INTEREST RATE MODEL CALIBRATION AND RISK-MANAGEMENT USING SEMIDEFINITE PROGRAMMING.

A. d'Aspremont

THÈSE
PRÉSENTÉE POUR OBTENIR LE TITRE DE
DOCTEUR DE L'ÉCOLE POLYTECHNIQUE.
SPÉCIALITÉ: MATHÉMATIQUES APPLIQUÉES.

SOUTENUE LE 15 FÉVRIER 2003 DEVANT LE JURY COMPOSÉ DE

NICOLE EL KAROUI	DIRECTRICE DE THÈSE
MARCO AVELLANEDA	RAPPORTEUR
RAMA CONT	EXAMINATEUR
BRUNO DUPIRE	EXAMINATEUR
LAURENT EL GHAOUI	RAPPORTEUR
OLIVIER SCAILLET	RAPPORTEUR

Préface et principaux résultats

Préface

Le modèle de Black & Scholes (1973) permet l'évaluation des options sur un seul actif dans un modèle lognormal et établit ainsi une bijection entre la volatilité d'un actif et le prix d'une option d'achat. Dans la pratique, les propriétés d'homogénéité de cette relation on fait de la volatilité implicite l'instrument privilégié de cotation des prix d'options. La formule de Black & Scholes (1973) n'a malheureusement pas d'équivalent simple dans le cas multivarié. Ainsi, la cotation des options sur portefeuilles s'effectue en termes de volatilité implicte, mais il n'est pas possible de relier simplement cette volatilité implicite de portefeuille à celle des actifs le composant. Le marché des options sur portefeuille d'actions étant majoritairement composé d'options sur indices, les risques d'incohérence dans la modélisation y sont très limités. La situation est cependant exactement inverse dans le marché des options sur taux d'intérêt.

L'activité de synthèse et de couverture des produits dérivés se décompose en trois phases. Dans la première, le modèle est calibré aux conditions présentes du marché. Cette phase implique la résolution d'un problème inverse, trivial dans le cas du modèle simple de Black & Scholes (1973), beaucoup plus difficile dans les modèles de taux. La seconde phase est l'évaluation propremennt dite des produits et de leur couverture. La troisième est la description et la gestion au quotidien du risque induit par les portefeuilles de dérivés portés. L'obtention de solutions numériques performantes et stables tout au long de cette chaîne de problèmes est indispensable au maintien d'une activité viable d'agrégation-désagrégation des risques financiers.

La liquidité du marché des produits dérivés de taux se concentre massivement autour des *caps* et *swaptions*, qui sont des options sur portefeuille. La modélisation jointe des taux et la possibilité de pouvoir évaluer de manière cohérente l'ensemble de ces options y est donc un prérequis fondamental. Les swaptions sont des options d'achat sur une combinaison convexe de taux forward. Les coéfficients de cette combinaison ont une variance très faible et les taux forwards ne sont pas des martingales sous une même probabilité. La première contribution de ce travail est de montrer que le prix des swaptions peut être approximé, sous une mesure martingale bien choisie (voir Jamshidian (1997)), par celui d'une option sur un portefeuille de martingales lognormales, l'erreur pouvant être bornée uniformément. Ceci réduit le problème de l'évaluation des swaptions dans les modèles linéaires, gaussiens, markoviens (voir El Karoui & Lacoste (1992), Duffie & Kan (1996) ou Musiela & Rutkowski (1997)) ou dans le modèle de marché (Brace, Gatarek & Musiela (1997) ou Sandmann & Sondermann (1997)), à celui de l'évaluation d'options d'achat dans un modèle de Black & Scholes (1973) multivarié.

Un deuxième chapitre s'intéresse donc a l'évaluation des options sur portefeuille dans le cadre d'un modèle de Black & Scholes (1973) multivarié. Un développement en série du prix d'une option d'achat est obtenu par des techniques d'approximation de diffusion similaires à celles utilisées dans Fournié, Lebuchoux & Touzi (1997). En plus de l'important gain en précison et rapidité par rapport aux méthodes de Monte-Carlo, les termes obtenus par cette méthode ont une interpretaion très naturelle. Le terme d'ordre zéro correspond a l'approximation classique d'une somme de variables lognormales par une variable lognormale, le terme d'ordre un correspond à un terme correcteur égal à l'espérance de l'erreur de couverture.

Cette approximation du prix des options permet d'écrire le problème de calibration d'un modèle de taux comme celui de trouver une matrice semidéfinie positive qui vérifie une série de contraintes linéaires. En d'autres termes, le problème de cablibration devient un programme semidéfini. Depuis les travaux de Nesterov & Nemirovskii (1994) et Vandenberghe & Boyd (1996) entre autres, ces programmes peuvent être résolus très efficacement, l'analyse et la preuve de la complexité polynomiale de ces problèmes étant similaire à celle obtenue pour les programmes linéaires (voir Nesterov & Todd (1998)).

Dans ce cadre, le dual du programme de calibration a également une interprétation très naturelle en termes de gestion des risques. Le cône des matrices semidéfinies positives étant symétrique, ce programme est également un programme semidéfini et sa solution fournit, selon l'objectif choisi, soit un portefeuille de couverture au sens de El Karoui & Quenez (1991) et Avellaneda & Paras (1996), soit la sensibilité de la solution à un changement des conditions de marché.

L'instabilité numérique a un coût direct pour les opérateurs de marché qui se traduit par une couverture imparfaite, des coûts de transaction et une description incomplète des risques. Par leur capacité à stabiliser ce processus quotidien de calibration, couverture et gestion des risques, nous espérons que les méthodes exposées dans ce travail vont réduire les coûts de transaction et améliorer la fiabilité et la transparence de la gestion des risques liés aux opérations sur produits dérivés exotiques.

Principaux résultats

Motivations, contributions et littérature associée

Les problèmes de calibration et de gestion des risques d'un modèles de taux ont comme paramètre naturels un operateur de covariance. Les méthodes actuelles qui consistent à fortement paramétrer cet opérateur ou a lui substituer des données historiques sont non convexes et donc intrinsèquement instables et inefficaces.

Origine du problème

- Dans le cadre de l'analyse de Heath, Jarrow & Morton (1992), on sait qu'un modèle de taux arbitré est entièrement paramétré par la donnée de la courbe des taux aujourd'hui et de leur fonction de covariance.

