentés ne permettent que de valoriser les options très proches de la monnaie, les autres présentant trop d'oscillations dans leur intégrande. Afin de voir plus large et de calibrer le smile, il faudra envisager d'autres méthodes d'intégration : des méthodes adaptatives type Gauss-Kronrod, assez sûres mais coûteuses, ou des quadratures type Filon, qui dépendent de la fréquence d'oscillation.

La calibration sur les prix de swaptions, calculés à partir des volatilités de marché et de la courbe de taux réglementaire, fut correcte. Mais en raison du grand nombre de paramètres, et de tous les minima locaux qu'ils induisent sur la fonction objective, trouver un minimum global se révèle difficile. La lenteur des calculs constitue aussi un obstacles à la calibration.

Enfin, la valorisation du BEL est cohérent avec le LMM+, mais la VIF augmente fortement. Il faudrait étudier plus en profondeur les résultats de simulation, voire en effectuer plus que 5000, pour pouvoir choisir un modèle pour les exercices à venir. En attendant, ce modèle possède certains avantages face au LMM+ qui est aujourd'hui un modèle de référence dans le milieu. Notamment, il s'affranchit du choix de la borne inférieure des taux, et, en calibrant davantage de paramètres, permet de ne pas s'appuyer sur des études historiques.

Références

- [1] E. S. Schwartz and A. B. Trolle, “A general stochastic volatility model for the pricing of interest rate derivatives,” *Review of Financial Studies*, vol. 22, no. 5, pp. 2007–2057, 2009.
- [2] D. Heath, R. Jarrow, and A. Morton, “Bond pricing and the term structure of interest rates : A new methodology for contingent claims valuation,” *Econometrica*, vol. 60, pp. 77–105, Jan. 1992.
- [3] C. Munk, “Stochastic duration and fast coupon bond option pricing in multi-factor models,” *Review of Derivatives Research*, no. 3, pp. 157–181, 1999.
- [4] S. Bochkhanov, “ALGLIB.”
- [5] R. Sedgewick, *Algorithms*. Addison-Wesley, 1984.
- [6] M. Matsumoto and T. Nishimura, “Mersenne twister : A 623-dimensionally equidistributed uniform pseudorandom number generator,” *ACM Transactions on Modeling and Computer Simulation*, vol. 8, pp. 3–30, Jan. 1998.
- [7] H. Geman, N. El Karoui, and J.-C. Rochet, “Changes of numéraire, changes of probability measure and option pricing,” *Journal of Applied Probability*, vol. 32, pp. 443–458, June 1995.
- [8] D. Lamberton and B. Lapeyre, “Introduction au calcul stochastique appliqué à la finance.” <http://cermics.enpc.fr/~bl/PS/PROAL/calculsto.pdf>, 1997. Seconde édition.
- [9] EIOPA, “Solvency II risk-free interest rate term structures.” <https://eiopa.europa.eu/regulation-supervision/insurance/solvency-ii-technical-information/risk-free-interest-rate-term-structures>. Consulté pour la dernière fois le 2 février 2017.

Annexes

A Quelques méthodes numériques

1 Algorithme de Nelder-Mead

Aussi appelé méthode des simplexes, cet algorithme d'optimisation s'appuie sur un simplexe (polyèdre de \mathbb{R}^n à $n + 1$ sommets), et va le déformer de sorte que tous ses sommets convergent vers le même minimum local d'une fonction $f : \mathcal{U} \subset \mathbb{R}^n \rightarrow \mathbb{R}$

On l'initialise avec $x_0, \dots, x_n \in \mathcal{U}$ non-alignés sur un même hyperplan. Ensuite, à chaque étape (on peut supposer que $f(x_0) \leq \dots \leq f(x_n)$, quitte à réarranger les points) :

- On calcule le centre de gravité $g = \frac{1}{n+1} \sum_{i=0}^n x_i$
- On calcule le symétrique de x_n par rapport à g : $r = 2g - x_n$
- Si $f(r) < f(x_{n-1})$: il faut étirer le simplexe avec $e = g + 2(g - x_n)$
 - Si $f(e) < f(r)$, on remplace x_n par e et on recommence
 - Sinon, on le remplace par r et on recommence
- Si $f(r) \geq f(x_{n-1})$: on peut réduire le simplexe avec $c = g + (g - x_n)/2$
 - Si $f(c) < f(x_{n-1})$, on remplace x_n par c et on recommence
 - Sinon, on réduit le simplexe par une homothétie de facteur $\frac{1}{2}$ par rapport à x_0 . On réitère.

2 Méthodes de Runge-Kutta : tabulation de coefficients

Conformément à la littérature sur le sujet, les coefficients seront notés de la manière suivante :

c_1				
c_2	a_1^1			
\vdots	\vdots	\ddots		
c_q	a_1^{q-1}	\dots	a_{q-1}^{q-1}	
(ordre)	b_1	\dots	b_{q-1}	b_q

2.1 Méthode RK4

0				
$\frac{1}{2}$	$\frac{1}{2}$			
$\frac{1}{2}$	0	$\frac{1}{2}$		
1	0	0	1	
(5)	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$

2.2 Méthode de Cash-Karp

0						
$\frac{1}{5}$	$\frac{1}{5}$					
$\frac{3}{10}$	$\frac{3}{40}$	$\frac{9}{40}$				
$\frac{3}{5}$	$\frac{3}{10}$	$-\frac{9}{10}$	$\frac{6}{5}$			
1	$-\frac{11}{54}$	$\frac{5}{2}$	$-\frac{70}{27}$	$\frac{35}{27}$		
$\frac{7}{8}$	$\frac{1631}{55296}$	$\frac{175}{512}$	$\frac{575}{138524}$	$\frac{44275}{110592}$	$\frac{253}{4096}$	
(5)	$\frac{37}{378}$	0	$\frac{250}{621}$	$\frac{125}{594}$	0	$\frac{512}{1771}$
(4)	$\frac{2825}{27648}$	0	$\frac{18575}{48384}$	$\frac{13525}{55296}$	$\frac{277}{14336}$	$\frac{1}{4}$

3 Quadratures de Gauss

3.1 Quadrature de Gauss-Legendre

Pour cette méthode, on intègre sur $I = [-1, +1]$ avec une fonction de pondération $w = 1$. On utilise comme nœuds les racines des polynômes de Legendre P_n et comme poids $\omega_i = \frac{2}{(1-x_i^2)P'(x_i)^2}$.