- Si on discrétise, le *paramètre naturel* de la calibration d'un modèle de taux est *une matrice semidéfinie positive*.

- Actuellement: la calibration est soit paramétrée par *un ou deux facteurs*, soit basée sur la *corrélation historique*.

- Les programmes de calibrations actuels sont donc non convexes et *intrinsèquement instables*.

- Ces méthodes *n'exploitent pas toute la richesse des modèles* sous-jacents.

- De plus, ces approches ne fournissent pas de résultats fiables sur la *sensibilité* de la solution à une variation des prix de marché. La technique la plus souvent utilisée est de *modifier* les données initiales et de *recalibrer* pour un certain nombre de scénarios précis. Cette technique, coûteuse numériquement, *amplifie l'instabilité* des résultats.

Contributions

La clé de tous les résultats qui vont suivre se trouve dans le développement récent d'alogrithmes de programmation linéaire sur l'espace des matrices semidéfinies, algorithmes dont la complexité (polynomiale) est comparable a celle des programmes linéaires classiques.

- Dans l'évaluation du prix d'une swaption , on peut *assimiler le taux swap a un panier de taux forwards***.**
- Parce que *la volatilité des zéros coupons est faible* comparée à celle des forwards, on peut supposer que les *poids* dans ce panier sont *constants* et que le swap et les forwards sont traités comme des *martingales sous une même mesure* dans l'évaluation des swaptions.
- Le prix de ces options panier peut ensuite être calculé (en première approximation) en utlisant la formule de Black-Scholes (*formule de marché* pour les swaptions) avec une variance bien choisie.
- Les autres termes du développement peuvent être calculés explicitement, le terme d'ordre un s'interprètant comme l'*erreur moyenne de couverture*.
- La **"***variance de marché***"** *est une forme linéaire* sur la matrice de covariance des forwards.
- Si l'on choisit d'optimiser un objective linéaire en variance, le problème de la calibration peut donc se résoudre comme *un programme semidéfini* canonique.
- Le *dual* de ce programme est un programme de couverture et *fourni la sensibilité* de la solution à une variation des prix de marchés.
- Enfin, l'optimisation du Gamma d'un portefeuille au moyen d'options vanille s'écrit également comme un programme semidéfini.

Littérature associée

- Les travaux de Nesterov & Nemirovskii (1994) et Vandenberghe & Boyd (1996) sur la programmation semidéfinie, Nesterov & Todd (1998) pour un traitement général de la complexité des programmes linéaires sur les cônes symétriques.
- Les résultats Rebonato (1998), Brace, Dun & Barton (1999) et Singleton & Umantsev (2001) sur l'évaluation des swaptions comme paniers de forwards. Rebonato (1999) sur la calibration du modèle de Brace et al. (1997) par paramétrisation des facteurs.
- Les travaux parallèles de Brace & Womersley (2000) sur la calibration du modèle de Brace et al. (1997) par programmaion semidéfinie et l'impact du nombre de facteurs sur l'évaluation de la Mid-Atlantique.
- Les articles de Fournié et al. (1997) et Lebuchoux & Musiela (1999) sur les approximations de diffusions.
- L'article de Douady (1995) sur l'optimisation du Gamma.

Première partie: évaluation des swaptions

Les modèles

On défini l'évolution de l'actif sans risque par $\beta_s = \exp\left(\int_t^s r(u,0)du\right)$ (1 euro placé a la date zéro au taux court) où $r(u,0)$ est le taux court spot a la date u. Dans l'analyse de Heath et al. (1992), et si on note $B(s,T)$ le prix en s du zero-coupon de maturité T

$$B(s,T) = E_s^{\mathbf{Q}}\left[\exp\left(-\int_t^t r(u,0)du\right)\right]$$

l'absence d'arbitrage entre les différents Z.C. impose:

$$\frac{B(t,T)}{\beta_t} = B(0,T)\exp\left(-\int_0^t \sigma^B(s,T-s)dW_s - \frac{1}{2}\int_0^t \left|\sigma^B(s,T-s)\right|^2 ds\right)$$

où $\{\sigma^B(t,\theta);\theta \geq 0\}$ est la volatilité des Z.C. et $W = \{W_t, t \geq 0\}$ est un M.B. de dimension d sous une probabilité risque-neutre **Q**. On définit ensuite le taux forward LIBOR de maturité δ (par ex. 3 mois) à la date t par:

$$1 + \delta L(t,\theta) = \exp\left(\int_\theta^{\theta+\delta} r(t,\nu)d\nu\right)$$

Le modèle de marché sur les LIBOR (lognormal en taux) Dans ce modèle, on suppose que le **taux** forward LIBOR a une volatilité lognormale:

$$dL(s,\theta) = (...)ds + L(s,\theta)\gamma(s,\theta)dW_s$$

avec $\gamma : \mathbb{R}_+^2 \longmapsto \mathbb{R}_+^d$ déterministe et si comme dans Brace et al. (1997) on impose $\gamma(s,\theta) = 0$, $\forall$ $\theta \in [0,\delta[$, on a spécifié la volatilité des Z.C. comme:

$$\sigma^B(t,\theta) = \sum_{k=1}^{\lfloor \delta^{-1}\theta \rfloor} \frac{\delta L(t,\theta - k\delta)}{1 + \delta L(t,\theta - k\delta)}\gamma(t,\theta - k\delta)$$

Le modèle affine (lognormal en prix) Dans ce modèle, on suppose que les **prix des Z.C.** ont une dynamique lognormale et sont donnés par:

$$\frac{dB(s,T)}{B(s,T)} = r(s,0)ds + \sigma^B(s,T-s)dW_s$$

où $\sigma^B(s,T-s)$ est ici déterministe (on obtient une dynamique lognormale shiftée sur les forwards).