La quadrature de Gauss-Legendre à cinq points utilise les valeurs suivantes :

Nœuds	Poids
0	$\frac{128}{225}$
$\pm \frac{1}{3}\sqrt{5 - 2\sqrt{\frac{10}{7}}}$	$\frac{322+13\sqrt{70}}{900}$
$\pm \frac{1}{3}\sqrt{5 + 2\sqrt{\frac{10}{7}}}$	$\frac{322-13\sqrt{70}}{900}$

3.2 Quadrature de Gauss-Kronrod

Cette méthode est double, car elle s'appuie sur une quadrature de Gauss à n nœuds pour établir un premier résultat, puis une quadrature avec $n + 1$ points supplémentaires (les n premiers sont réutilisés) pour renvoyer un second résultat d'ordre supérieur. La différence donne une estimation de l'erreur faite, qui peut être ensuite utilisée pour adapter la taille de l'intervalle d'intégration.

Voici les valeurs utilisées pour la méthode de Gauss-Kronrod la plus courante, (G7, K15) :

Quadrature de Gauss	
Nœuds	Poids
$\pm 0.94910\ 79123\ 42759$	0.12948 49661 68870
$\pm 0.74153\ 11855\ 99394$	0.27970 53914 89277
$\pm 0.40584\ 51513\ 77397$	0.38183 00505 05119
0	0.41795 91836 73469

Quadrature de Kronrod	
Nœuds	Poids
$\pm 0.99145\ 53711\ 20813$	0.02293 53220 10529
$\pm 0.94910\ 79123\ 42759$	0.06309 20926 29979
$\pm 0.86486\ 44233\ 59769$	0.10479 00103 22250
$\pm 0.74153\ 11855\ 99394$	0.14065 32597 15525
$\pm 0.58608\ 72354\ 67691$	0.16900 47266 39267
$\pm 0.40584\ 51513\ 77397$	0.19035 05780 64785
$\pm 0.20778\ 49550\ 07898$	0.20443 29400 75298
0	0.20948 21410 84728

4 Simulation de variables aléatoires

Afin de réaliser des simulations de Monte-Carlo, on a besoin de simuler les variables aléatoires sous-jacentes. Or, les machines sont incapables de réaliser des tirages réellement aléatoires : on a donc recours à des algorithmes de générations de nombres *pseudo-aléatoires*. Ce sont des algorithmes déterministes mais qui donnent des suites de nombres difficilement prévisibles, et aux propriétés statistiques proches des suites i.i.d. de variables uniformes sur le segment $[0, 1]$.

C'est à partir de simulations de variables uniformes que l'on peut ensuite simuler d'autres lois (gaussienne, exponentielle...).

4.1 Générateurs de nombres pseudo-aléatoires

L'algorithme le plus simple reste le générateur par congruence linéaire ou LCG. Il a en paramètre un entier M qui déterminera la précision de l'algorithme, et deux entiers a et b . On définit alors la suite $(x_n)_{n \in \mathbb{N}}$ ainsi :

$$\begin{cases} x_0 & \in \{0, \dots, M-1\} \\ x_{n+1} & = a \times x_n + b \mod M \end{cases}$$

Sedgewick (dans [5]) préconise les paramètres suivants :

$$\begin{aligned} M &= 10^8 \\ a &= 31415821 \\ b &= 1 \end{aligned}$$

On en tire une suite $\mathcal{U}([0, 1])$ en divisant x_n par M .

Le générateur de référence aujourd'hui est le Mersenne-Twister que l'on trouvera dans son article original de 1998 ([6]). Celui utilisé dans nos programmes est en particulier MT19937-64, qui a une période de $2^{19937} - 1$. Il fonctionne en deux temps : une étape génère une suite d'entiers selon une récurrence d'ordre 312, et une étape transforme cette suite afin de la rendre moins prévisible. L'algorithme n'utilise que des opérations bit-à-bit.

4.2 Génération de variables gaussiennes

La méthode la plus courante pour simuler des variables aléatoires gaussiennes centrées et réduites est la méthode de Box-Muller : si $U, V \sim \mathcal{U}[0, 1]$ alors on

définit les deux variables suivantes :

$$X = \sqrt{-2 \ln U} \cos(2\pi V)$$

$$Y = \sqrt{-2 \ln U} \sin(2\pi V)$$

Alors X et Y sont indépendantes, de loi $\mathcal{N}(0, 1)$.

5 Comparaison des performances

5.1 Résolution de l'équation de Riccati

Ces tableaux comparent les performances de résolution de l'équation vérifiée par M , N_1 et N_2 par différents algorithmes. La même équation est résolue mille fois à chaque fois. Le temps requis est reporté dans la table, ainsi que la différence relative sur le module de $M(T_0)$ avec le résultat donné par ALGLIB, une librairie que nous prenons arbitrairement comme référence.

La mention "NC" indique que l'algorithme diverge, et renvoie un résultat non numérique au sens de C#.