Instruments de base: les swaptions

Taux swap Le swap est le taux qui équilibre les PV d'une branche fixe et d'une branche variable. Il est défini par:

$$swap(t,T,T_n) = \frac{B(t,T^{floating}) - B(t,T_{n+1}^{floating})}{Level(t,T^{fixed},T_n^{fixed})}$$

avec $Level(t,T^{fixed},T_n^{fixed}) = \sum_{i=i_T}^{n} coverage(T_i^{fixed},T_{i+1}^{fixed})B(t,T_i^{fixed})$ et $T_{i_T} = T$. On peut encore écrire ce taux comme:

$$swap(t,T_0,T_n) = \sum_{i=i_T}^{n} \omega_i(t)K(t,T_i)$$

où

$$\omega_i(t) = \frac{coverage(T_i^{float},T_{i+1}^{float})B(t,T_{i+1}^{float})}{Level(t,T^{fixed},T_n^{fixed})} \text{ et } K(t,T) = L(t,T-t)$$

En pratique, les poids $\omega_i(t)$ sont remarquablement stables. En pratique, on considère la fréquence des payements flottants comme étant un multiple de celle des payements fixes.

Swaption (formule en taux) Si on suppose que les taux suivent la dynamique du modèle de marché sur les taux LIBOR, on définit le prix de la swaption comme une somme de Calls sur le taux swap prévalant a la date T:

$$Ps(t) = B(t,T)E_t^{Q_T}\left[\sum_{i=i_T}^{N} \frac{\beta(T)\delta cvg(i,b)}{\beta(T_{i+1})}\left(swap(T,T,T_N)-k\right)^+\right]$$

où $\mathbf{Q}^T$ est la probabilité forward en T. Si on défini une nouvelle probabilité martingale $\mathbf{Q}^{LVL}$ associée au forward swap:

$$\frac{d\mathbf{Q}^{LVL}}{d\mathbf{Q}^T}|_t = B(t,T)\beta(T)\sum_{i=1}^{N}\frac{\delta cvg(i,b)\beta^{-1}(T_{i+1})}{Level(t,T,T_N)}$$

on peut réécrire le prix de la swaption comme une option sur le taux swap:

$$Ps(t) = Level(t,T,T_N)E_t^{Q_{LVL}}\left[(swap(T,T,T_N)-k)^+\right]$$

ou encore comme une option sur un panier de forwards:

$$Ps(t) = Level(t,T,T_N)E_t^{Q_{LVL}}\left[\left(\sum_{i=0}^{n}\omega_i(T)K(T,T_i)-k\right)^+\right]$$

Dans le modèle de marché sur les LIBOR, on constate que *la stabilité empirique est bien reproduite par le modèle*. En effet, on a:

$$dswap(s,T,T_N) = \sum_{i=i_T}^{N}\omega_i(s)K(s,T_i)\left(\gamma(s,T_i-s)+\eta(s,T_i)\right)dW_s^{LVL}$$

où la contribution des poids est donnée par:

$$\eta(s,T_i) = \left(\sum_{k=i_T}^{N}\omega_i(s)\left(\sigma^B(s,T_i-s)-\sigma^B(s,T_k-s)\right)\right)$$

où $\sigma^B(t,\theta)$ est la volatilité des Z.C. D'autre part le changement de probabilité se traduit en termes de drift par:

$$dW_s^{LVL} = dW_s^T + \sum_{i=i_T}^{N}\left(\omega_i(s)\sum_{j=1}^{i}\frac{\delta K(s,T_j)}{1+\delta K(s,T_j)}\gamma(s,T_j-s)\right)ds$$

En première approximation, avec en pratique:

$$\delta K(s,T_j) \simeq 1\%$$

et

$$\sum_{i=i_T}^{N} \omega_i(s) K(s,T_i) \eta(s,T_i) = \sum_{i=i_T}^{N} \omega_i(s) \left(K(s,T_i) - swap(s,T,T_N)\right) \eta(s,T_i)$$

avec $\sum_{i=i_T}^{N} \omega_i(s) = 1$ et $0 \leq \omega_i(s) \leq 1$, on peut considérer que la contribution des poids dans la volatilité du swap peut être négligée face à celle des fowards. On peut également négliger le drift introduit par le passage de la probabilité forward à la probabilité forward swap. La swaption dans le modèle lognormal en taux peut donc être évaluée comme une *option sur un panier de taux lognormaux*.

Swaption (formule en prix) On peut aussi écrire le prix de la Swaption de strike k et de maturité T comme celui d'un put sur un panier de Z.C.:

$$Ps(t) = B(t,T) E_t^{Q_T} \left[\left(1 - B(t,T_{N+1}) - k\delta \sum_{i=i_T}^{N} B(t,T_i) \right)^+ \right]$$

les coefficients dans le panier sont ici constants. Dans le modèle lognormal en prix, la swaption peut donc ici aussi être évaluée comme une *option sur un panier d'actifs lognormaux*.

Évaluation des options sur un panier d'actifs

La dynamique du forward Dans les deux types de modèles qui précèdent, on écrit le prix de la swaption comme celui d'une option sur un panier d'actifs lognormaux. On va donc chercher à approximer ce prix dans les cas général où ces n actifs on une corrélation de dimension a priori égale à n, en utilisant les méthodes de développement du prix détaillées par Fournié et al. (1997) et Lebuchoux & Musiela (1999). On se place directement dans le marché forward où la dynamique des actifs sous-jacents (prix ou taux) est donnée par:

$$dF_s^i = F_s^i \sigma_s^i dW_s$$

où W_t un $\mathbf{Q}^T$-Brownien d-dimensionel et $\sigma_s = \left(\sigma_s^i\right)_{i=1,\ldots,n} \in \mathbb{R}^{n \times d}$ est la matrice de volatilité. Dans toute la suite on notera $\Gamma_s \in \mathbb{R}^{n \times n}$ la matrice de covariance correspondante. On cherche à calculer le prix d'un Call sur panier dont le payoff à maturité est donné par:

$$h\left(F_T^\omega\right) = \left(\sum_{i=1}^{n} \omega_i F_T^i - k \right)^+ \text{ avec } \sum_{i=1}^{n} \omega_i = 1$$

Pour ce faire on commence par écrire la dynamique du sous-jacent F_s^ω sous forme lognormale:

$$dF_s^\omega = F_s^\omega \left(\sum_{i=1}^{n} \widehat{\omega}_{i,s} \sigma_s^i \right) dW_s$$

avec

$$\widehat{\omega}_{i,s} = \frac{\omega_i F_s^i}{\sum_{i=1}^{n} \omega_i F_s^i}$$

La dynamique de ces poids est donc donnée par

$$\frac{d\widehat{\omega}_{i,s}}{\widehat{\omega}_{i,s}}=\left(\sum_{j=1}^{n}\widehat{\omega}_{j,s}\left(\sigma_s^i-\sigma_s^j\right)\right)\left(dW_s+\sum_{j=1}^{n}\widehat{\omega}_{j,s}\sigma_s^j ds\right)$$

et on vérifie tres naturlellement que si les volatilités σ_s^i sont toutes identiques, la duynamique du forward F_s^ω est exactement lognormale avec comme volatilité σ_s^ω, on définit donc la volatilité résiduelle de chaque actif par rapport à cette volatilité centrale comme:

$$\xi_s^i=\sigma_s^i-\sum_{j=1}^{n}\widehat{\omega}_{j,t}\sigma_s^j \text{ avec } \sigma_s^\omega=\sum_{j=1}^{n}\widehat{\omega}_{i,t}\sigma_s^j$$

où σ_s^ω est $\mathcal{F}_t-mesurable.$

Développement du prix En pratique, la volatilité résiduelle et les moyennes $\sum_{j=1}^n\widehat{\omega}_{j,s}\tilde{\sigma}_s^j$ sont supposées petites et on va donc développer la dynamique du forward en remplaçant $\sum_{j=1}^n\widehat{\omega}_{j,s}^\varepsilon\xi_s^j$ par $\varepsilon\sum_{j=1}^n\widehat{\omega}_{j,s}^\varepsilon\xi_s^j$ pour un $\varepsilon>0$ petit, pour écrire:

$$\begin{cases} dF_s^{\omega,\varepsilon}=F_s^{\omega,\varepsilon}\left(\sigma_s^\omega+\varepsilon\sum_{j=1}^n\widehat{\omega}_{j,s}\xi_s^j\right)dW_s \\ d\widehat{\omega}_{i,s}^\varepsilon=\widehat{\omega}_{i,s}^\varepsilon\left(\xi_s^i-\varepsilon\sum_{j=1}^n\widehat{\omega}_{j,s}^\varepsilon\xi_s^j\right)\left(dW_s+\sigma_s^\omega ds+\varepsilon\sum_{j=1}^n\widehat{\omega}_{j,s}\xi_s^j ds\right)\end{cases}$$

Comme dans Fournié et al. (1997) et Lebuchoux & Musiela (1999) on cherche donc a évaluer:

$$C^\varepsilon=E\left[\left(F_T^{\omega,\varepsilon}-k\right)^+\mid\left(F_t^\omega,\widehat{\omega}_t\right)\right]$$

en l'approximant par sont développement de Taylor autour de $\varepsilon=0$:

$$C^\varepsilon=C^0+C^{(1)}\varepsilon+C^{(2)}\frac{\varepsilon^2}{2}+o(\varepsilon^2)$$

Terme d'ordre zéro Le terme d'ordre zéro se calcule directement comme la solution de l'E.D.P. limite:

$$\begin{cases}\frac{\partial C^0}{\partial s}+\|\sigma_s^\omega\|^2\frac{x^2}{2}\frac{\partial^2 C^0}{\partial x^2}=0 \\ C^0=(x-K)^+ \text{ for } s=T\end{cases}$$

et on peut donc obtenir C^0 par la formule de Black & Scholes (1973) avec comme variance $\|\sigma_s^\omega\|^2$:

$$C^0=BS(T,F_t^\omega,V_T)=F_t^\omega N(h(V_T))-\kappa N\left(h(V_T)-\sqrt{V_T}\right)$$

avec

$$h\left(V_T\right)=\frac{\left(\ln\left(\frac{F_t^\omega}{\kappa}\right)+\frac{1}{2}V_T\right)}{\sqrt{V_T}} \text{ et } V_T=\int_t^T\|\sigma_s^\omega\|^2 ds$$

Terme d'ordre un On peut ensuite s'intéresser à l'E.D.P. vérifiée par $\partial C^\varepsilon/\partial\varepsilon$:

$$\begin{cases} L_0^\varepsilon C^\varepsilon = 0 \\ C^\varepsilon = (x-k)^+ \text{ en } s = T \end{cases}$$

où l'on a noté:

$$\begin{aligned} L_0^\varepsilon &= \frac{\partial C^\varepsilon}{\partial s} + \left\| \sigma_s^\omega + \varepsilon \sum_{j=1}^n y_j \xi_s^j \right\|^2 \frac{x^2}{2} \frac{\partial^2 C^\varepsilon}{\partial x^2} \\ &+ \sum_{j=1}^n \left(\left\langle \xi_s^j, \sigma_s^\omega \right\rangle + \varepsilon \sum_{k=1}^n y_k \left\langle \xi_s^j - \sigma_s^\omega, \xi_s^k \right\rangle - \varepsilon^2 \left\| \sum_{k=1}^n y_k \xi_s^k \right\|^2 \right) x y_j \frac{\partial^2 C^\varepsilon}{\partial x \partial y_j} \\ &+ \sum_{j=1}^n \left\| \xi_s^j - \varepsilon \sum_{k=1}^n y_k \xi_s^k \right\|^2 \frac{y_j^2}{2} \frac{\partial^2 C^\varepsilon}{\partial y_j^2} \\ &+ \sum_{j=1}^n \left(\left\langle \xi_s^j, \sigma_s^\omega \right\rangle + \varepsilon \sum_{k=1}^n y_k \left\langle \xi_s^j - \sigma_s^\omega, \xi_s^k \right\rangle - \varepsilon^2 \left\| \sum_{k=1}^n y_k \xi_s^k \right\|^2 \right) y_j \frac{\partial C^\varepsilon}{\partial y_j} \end{aligned}$$

on peut passer à la limite en $\varepsilon = 0$ (en s'accordant comme dans Fournié et al. (1997) un peu de liberté avec les conditions de régularité), ce qui donne:

$$\begin{cases} L_0^0 C^{(1)} + \left(\sum_{j=1}^n y_j \left\langle \xi_s^j, \sigma_s^\omega \right\rangle \right) x^2 \frac{\partial^2 C^0}{\partial x^2} = 0 \\ C^\varepsilon = 0 \text{ en } s = T \end{cases}$$