$u = 10$	$T_0 = 20, T_1 = 30$		$T_0 = 1, T_1 = 21$		$T_0 = 20, T_1 = 21$		$T_0 = 1, T_1 = 2$	
Méthode	Erreur	Temps	Erreur	Temps	Erreur	Temps	Erreur	Temps
Chgt. var.	-0.033%	0.499	-0.104%	0.483	-0.02%	0.483	-9.02E-06	0.499
RK4 300 pas	NC	NC	NC	NC	NC	NC	-1.454%	0.608
ALGLIB	-	53.42	-	2.901	-	2.558	-	0.124
Dormand-Prince	1.08E-10	50.74	-9.20E-10	2.761	-1.20E-08	5.210	-1.05E-07	0.234
Cash-Karp	1.08E-10	40.18	-9.26E-10	3.12	-1.30E-08	3.634	-1.33E-07	0.187

$u = 1000$	$T_0 = 20, T_1 = 30$		$T_0 = 1, T_1 = 21$		$T_0 = 20, T_1 = 21$		$T_0 = 1, T_1 = 2$	
Méthode	Erreur	Temps	Erreur	Temps	Erreur	Temps	Erreur	Temps
Chgt. var.	0.0047%	0.483	-7,68E-06	0.499	-0.009%	0.483	-2,09E-06	0.483
RK4 300 pas	0.0574%	0.592	-3.54E-09	0.592	2,04E-09	0.592	4,59E-09	0.608
ALGLIB	-	0.889	-	0.171	-	0.405	-	0.093
Dormand-Prince	1.07E-10	1.700	6.04E-09	0.171	-1.08E-08	0.280	-3,88E-05	0.031
Cash-Karp	8.37E-11	1.200	-1.22E-08	0.146	1.22E-09	0.202	-2,13E-05	0.031

B Rappels de calcul stochastique

1 Formule d'Itô

Soient X^1, \dots, X^n des semi-martingales continues et $f : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$, de classe C^1 . Alors pour tout $t \geq 0$:

$$df(t, X_t^1, \dots, X_t^n) = \frac{\partial f}{\partial t} dt + \sum_{i=1}^n \frac{\partial f}{\partial X_i} dX_t^i + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial X^i \partial X^j} d\langle X_t^i, X_t^j \rangle$$

2 Changement de numéraire

On sait que sous la mesure risque-neutre Q , les actifs actualisés sont des martingales. Dans leur article de 1995 [7], El Karoui, Geman et Rochet ont généralisé cette propriété.

Un numéraire est un processus de prix $(N_t)_{t \geq 0}$, presque-sûrement positif et qui ne délivre pas de dividende. Le processus d'accumulation $\left(e^{\int_0^t r_s ds}\right)_{t \geq 0}$ constitue par exemple un numéraire.

À tout numéraire N , on peut associer une mesure Q^N par sa dérivée de Radon-Nikodym :

$$\frac{dQ^N}{dQ} \Big|_{\mathcal{F}_t} = \frac{N(t)}{N(0)e^{\int_0^t r_s ds}}$$

Il a été montré que, sous cette nouvelle probabilité, les actifs exprimés sous ce numéraire sont des martingales. Ainsi pour tout actif S et tout $t \leq T$:

$$\mathbb{E}_t^{Q^N} \left[\frac{S_T}{N(T)} \right] = \frac{S_t}{N(t)}$$

Plus généralement, pour deux numéraires M et N :

$$\frac{dQ^M}{dQ^N} \Big|_{\mathcal{F}_t} = \frac{M(t)N(0)}{M(0)N(t)}$$

3 Formule de Bachelier

Soit X une variable gaussienne : $X \sim \mathcal{N}(X_0, s^2)$. Alors :

$$\begin{aligned} \mathbb{E}[(X - K)_+] &= \int_K^{+\infty} (x - K) \frac{e^{-\frac{(x-X_0)^2}{2s^2}}}{\sqrt{2\pi}s} ds \\ &= \int_{\frac{K-X_0}{s}}^{+\infty} ys \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy + (X_0 - K) \int_{\frac{K-X_0}{s}}^{+\infty} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy \\ &= \frac{s}{\sqrt{2\pi}} \left[-e^{-\frac{y^2}{2}} \right]_{\frac{K-X_0}{s}}^{+\infty} + (X_0 - K) \left(1 - F_{\mathcal{N}} \left(\frac{K - X_0}{s} \right) \right) \\ &= s \times \frac{e^{-\frac{(X_0-K)^2}{2s^2}}}{\sqrt{2\pi}} + (X_0 - K) F_{\mathcal{N}} \left(\frac{X_0 - K}{s} \right) \end{aligned}$$

4 Formule de Black

Considérons X , une variable aléatoire log-normale : $X = X_0 e^{-\frac{s^2}{2} + Y}$, avec $Y \sim \mathcal{N}(0, s^2)$ de sorte que $\mathbb{E}[X] = X_0$. On a :

$$\begin{aligned} \mathbb{E}[(X - K)_+] &= \int_{-\infty}^{+\infty} \left(X_0 e^{-\frac{s^2}{2} + sy} - K \right)_+ \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy \\ &= X_0 \int_{-d_2}^{+\infty} e^{-\frac{s^2}{2} + y - \frac{y^2}{2}} \frac{dy}{\sqrt{2\pi}} - K \int_{-d_2}^{+\infty} e^{-\frac{y^2}{2}} \frac{dy}{\sqrt{2\pi}} \\ &= X_0 \int_{-d_2 - s^2}^{+\infty} e^{-\frac{z^2}{2}} \frac{dz}{\sqrt{2\pi}} - K(1 - F_{\mathcal{N}}(-d_2)) \\ &= X_0(1 - F_{\mathcal{N}}(-d_1)) - K F_{\mathcal{N}}(d_2) \\ &= X_0 F_{\mathcal{N}}(d_1) - K F_{\mathcal{N}}(d_2) \end{aligned}$$

avec :

$$\begin{aligned} d_2 &= \frac{1}{s} \left(\ln \frac{X_0}{K} - \frac{s^2}{2} \right) \\ d_1 &= d_2 + s^2 = \frac{1}{s} \left(\ln \frac{X_0}{K} + \frac{s^2}{2} \right) \end{aligned}$$

D'où la formule suivante :

$$\text{Black}(X_0, K, s) = \mathbb{E}[(X - K)_+] = X_0 F_{\mathcal{N}} \left(\frac{1}{s} \ln \frac{X_0}{K} + \frac{s}{2} \right) - K F_{\mathcal{N}} \left(\frac{1}{s} \ln \frac{X_0}{K} - \frac{s}{2} \right)$$

5 Volatilités implicites

Ces deux formules (Bachelier et Black-Scholes) sont croissantes avec la volatilité et continues, ce qui permet, pour tout prix de call (ou de put) compris dans un certain intervalle, de lui définir une volatilité implicite comme l'inverse du prix de marché par la fonction de pricing. C'est cette volatilité qui est donnée, plutôt que le prix, pour certains produits dérivés courants dans les logiciels d'information financière comme Bloomberg.