Ceci permet de calculer $C^{(1)}$ en utilisant la représentation de Feynman-Kac:

$$\begin{aligned} C^{(1)} &= F_t^\omega \int_t^T \sum_{j=1}^n \widehat{\omega}_{j,t} \left\langle \xi_s^j, \sigma_s^\omega \right\rangle \exp\left(\int_t^s -\frac{1}{2} \left\| \xi_u^j - \sigma_u^\omega \right\|^2 du \right) \\ E &\left[\frac{\exp\left(\int_t^s \left(\sigma_u^\omega + \xi_u^j \right) dW_u \right)}{\sqrt{V_{s,T}}} n \left(\frac{\ln \frac{F_t^\omega}{K} + \int_t^s \sigma_u^\omega dW_u - \frac{1}{2} V_{t,s} + \frac{1}{2} V_{s,T}}{\sqrt{V_{s,T}}} \right) \right] ds \end{aligned}$$

pour obtenir:

$$\begin{aligned} C^{(1)} &= F_t^\omega \int_t^T \sum_{j=1}^n \widehat{\omega}_{j,t} \frac{\left\langle \xi_s^j, \sigma_s^\omega \right\rangle}{\sqrt{V_{t,T}}} \exp\left(2 \int_t^s \left\langle \xi_u^j, \sigma_u^\omega \right\rangle du \right) \\ & \quad n \left(\frac{\ln \frac{F_t^\omega}{K} + \int_t^s \left\langle \xi_u^j, \sigma_u^\omega \right\rangle du + \frac{1}{2} V_{t,T}}{\sqrt{V_{t,T}}} \right) ds \end{aligned}$$

Calcul du prix de l'option sur panier En résumé on peut donc obtenir une formule approximant le prix de l'option sur panier:

$$E_t \left[(F_T^\omega - k)^+ \right] = BS(T, F_t^\omega, V_T) + C^{(1)}$$

où

$$V_T = \int_t^T \left\| \sum_{i=1}^N \hat{\omega}_i(t)\gamma(s, T_i - s) \right\|^2 ds$$

et

$$\begin{aligned} C^{(1)} &= F_t^\omega \int_t^T \sum_{j=1}^n \widehat{\omega}_{j,t} \frac{\left\langle \xi_s^j, \sigma_s^\omega \right\rangle}{\sqrt{V_{t,T}}} \exp\left(2 \int_t^s \left\langle \xi_u^j, \sigma_u^\omega \right\rangle du\right) \\ & \quad n \left(\frac{\ln \frac{F_t^\omega}{K} + \int_t^s \left\langle \xi_u^j, \sigma_u^\omega \right\rangle du + \frac{1}{2} V_{t,T}}{\sqrt{V_{t,T}}} \right) ds \end{aligned}$$

Application aux swaptions Dans le cas de la swaption, la formule a l'ordre zéro s'écrit:

$$Level(t, T, T_N) \left(swap(t, T, T_N) N(h) - \kappa N(h - \sqrt{V_T}) \right)$$

avec

$$h = \frac{\left(\ln \left(\frac{swap(t,T,T_N)}{\kappa} \right) + \frac{1}{2} V_T \right)}{\sqrt{V_T}}$$

et où

$$V_T = \int_t^T \left\| \sum_{i=1}^N \hat{\omega}_i(t)\gamma(s, T_i - s) \right\|^2 ds \text{ avec } \hat{\omega}_i(t) = \omega_i(t) \frac{K(t, T_i)}{swap(t, T, T_N)}$$

Précision de la formule sur les paniers simples

On peut étudier la précision de cette approximation en comparant avec un Monte-Carlo (figure 3.2). Ces valeurs répliquent les paramètres utlisés pour une swaption (5 ans, 5ans). La matrice de covariance est issue de données historiques sur la covariance des FRA.

Précisicion de la formule dans le modèle lognormal sur LIBOR.

On peut aussi tester la qualité de l'approximation à l'ordre zéro dans le cadre du modèle de marché en comaprant encore une fois avec les résultats obtenus par Monte-Carlo (figure 3.1).

Deuxième partie: Calibration

Le programme de calibration

Comme on l'a vu dans la partie précédente, le prix de la swaption peut s'approximer par son prix de Black calculé avec ue variance de marché bien choisie. Avec

$$\hat{\omega}_i(t) = \omega_i(t) \frac{K(t, T_i)}{swap(t, T, T_N)}$$

où les $\omega_i(t)$ proviennent de la décomposition du Swap en panier de FRA, cette variance variance s'obtient comme:

$$V_T = \int_t^T \left\| \sum_{i=1}^N \hat{\omega}_i(t)\gamma(s, T_i - s) \right\|^2 ds$$

ou encore

$$V_T = \int_t^T Tr\left(\Omega_t \Gamma_s\right) ds$$

ou on a noté

$$\Omega_t = \hat{\omega}(t)\hat{\omega}(t)^T = \left(\hat{\omega}_i(t)\hat{\omega}_j(t)\right)_{i,j\in[1,N]} \succeq 0$$

Si on se donne une série de varinaces de marché $\sigma_k^2 T_k$ correspondants à des Swaptions (ou Caplets) de poids $\hat{\omega}_k$ et de maturité T_k et si on suppose que la covariance des LIBOR est constante par morceaux, on peut écrire le programme de calibration comme:

$$\begin{array}{ll} \text{Trouver} & X_i \\ \text{avec} & \sum_{i=t}^{T} \delta Tr\left(\Omega_{t,k} X_i\right) = \sigma_k^2 T_k \text{ où } k = 1, ..., M \\ & X_i \succeq 0 \text{ pour } i = 0, ..., T \end{array}$$

ou encore, sous-forme bloc-diagonale:

$$\begin{array}{ll} \text{Trouver} & X \\ \text{avec} & Tr\left(\Omega_k X\right) = \sigma_k^2 T_k \text{ où } k = 1, ..., M \\ & X \succeq 0 \text{ pour } i = 0, ..., T \end{array}$$

Le programme de calibration s'exprime donc comme un programme semidéfini (SDP) avec comme inconnue la matrice de covariance des FRA.

La résolution simultanée de ce programme et de son dual donne une **preuve de convergence** sous forme du gap de dualité, ou une preuve de non faisabilité si les prix sont incompatibles avec les hypothèses du modèle.