Cette grandeur permet de mieux comparer les prix des options entre eux : on remarque par exemple que le prix d'une swaption est globalement proportionnel à son ténor, et est très sensible à la courbe de taux en vigueur. Prendre la volatilité implicite retire ces dépendances.

6 Convergence des schémas d'approximation des EDS

Prenons un schéma d'approximation de Y solution d'une EDS sur $[0, T]$, donnant une solution $Y_h(T)$ dépendant d'un pas $h > 0$. On dit que ce schéma a un ordre fort de convergence γ si :

$$\mathbb{E} [|Y_h(T) - Y(T)|] \underset{h \rightarrow 0}{=} \mathcal{O}(h^\gamma)$$

et on dit qu'il a un ordre faible de convergence γ si :

$$|\mathbb{E} [g(Y_h(T))] - \mathbb{E} [g(Y(T))]| \underset{h \rightarrow 0}{=} \mathcal{O}(h^\gamma)$$

B Rappels de calcul stochastique

pour toute fonction g au moins $2\gamma + 1$ fois continûment différentiable et à croissance polynomiale.

C Modèles classiques : détails de calcul

1 Modèle de Vašíček

Calculons les moments de l'intégrale de r_t . Sachant que :

$$r_t = \theta + (r_s - \theta)e^{-\kappa(t-s)} + \sigma \int_s^t e^{-\kappa(t-u)} dW_u^Q$$

L'intégrale de r_t vaut :

$$\begin{aligned} \int_t^T r_s ds &= \int_t^T \theta ds + (r_t - \theta) \int_t^T e^{-\kappa(s-t)} ds + \sigma \int_t^T \int_t^s e^{-\kappa(s-u)} dW_u^Q ds \\ &= \theta(T-t) + (r_t - \theta) \frac{1 - e^{-\kappa(T-t)}}{\kappa} + \sigma \int_t^T \int_u^T e^{-\kappa(s-u)} ds dW_u^Q \\ &= \theta(T-t) + (r_t - \theta) \frac{1 - e^{-\kappa(T-t)}}{\kappa} + \frac{\sigma}{\kappa} \int_t^T (1 - e^{-\kappa(T-u)}) dW_u^Q \end{aligned}$$

On obtient alors :

$$\begin{aligned} \mathbb{E}_t^Q \left[\int_t^T r_s ds \right] &= \theta(T-t) + (r_t - \theta) \frac{1 - e^{-\kappa(T-t)}}{\kappa} \\ \mathbb{V}_t^Q \left[\int_t^T r_s ds \right] &= \mathbb{V}_t^Q \left[\frac{\sigma}{\kappa} \int_t^T (1 - e^{-\kappa(T-u)}) dW_u^Q \right] \\ &= \frac{\sigma^2}{\kappa^2} \int_t^T (1 - e^{-\kappa(T-u)})^2 du \\ &= \frac{\sigma^2}{\kappa^3} \left[\kappa(T-t) - 2(1 - e^{-\kappa(T-t)}) + (1 - e^{-2\kappa(T-t)}) \right] \end{aligned}$$

2 Modèle CIR

Rappelons que le processus CIR vérifie l'EDS suivante :

$$dr_t = \kappa(\theta - r_t)dr_t + \sigma\sqrt{r_t}dW_t^Q$$

Proposition C.1 (Condition de Feller). *Le processus $(r_t)_{t \geq 0}$ étant défini comme ci-dessus :*

- si $\frac{2\kappa\theta}{\sigma^2} < 1$, r possède presque-sûrement un point d'annulation sur $]0, +\infty[$
- si $\frac{2\kappa\theta}{\sigma^2} \geq 1$ r est presque-sûrement strictement positif sur $[0, +\infty[$ entier

Démonstration. Cette démonstration est largement inspirée de [8]. Introduisons la fonction suivante :

$$s : x \in]0, +\infty[\mapsto \int_1^x e^{t\frac{2\kappa}{\sigma^2}} t^{-\frac{2\kappa\theta}{\sigma^2}} dt$$

Et pour $0 < \varepsilon < r_0 < M$ les temps d'arrêt :

$$\begin{aligned} \tau_\varepsilon &= \inf \{t \geq 0 \mid r_t < \varepsilon\} \\ \tau_M &= \inf \{t \geq 0 \mid r_t > M\} \\ \tau &= \tau_\varepsilon \wedge \tau_M \end{aligned}$$

En dérivant s deux fois, on arrive à la relation suivante pour tout $x > 0$:

$$\frac{\sigma^2}{2} s''(x) = -(\theta - x) s'(x)$$

Or l'application du lemme d'Itô au processus $(s(r_{t \wedge \tau}))_{t \geq 0}$ donne :

$$\begin{aligned} s(r_{t \wedge \tau}) &= s(r_0) + \int_0^t s'(r_s) \mathbb{I}_{s < \tau} dr_s + \frac{1}{2} \int_0^t s''(r_s) \mathbb{I}_{s < \tau} d\langle r \rangle_s \\ &= s(r_0) + \int_0^{t \wedge \tau} s'(r_s) [\kappa(\theta - r_s) ds + \sigma \sqrt{r_s} dW_s] - \frac{1}{2} \int_0^{t \wedge \tau} \frac{2(\theta - x)}{\sigma^2} s'(r_s) \sigma^2 r_s ds \\ &= s(r_0) + \int_0^{t \wedge \tau} \sigma s'(r_s) \sqrt{r_s} dW_s \end{aligned}$$

Sachant que s est continue sur son intervalle de définition, et que $r_{t \wedge \tau}$ est borné, $s(r_{t \wedge \tau})$ est aussi borné. De plus, pour tout $s \in [0, \tau]$:

$$s'(r_s) = e^{r_s \frac{2\kappa}{\sigma^2} - \frac{2\kappa\theta}{\sigma^2}} > e^{\varepsilon \frac{2\kappa}{\sigma^2} - \frac{2\kappa\theta}{\sigma^2}} = s_{min} > 0$$

Et, en prenant la variance de l'expression précédente :

$$\begin{aligned} \mathbb{V}[s'(r_s)] &= \int_0^t \mathbb{E}[\sigma^2 s'(r_s)^2 r_s \mathbb{I}_{s < \tau}] ds \\ &\geq \sigma^2 s_{min}^2 \varepsilon \int_0^t \mathbb{E}[\mathbb{I}_{s < \tau}] ds = \sigma^2 s_{min}^2 \varepsilon \mathbb{E}[t \wedge \tau] \end{aligned}$$

Donc $\mathbb{E}[t \wedge \tau] < \frac{\mathbb{V}[s'(r_s)]}{\sigma^2 s_{min}^2 \varepsilon}$ pour tout $t \geq 0$, ce qui implique que τ est presque-sûrement fini.

En prenant cette fois l'espérance :

$$\mathbb{E}[s(r_{t \wedge \tau})] = s(r_0)$$

Le temps d'arrêt τ est fini presque-sûrement, et $s(r_{t \wedge \tau})$ est borné. On fait donc tendre t vers l'infini pour aboutir à :

$$\begin{aligned} s(r_0) &= \mathbb{E}[s(r_\tau)] \\ &= \mathbb{P}(\tau_\varepsilon < \tau_M) \mathbb{E}[s(r_\tau) | \tau_\varepsilon < \tau_M] + \mathbb{P}(\tau_\varepsilon > \tau_M) \mathbb{E}[s(r_\tau) | \tau_\varepsilon > \tau_M] \\ &= \mathbb{P}(\tau_\varepsilon < \tau_M) s(\tau) + \mathbb{P}(\tau_M < \tau_\varepsilon) s(M) \end{aligned}$$

Remarquons que si $\frac{2\kappa\theta}{\sigma^2} \geq 1$, l'intégrale qui définit s ne converge pas au voisinage de zéro et $\lim_{x \rightarrow 0} s(x) = -\infty$. Dans ce cas, d'après la dernière relation, $\mathbb{P}(\tau_0 < \tau_M) = 0$. Cela est vrai indépendamment de M , or r est continu, donc $\tau_M \xrightarrow{M \rightarrow \infty} +\infty$ puis $\mathbb{P}(\tau_0 < \infty) = 0$ donc r ne s'annule jamais.

Dans le cas contraire, $s(0)$ existe. Par continuité de r sur chaque trajectoire, τ_ε croît continûment. De plus s est continue, on peut alors passer à la limite :

$$s(r_0) = \mathbb{P}(\tau_0 < \tau_M) s(0) + \mathbb{P}(\tau_M < \tau_0) s(M)$$

Sachant que s diverge au voisinage de l'infini, on est alors assurés que $\lim_{M \rightarrow \infty} \mathbb{P}(\tau_M < \tau_0) = 0$, i.e. $\mathbb{P}(\tau_0 < \infty) = 1$. \square

D Modèle HJM de Schwartz et Trolle : détails de calculs

1 Taux forward

Exprimons le taux forward en fonction des facteurs de risques :

$$\begin{aligned} f(t, T) - f(0, T) &= \int_0^t df(u, T) \\ &= \int_0^t \mu(u, T) du + S \end{aligned}$$

avec

$$\begin{aligned} S &= \sum_{i=1}^N \int_0^t \sqrt{v_i(u)} (\alpha_0^i + \alpha_1^i (T - u)) e^{-\gamma_i (T - u)} dW_i^Q(u) \\ &= \sum_{i=1}^N \left[(\alpha_0^i + \alpha_1^i (T - t)) e^{-\gamma_i (T - t)} \int_0^t \sqrt{v_i(u)} e^{-\gamma_i (t - u)} dW_i^Q(u) \right. \\ &\quad \left. + \alpha_1^i e^{-\gamma_i (T - t)} \int_0^t \sqrt{v_i(u)} (t - u) e^{-\gamma_i (t - u)} dW_u^i \right] \end{aligned}$$

En séparant tous les termes $T - u$ en $(T - t) + (t - u)$:

$$\begin{aligned} \int_0^t \mu(u, T) du &= \sum_{i=1}^N \sum_{j=2}^6 I_j^i(t, T) \\ I_2^i(t, T) &= \int_0^t v_i(u) (\alpha_0^i + \alpha_1^i (T - t)) \left(\frac{\alpha_0^i}{\gamma_i} + \frac{\alpha_1^i}{\gamma_i^2} \right) e^{-\gamma_i (T - s)} ds \\ &= (\alpha_0^i + \alpha_1^i (T - t)) \left(\frac{\alpha_0^i}{\gamma_i} + \frac{\alpha_1^i}{\gamma_i^2} \right) e^{-\gamma_i (T - t)} \int_0^t v_i(u) e^{-\gamma_i (t - s)} ds \\ I_3^i(t, T) &= - \int_0^t v_i(u) (\alpha_0^i + \alpha_1^i (T - t)) \left(\frac{\alpha_0^i}{\gamma_i} + \frac{\alpha_1^i}{\gamma_i^2} + \frac{\alpha_1^i}{\gamma_i} (T - t) \right) e^{-2\gamma_i (T - u)} du \\ &= -(\alpha_0^i + \alpha_1^i (T - t)) \left(\frac{\alpha_0^i}{\gamma_i} + \frac{\alpha_1^i}{\gamma_i^2} + \frac{\alpha_1^i}{\gamma_i} (T - t) \right) e^{-2\gamma_i (T - t)} \int_0^t v_i(u) e^{-2\gamma_i (t - u)} du \\ I_4^i(t, T) &= \int_0^t v_i(u) \alpha_1^i (t - u) \left(\frac{\alpha_0^i}{\gamma_i} + \frac{\alpha_1^i}{\gamma_i^2} \right) e^{-\gamma_i (T - u)} du \\ &= \alpha_1^i \left(\frac{\alpha_0^i}{\gamma_i} + \frac{\alpha_1^i}{\gamma_i^2} \right) e^{-\gamma_i (T - t)} \int_0^t v_i(u) (t - u) e^{-\gamma_i (t - u)} du \\ I_5^i(t, T) &= - \int_0^t v_i(u) \left[\alpha_1^i \left(\frac{\alpha_0^i}{\gamma_i} + \frac{\alpha_1^i}{\gamma_i^2} \right) + \frac{\alpha_1^i{}^2}{\gamma_i} (T - t) + (\alpha_0^i + \alpha_1^i (T - t)) \frac{\alpha_1^i}{\gamma_i} \right] (t - u) e^{-2\gamma_i (T - u)} du \\ &= - \left(\frac{\alpha_1^i{}^2}{\gamma_i^2} + 2 \frac{\alpha_0^i \alpha_1^i}{\gamma_i} + 2 \frac{\alpha_1^i{}^2}{\gamma_i} (T - t) \right) e^{-2\gamma_i (T - t)} \int_0^t v_i(u) (t - u) e^{-2\gamma_i (t - u)} du \\ I_6^i(t, T) &= - \int_0^t v_i(u) \frac{\alpha_1^i{}^2}{\gamma_i} (t - u)^2 e^{-2\gamma_i (T - u)} du \end{aligned}$$

$$= -\frac{\alpha_1^i{}^2}{\gamma_i} e^{-2\gamma_i(T-t)} \int_0^t v_i(u)(t-u)^2 e^{-2\gamma_i(t-u)} du$$

On peut donc mettre le taux forward sous la forme :

$$f(t, T) = f(0, T) + \sum_{i=1}^N B_i(T-t)x_i(t) + \sum_{i=1}^N \sum_{j=1}^6 D_i^j(T-t)\varphi_{i,j}(t)$$

avec les fonctions suivantes :

$$\begin{aligned} B_i(\tau) &= (\alpha_0^i + \alpha_1^i \tau) e^{-\gamma_i \tau} \\ D_i^1(\tau) &= \alpha_1^i e^{-\gamma_i \tau} \\ D_i^2(\tau) &= (\alpha_0^i + \alpha_1^i \tau) \left(\frac{\alpha_0^i}{\gamma_i} + \frac{\alpha_1^i}{\gamma_i^2} \right) e^{-\gamma_i \tau} \\ D_i^3(\tau) &= -(\alpha_0^i + \alpha_1^i \tau) \left(\frac{\alpha_0^i}{\gamma_i} + \frac{\alpha_1^i}{\gamma_i^2} + \frac{\alpha_1^i}{\gamma_i} \tau \right) e^{-2\gamma_i \tau} \\ D_i^4(\tau) &= \alpha_1^i \left(\frac{\alpha_0^i}{\gamma_i} + \frac{\alpha_1^i}{\gamma_i^2} \right) e^{-\gamma_i \tau} \\ D_i^5(\tau) &= \left(\frac{\alpha_1^i{}^2}{\gamma_i^2} + 2 \frac{\alpha_0^i \alpha_1^i}{\gamma_i} + 2 \frac{\alpha_1^i{}^2}{\gamma_i} \tau \right) e^{-2\gamma_i \tau} \\ D_i^6(\tau) &= -\frac{\alpha_1^i{}^2}{\gamma_i} e^{-2\gamma_i \tau} \end{aligned}$$

et les x_i , $\varphi_{i,j}$ qui vérifient :

$$\begin{aligned} dx_i(t) &= \sqrt{v_i(t)} dW_t^i - \gamma_i x_i(t) dt \\ d\varphi_{i,1}(t) &= (x_i(t) - \gamma_i \varphi_{i,1}(t)) dt \\ d\varphi_{i,2}(t) &= (v_i(t) - \gamma_i \varphi_{i,2}(t)) dt \\ d\varphi_{i,3}(t) &= (v_i(t) - \gamma_i 2\varphi_{i,3}(t)) dt \\ d\varphi_{i,4}(t) &= \int_0^t \left[v_i(u) e^{-\gamma_i(t-u)} - v_i(u)(t-u) \gamma_i e^{-\gamma_i(t-u)} \right] du dt \\ &= (\varphi_{i,2} - \gamma_i \varphi_{i,4}(t)) dt \\ d\varphi_{i,5}(t) &= (\varphi_{i,3}(t) - 2\gamma_i \varphi_{i,5}(t)) dt \\ d\varphi_{i,6}(t) &= 2(\varphi_{i,5}(t) - \gamma_i \varphi_{i,6}(t)) dt \\ x_i(0) &= \varphi_{i,j}(0) = 0 \end{aligned}$$