Un programme convexe On peut comparer les deux types de paramétrage du problème de calibration sur un example simple. Dans les programmes paramétrés par facteurs de volatilité, si on cherche a résoudre le programme suivant:

$$\begin{array}{ll} \max & Tr\left(\left[\begin{array}{cc} 1 & -1 \\ -1 & 1 \end{array}\right] X\right) \\ \text{avec} & Tr\left(\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right] X\right) = 1 \\ & X \succeq 0 \end{array}$$

et qu'on le paramètre comme dans Rebonato (1999), on obtient:

$$X(u,v) = \left(\begin{array}{cc} \cos^2(u) & \cos(v)\cos(u)\sin(u) \\ \cos(v)\cos(u)\sin(u) & \sin^2(u) \end{array}\right)$$

on peut représenter la fonction $Tr([1,-1;-1,1]X(u,v))$: En général, le programme paramétré par les facters revient a trouver une solution de rang minimal a un programme semidéfini. Ceci fait apparaitre le programme de calibration comme NP-dur (et même NP-complet).

Par contre, la version SDP s'écrit comme l'optimisation d'une forme linéaire sur l'intersection du cône des matrices semidéfinies positives. Ce cone est représenté ici par

$$X = \left(\begin{array}{cc} x & y \\ y & z \end{array}\right) \succeq 0 \Longleftrightarrow \min_i \lambda_i(X) \geq 0$$

ce qui donne: et le domaine d'un programme semidéfini, l'intersection de ce cône avec un plan peut donc être représentée comme dans la figure 4.3.

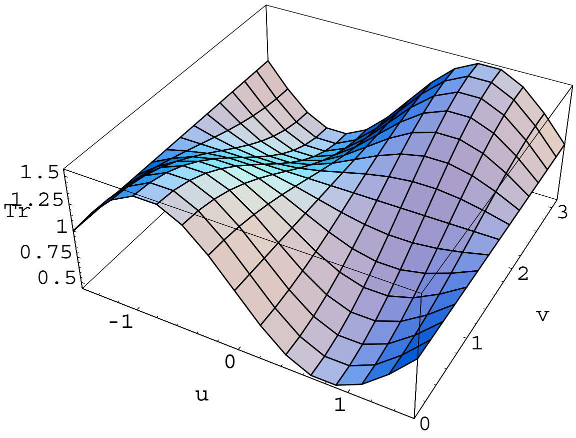

Figure 1: La fonction objectif paramétrée par facteurs.

Objectifs Comme l'ont souligné Nesterov & Nemirovskii (1994), la classe des objectifs représentables par SDP est très vaste. Elle inclut évidemment les objectifs linéaires mais aussi quadratiques par complément de Schur:

$$\|x\|^2 \leq t \text{ pour } x \in \mathbb{R}^n$$

peut encore s'écrire

$$\begin{pmatrix} tI & x^T \\ x & 1 \end{pmatrix} \succeq 0$$

On peut également représenter la norme spectrale en termes d'inégalités matricielles et donc de SDP:

$$\begin{array}{ll} \text{minimiser} & \|X - A\| \\ \text{pour} & Tr(X\Omega_{T_i}) = \left(\sigma^2_{market}\right)_i T_i \\ & X \succeq 0 \end{array}$$

devient:

$$\begin{array}{ll} \text{minimiser} & t \\ \text{pour} & Tr(X\Omega_{T_i}) = \left(\sigma^2_{market}\right)_i T_i \\ & X - A \preceq tId \\ & X - A \succeq -tId \\ & X \succeq 0 \text{ et } t \geq 0 \end{array}$$

Programme dual Le dual du SDP de calibration est un programme avec objectif linéaire, où les contraintes sont données par une inégalité matricielle linéaire. Le cône des matrices semidéfinies positives est autoadjoint, et on peut former le Lagrangien

$$\begin{aligned} L(X, y) &= -Tr(CX) + \sum_{k=1}^{M} y_k \left(Tr(\Omega_k X) - \sigma_k^2 T_k\right) \\ &= Tr\left(\sum_{k=1}^{M} (y_k \Omega_k - C)\, X\right) - \sum_{k=1}^{M} y_k \sigma_k^2 T_k \end{aligned}$$

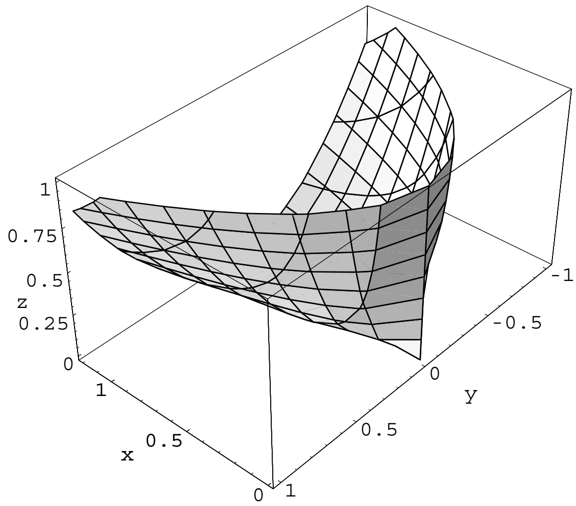

Figure 2: Le cône des matrices semidéfinies positives en dimension trois.

pour obtenir:

$$\begin{array}{ll} \text{maximiser} & \sum_{k=1}^{M} y_k \sigma_k^2 T_k \\ \text{pour} & 0 \preceq \left(\sum_{k=1}^{M} y_k \Omega_k - C\right) \end{array}$$

Calcul des sensibilités On peut maintenant utiliser les résultats de Todd & Yildirim (1999) pour calculer la sensibilité de la solution a un changement dans les conditions du marché. On note X^{opt}, y^{opt} et

$$Z^{opt} = \left(C - \sum_{k=1}^{M} y_k^{opt} \Omega_k\right)$$

la solution du programme de calibration. On note encore pour $P, Q, X \in \mathbb{R}^{n \times n}$:

$$(P \circledast Q) K := \frac{1}{2}\left(PKQ^T + QKP^T\right)$$

et

$$\begin{array}{ll} A : \mathbf{S}^M \longrightarrow \mathbb{R}^m & A^* : \mathbb{R}^m \longrightarrow \mathbf{S}^M \\ X \longmapsto AX := (Tr\,(A_i X))_{i=1,\ldots,m} & y \longmapsto A^* y := \sum_{i=1}^{m} y_i \Omega_i \end{array}$$

On définit $M = I$ (direction de recherche A.H.O.) ou $M = Z^{opt}$ (H.K.M.) et enfin les operateurs $E = Z^{opt} \odot M,\ F = MX^{opt} \odot I$ et leurs adjoints $E^* = Z^{opt} \odot M$ and $F^* = X^{opt} M \odot I$.