2 Zéro-coupons

Il s'agit d'intégrer les fonctions B_i et D_i^j pour aboutir aux \mathcal{B}_i et \mathcal{D}_i^j :

$$\begin{aligned}
\mathcal{B}_i(\tau) &= - \int_0^\tau (\alpha_0^i + \alpha_1^i s) e^{-\gamma_i s} ds \\
&= - \left(\frac{\alpha_0^i}{\gamma_i} + \frac{\alpha_1^i}{\gamma_i^2} \right) (1 - e^{-\gamma_i \tau}) + \frac{\alpha_1^i}{\gamma_i} \tau e^{-\gamma_i \tau} \\
\mathcal{D}_i^1(\tau) &= - \int_0^\tau \alpha_1^i e^{-\gamma_i s} ds \\
&= - \frac{\alpha_1^i}{\gamma_i} (1 - e^{-\gamma_i \tau}) \\
\mathcal{D}_i^2(\tau) &= - \int_0^\tau (\alpha_0^i + \alpha_1^i s) \left(\frac{\alpha_0^i}{\gamma_i} + \frac{\alpha_1^i}{\gamma_i^2} \right) e^{-\gamma_i s} ds \\
&= \left(\frac{\alpha_0^i}{\gamma_i} + \frac{\alpha_1^i}{\gamma_i^2} \right) \left[\frac{\alpha_1^i}{\gamma_i} \tau e^{-\gamma_i \tau} - \left(\frac{\alpha_0^i}{\gamma_i} + \frac{\alpha_1^i}{\gamma_i^2} \right) (1 - e^{-\gamma_i \tau}) \right] \\
\mathcal{D}_i^3(\tau) &= \int_0^\tau (\alpha_0^i + \alpha_1^i s) \left(\frac{\alpha_0^i}{\gamma_i} + \frac{\alpha_1^i}{\gamma_i^2} + \frac{\alpha_1^i}{\gamma_i} s \right) e^{-2\gamma_i s} ds \\
&= \left(\frac{\alpha_0^{i2}}{2\gamma_i^2} + \frac{\alpha_0^i \alpha_1^i}{\gamma_i^3} + \frac{\alpha_1^{i2}}{2\gamma_i^4} \right) (1 - e^{-2\gamma_i \tau}) - \left(\frac{\alpha_0^i \alpha_1^i}{\gamma_i^2} \tau + \frac{\alpha_1^{i2}}{\gamma_i^3} \tau + \frac{\alpha_1^{i2}}{2\gamma_i^2} \tau^2 \right) e^{-2\gamma_i \tau} \\
\mathcal{D}_i^4(\tau) &= - \int_0^\tau \alpha_1^i \left(\frac{\alpha_0^i}{\gamma_i} + \frac{\alpha_1^i}{\gamma_i^2} \right) e^{-\gamma_i s} ds \\
&= - \frac{\alpha_1^i}{\gamma_i} \left(\frac{\alpha_0^i}{\gamma_i} + \frac{\alpha_1^i}{\gamma_i^2} \right) (1 - e^{-\gamma_i \tau}) \\
\mathcal{D}_i^5(\tau) &= \int_0^\tau \left(\frac{\alpha_1^{i2}}{\gamma_i^2} + 2 \frac{\alpha_0^i \alpha_1^i}{\gamma_i} + 2 \frac{\alpha_1^{i2}}{\gamma_i} s \right) e^{-2\gamma_i s} ds \\
&= \left(\frac{\alpha_1^{i2}}{\gamma_i^3} + \frac{\alpha_0^i \alpha_1^i}{\gamma_i^2} \right) (1 - e^{-2\gamma_i \tau}) - \frac{\alpha_1^{i2}}{\gamma_i^2} \tau e^{-2\gamma_i \tau} \\
\mathcal{D}_i^6(\tau) &= \int_0^\tau \frac{\alpha_1^{i2}}{\gamma_i} e^{-2\gamma_i s} ds \\
&= \frac{\alpha_1^{i2}}{2\gamma_i^2} (1 - e^{-2\gamma_i \tau})
\end{aligned}$$

E Données de l'étude

1 Prix des swaptions

Voici les prix des swaptions, présentées sous formes de leurs volatilités implicites normales.

Mois de mars (consulté le 01/04/16)

		Ténor		
		1	5	10
Maturité	1	0.2537%	0.4449%	0.6432%
	2	0.3742%	0.5341%	0.6842%
	3	0.8429%	0.6095%	0.716%
	4	0.5941%	0.6633%	0.7342%
	5	0.6554%	0.7021%	0.7479%
	7	0.7089%	0.7405%	0.7582%
	10	0.7276%	0.7535%	0.7497%
	15	0.7141%	0.727%	0.7138%
	20	0.6845%	0.689%	0.6725%

Mois de juin (consulté le 01/07/16)

		Ténor		
		1	5	10
Maturité	1	0.2496%	0.4085%	0.6088%
	2	0.3526%	0.4979%	0.6541%
	3	0.4758%	0.5751%	0.6943%
	4	0.5808%	0.6397%	0.7189%
	5	0.6283%	0.6824%	0.7336%
	7	0.6998%	0.725%	0.7371%
	10	0.7171%	0.7335%	0.7263%
	15	0.6895%	0.7065%	0.6933%
	20	0.6589%	0.6708%	0.655%

2 Courbes de taux

Courbe des taux EIOPA au 31 mars

Mat.	Fac. Act.	Mat.	Fac. Act.	Mat.	Fac. Act.
0M	1	19A	0.81091717	71A	0.11445725
1M	1.00006318	20A	0.79707286	72A	0.10984631
2M	1.00011571	21A	0.78095075	73A	0.10542083
3M	1.00015912	22A	0.76295153	74A	0.1011734
4M	1.00019494	23A	0.74354979	75A	0.09709688
5M	1.00022468	24A	0.72313805	76A	0.09318445
6M	1.00024983	25A	0.70203993	77A	0.08942951
7M	1.00027185	26A	0.68052137	78A	0.08582576
8M	1.00029219	27A	0.65879995	79A	0.08236712
9M	1.00031229	28A	0.63705282	80A	0.07904778
10M	1.00033355	29A	0.61542329	81A	0.07586212
11M	1.00035738	30A	0.59402645	82A	0.07280479
12M	1.00038514	31A	0.57295383	83A	0.06987062
13M	1.00041778	32A	0.55227733	84A	0.06705465
14M	1.00045449	33A	0.53205255	85A	0.06435214
15M	1.00049408	34A	0.51232154	86A	0.06175852
16M	1.00053538	35A	0.49311515	87A	0.0592694
17M	1.0005772	36A	0.47445497	88A	0.05688058
18M	1.0006184	37A	0.45635496	89A	0.05458803
19M	1.00065786	38A	0.43882286	90A	0.05238786
20M	1.00069445	39A	0.4218613	91A	0.05027635
21M	1.00072708	40A	0.40546877	92A	0.04824993
22M	1.00075466	41A	0.38964043	93A	0.04630519
23M	1.00077613	42A	0.3743688	94A	0.04443882
24M	1.00079044	43A	0.35964425	95A	0.04264766
25M	1.00079662	44A	0.34545555	96A	0.0409287
26M	1.00079403	45A	0.33179021	97A	0.03927902
27M	1.00078212	46A	0.31863479	98A	0.03769582
28M	1.00076033	47A	0.3059752	99A	0.03617644
29M	1.00072813	48A	0.29379689	100A	0.03471829
30M	1.00068497	49A	0.28208502	101A	0.03331891
31M	1.00063034	50A	0.27082467	102A	0.03197594
32M	1.00056371	51A	0.26000088	103A	0.03068709
33M	1.00048458	52A	0.24959878	104A	0.0294502
34M	1.00039245	53A	0.23960367	105A	0.02826315
35M	1.00028681	54A	0.2300011	106A	0.02712395
3A	1.00016719	55A	0.22077687	107A	0.02603067
4A	0.99765879	56A	0.21191711	108A	0.02498146
5A	0.99309294	57A	0.20340829	109A	0.02397453
6A	0.98575933	58A	0.19523723	110A	0.02300819
7A	0.97585858	59A	0.18739115	111A	0.0220808
8A	0.96358614	60A	0.17985762	112A	0.02119079
9A	0.95071575	61A	0.17262465	113A	0.02033665
10A	0.93500638	62A	0.1656806	114A	0.01951694
11A	0.92019843	63A	0.15901425	115A	0.01873027
12A	0.90535017	64A	0.15261474	116A	0.01797531
13A	0.88890037	65A	0.14647161	117A	0.01725078
14A	0.87223182	66A	0.14057478	118A	0.01655545
15A	0.85715168	67A	0.13491453	119A	0.01588815
16A	0.84467218	68A	0.12948149	120A	0.01524774
17A	0.83368356	69A	0.12426665		
18A	0.82280918	70A	0.11926135		

Courbe des taux EIOPA au 30 juin

Mat.	Fac. Act.	Mat.	Fac. Act.	Mat.	Fac. Act.
0M	1	19A	0.83938301	71A	0.11888871
1M	1.00012488	20A	0.82554359	72A	0.11409923
2M	1.00023878	21A	0.80923608	73A	0.1095024
3M	1.00034334	22A	0.79089973	74A	0.10509051
4M	1.0004402	23A	0.77104156	75A	0.10085615
5M	1.00053097	24A	0.75008083	76A	0.09679223
6M	1.00061722	25A	0.72836324	77A	0.09289191
7M	1.00070052	26A	0.70617286	78A	0.08914862
8M	1.00078243	27A	0.68374218	79A	0.08555608
9M	1.00086446	28A	0.66126053	80A	0.08210821
10M	1.00094812	29A	0.63888115	81A	0.07879921
11M	1.0010349	30A	0.61672718	82A	0.07562351
12M	1.00112626	31A	0.59489667	83A	0.07257573
13M	1.0012233	32A	0.57346672	84A	0.06965074
14M	1.00132563	33A	0.55249709	85A	0.06684359
15M	1.00143252	34A	0.53203308	86A	0.06414955
16M	1.00154325	35A	0.51210809	87A	0.06156407
17M	1.0016571	36A	0.49274562	88A	0.05908276
18M	1.00177338	37A	0.47396111	89A	0.05670144
19M	1.00189141	38A	0.4557633	90A	0.05441609
20M	1.00201049	39A	0.43815554	91A	0.05222283
21M	1.00212997	40A	0.42113675	92A	0.05011796
22M	1.00224919	41A	0.4047023	93A	0.04809792
23M	1.0023675	42A	0.38884473	94A	0.04615929
24M	1.00248426	43A	0.37355434	95A	0.04429879
25M	1.00259881	44A	0.35881966	96A	0.04251328
26M	1.00271025	45A	0.34462787	97A	0.04079973
27M	1.00281767	46A	0.33096519	98A	0.03915524
28M	1.00292016	47A	0.31781707	99A	0.03757703
29M	1.00301683	48A	0.3051685	100A	0.03606243
30M	1.00310681	49A	0.29300416	101A	0.03460887
31M	1.00318922	50A	0.28130857	102A	0.0332139
32M	1.00326322	51A	0.27006627	103A	0.03187516
33M	1.00332795	52A	0.25926184	104A	0.03059037
34M	1.00338261	53A	0.24888006	105A	0.02935737
35M	1.00342636	54A	0.23890593	106A	0.02817407
3A	1.0034584	55A	0.22932472	107A	0.02703846
4A	1.00285208	56A	0.22012204	108A	0.02594862
5A	1.00014262	57A	0.21128384	109A	0.02490271
6A	0.99466195	58A	0.20279646	110A	0.02389896
7A	0.98683079	59A	0.19464662	111A	0.02293567
8A	0.97665802	60A	0.18682142	112A	0.0220112
9A	0.96482259	61A	0.1793084	113A	0.02112399
10A	0.95099758	62A	0.17209548	114A	0.02027255
11A	0.93711498	63A	0.165171	115A	0.01945542
12A	0.92311303	64A	0.1585237	116A	0.01867123
13A	0.9086522	65A	0.1521427	117A	0.01791865
14A	0.89459466	66A	0.14601754	118A	0.0171964
15A	0.88195077	67A	0.14013811	119A	0.01650326
16A	0.87120287	68A	0.1344947	120A	0.01583807
17A	0.86129872	69A	0.12907795		
18A	0.85102531	70A	0.12387884		