On suppose que ces données de marché $\sigma_k^2 T_k$ on été modifiées dans une direction donnée par un vecteur $u \in \mathbb{R}^n$ petit, la nouvelle solution du programme de calibration devient:

$$\Delta X = E^{-1} F A^* \left[\left(A E^{-1} F A^*\right)^{-1} u\right]$$

de plus, celle-ci est garantie valable si

$$\left\| \left(X^{opt}\right)^{-\frac{1}{2}} \left(E^{-1} F A^* \left[\left(A E^{-1} F A^*\right)^{-1} u\right]\right) \left(X^{opt}\right)^{-\frac{1}{2}} \right\| \leq 1$$

On dispose donc d'une matrice donnée par

$$S = E^{-1}FA^* \left[\left(AE^{-1}FA^* \right)^{-1} \right]$$

qui permet de *calculer directement la sensibilité de la solution à l'ensemble des scenarios de marché* possibles.

Bornes sur le prix On peut placer en objectif une matrice correpsondant au prix d'une autre Swaption.

$$\begin{array}{ll} \text{maximiser} & \sigma_{\max}^2 T = Tr(\Omega_0 X) \\ \text{s.t.} & Tr(\Omega_k X) = \sigma_k^2 T_k \text{ for } k = 1, ..., M \\ & X \succeq 0 \end{array}$$

Le dual de ce programme peut s'interpréter à la Avellaneda & Paras (1996) comme un programme de couverture, si on note $BS_k(v)$, le prix de Black Scholes de l'option k pour une variance v:

$$\inf_{\lambda} \left\{ \sum_{k=1}^{M} \lambda_k C_k + \sup_{X \succeq 0} \left(BS_0(Tr(\Omega_0 X)) - \sum_{k=1}^{M} \lambda_k BS_k \left(Tr(\Omega_k X) \right) \right) \right\}$$

ou encore

Prix = Min $\{$Valeur **de la couverture** $+$ Max $($PV du **résidu**$)\}$

Ce prix est donc calculé en introduisant dans la calibration les instruments de couverture et en choisissant les paramètres de calibration les plus conservateurs possibles. L'addition d'instruments dans la calibration améliore la diversification du risque (sous-additivité du max).

La figure 7.9 donne un exemple de bornes sur les prix (6 Novembre 2000). On calibre en utilisant tous les caplets et les swaptions suivantes: 5Y into 5Y, 5Y into 2Y, 5Y into 10Y, 2Y into 2Y, 2Y into 5Y, 7Y into 5Y, 10Y into 5Y, 10Y into 2Y, 10Y into 10Y, 7Y into 3Y, 4Y into 6Y, 17Y into 3Y. Considérant la simplicité du modèle utilisé (covariance stationaire des FRA), il est surprenant de constater que le modèle restitue bien la volatilité des swaptions de sous-jacent inférieur à dix ans.

Rang faible ou matrice régulière? Comme l'ont observé Fazel, Hindi & Boyd (2000), si on place une matrice définie positive comme objectif on obtient en général une matrice de rang faible (figure 4.4) dont les valeurs propres sont rapidement décroissantes (figure 4.5). On constate que la matrice est de rang deux. Cette méthode empirique donne d'excellents résultats en pratique mais aucune garantie ne peut être obtenue quant au rang de la solution (le problème devient alors NP-complet).

Si on impose en plus des contraintes de lissage à la matrice de covariance, on obtient un résultat plus intuitif (figure 4.6) mais cela se fait au prix d'une augmentation du rang de la solution (figure 4.7). Cependant, comme la minimisation de la surface de la matrice de covariance revient a minimiser une entropie (quadratique), on s'attend à ce que cette matrice varie moins au cours du temps que celle obtenue par diminution du rang (on constate tout de même que cette matrice a uniquement deux valeurs propres dominantes, conformément aux résultats empiriques).

Remerciements

Mes remerciements vont tout d'abord à Nicole El Karoui qui a guidé mes premiers pas en finance et n'a cessé d'accompagner mes progrès depuis. Ma dette envers elle est immense. Son enseignement, ses conseils et sa disponibilité, d'abord à l'Ecole Polytechnique, puis au DEA de Paris VI et enfin pendant cette thèse ont été le moteur de ce travail, effectué en liaison étroite depuis différents points du globe (Paris, Londres, Stanford ...).

Mon passage chez Paribas a été un champ d'expérimentation extraordinairement fécond. En particulier, les longues discussions sur le métier de quant avec Jérôme Lebuchoux, ses conseils, ainsi que sa vision du fonctionnement des marchés dérivés ont été une référence fondamentale dans la construction de ce travail. Au-delà de sa fonction à la tête de l'équipe, c'est comme ami que je tiens à lui exprimer ma gratitude. Je souhaite que d'une manière ou d' une autre l'obstacle de la distance soit un jour levé, afin que nous puissions poursuivre ensemble les grands projets entamés à Londres.

Marco Avallaneda a bien voulu accepter de lire ma thèse. Je lui en suis extrêmement reconnaissant, d'autant plus que ses travaux fondamentaux sur la volatilité stochastique et sur la couverture, ainsi que ses résultats récents sur le smile des options paniers en equity, ont été une très grande inspiration pour moi.

Laurent El Ghaoui, avec qui j'ai le grand plaisir de travailler en ce moment sur l'arbitrage statique des options sur paniers, a aimablement accepté d'être rapporteur de cette thèse. Son travail sur la robustesse trouve un écho direct dans les problèmes de calibration et de gestion de portefeuille. Je le remercie vivement pour les très nombreuses indications qu'il m'a données à ce sujet.

Je veux aussi remercier Olivier Scaillet que j'ai rencontré à la conférence de l'AFFI à Strasbourg et qui a également bien voulu être rapporteur de cette thèse. J'espère que nous aurons l'occasion de poursuivre la longue discussion que nous avons entamée à cette occasion.

Rama Cont m'a invité à participer à un très grand nombre de séminaires et conférences qui ont enrichi ma réflexion sous plusieurs aspects. Il a accepté d'être membre du jury. Son aide et ses encouragements m'ont été extrêmement précieux. Les résultats sur la stabilisation de la calibration lui doivent beaucoup.

Les applications numériques des résultats qui vont suivre seraient restées lettre morte sans l'assistance quasi-permanente (d'abord locale, puis trans-Atlantique) de Cyril Godart. Avec Rémy Croisille, il a fait de ce passage à Londres un moment inoubliable.

Jean-Michel Lasry m'a fait l'honneur de m'accueillir dans son équipe. Ses suggestions au cours des entretiens qu'il m'a accordés m'ont été infiniment précieuses.

Marek Musiela est le premier à m'avoir dit que résoudre un problème avec contrainte linéaire sur une matrice semidéfinie positive s'appelait un semidefinite program et a ainsi mis fin à mes expérimentations artisanales sur cette question. Je le remercie ici également pour l'ensemble des idées qu'il a bien voulu partager. Certaines d'entre elles sont toujours en cours d'exploration. Aussi j'espère avoir la chance de pouvoir en discuter avec lui dans l'avenir.

Darrell Duffie m'a accueilli en stage de fin d'Ecole à Stanford. Je tiens ici à le remercier très vivement pour tout ce que j'ai appris durant ce stage ainsi que durant ses cours que j'ai eu le privilège de suivre à Dauphine et à Stanford. Je tiens enfin à lui exprimer ma gratitude pour son aide et ses très nombreux conseils sur la suite de mes études.

Je veux également remercier toute les équipes FIRST Swaps et SPG chez Paribas et en particulier Guillaume Amblard, Régis Bénichou, Philippe Challande, Stefano Galluccio, Ezra Nahum et Eric Fournié. La gentillesse de ce dernier et son expertise m'ont énormément aidé. Je lui dois d'avoir été recruté chez Paribas lors de ses cours à Paris VI. Je regrette le temps où, à Paribas, il me suffisait de me tourner vers la droite pour avoir une réponse à mes questions sur le calcul de Maillavin ou les E.D.P.

Ma reconnaissance va aussi à Jeanne Bailleul, au CMAP, qui m'a permis de surmonter bien des difficultés. Le fait que le présent travail ait abouti montre que ses efforts n'ont pas été vains.

Je remercie également pour leurs commentaires et suggestions toutes les personnes que j'ai rencontrées aux séminaires de Paribas et les participants aux séminaires et conférences où ce travail a été présenté, en particulier Vlad Bally, Jérôme Busca, Claude Lemaréchal, Pierre-Louis Lions, Yurii Nesterov, Pablo Parillo, Antoon Pelsser, Ioanna Popescu, Robert Womersley.

”I don’t want to achieve immortality through my work. I want to achieve it through not dying.”

Woody Allen

Preface

In the original Black & Scholes (1973) model, there is a one-to-one correspondence between the price of an option and the volatility of the underlying asset. In fact, options are most often directly quoted in terms of their Black & Scholes (1973) implied volatility. In the case of options on multiple assets such as basket options, that one-to-one correspondence between market prices and covariance is lost. The market quotes basket options in terms of their Black & Scholes (1973) volatility but has no direct way of describing the link between this volatility and that of the individual assets composing the basket. Today, this is not yet critically important in equity markets where most of the trading in basket options is concentrated among a few index options, we will see however that it is crucial in interest rate derivative markets where most of the volatility information is contained in a rather diverse set of basket options.

Indeed, a large part of the liquidity in interest rate option markets is concentrated in European caps and swaptions. In the first chapter of this work we will show how one can express the price of swaptions (and caplets) as that of an option on a basket of zero-coupon bonds in one approach, or a basket of forward Libor rates in another. This basket option representation is exact in the first case and we will show how it provides an excellent pricing approximation in the second.

In particular, this will allow us to reduce the problem of pricing swaptions in both the Gaussian H.J.M. model (see El Karoui & Lacoste (1992), Duffie & Kan (1996) or Musiela & Rutkowski (1997)) and the Libor market model (see Brace et al. (1997), Miltersen, Sandmann & Sondermann (1995) or Miltersen, Sandmann & Sondermann (1997)) to that of pricing swaptions in a multidimensional Black & Scholes (1973) lognormal model. The second chapter is then focused on finding a good pricing approximation for basket calls in this generic model. We derive price expansion where the first term is computed as the usual Black & Scholes (1973) price with an appropriate variance and the second term can be interpreted as the expected value of the tracking error obtained when hedging with the approximate volatility.

Besides its radical numerical performance compared to Monte-Carlo methods, the formula we obtain has the advantage of expressing the price of a basket option in terms of a Black & Scholes (1973) covariance that is a linear form in the underlying covariance matrix. This sets the multidimensional model calibration problem as that of finding a positive semidefinite (covariance) matrix that satisfies a certain number of linear constraints. In other words, the calibration becomes a *semidefinite program*. Recent advances in optimization (see Nesterov & Nemirovskii (1994) or Vandenberghe & Boyd (1996)) have led to algorithms that solve these problems very efficiently, with a complexity analysis that is comparable to that of linear programs (see Nesterov & Todd (1998)). This means that the general multidimensional market covariance calibration problem can be solved very efficiently.

Finally, we show that these same semidefinite programming techniques provide key sensitivity and risk-management results together with the calibration solution. For instance, we show how all sensitivities of the solution matrix to changes in market conditions can be directly obtained from the optimal solution of the dual calibration problem.

Numerical instability has a direct cost in both poor hedging and incomplete risk description. By reducing the amount of numerical noise in the daily recalibration process and improving the risk-management of interest rate models, we hope these methods will significantly reduce hedging costs and improve the reliability of risk-management computations.

Contents

List of Figures

Part I

Calibration

Introduction

A robust and efficient calibration algorithm is a central element in the successful implementation of a derivatives pricing model. On one hand, the arbitrage-free price derived from a dynamic hedging strategy à la Black & Scholes (1973) and Merton (1973) has now become a central reference in the pricing and risk-management of financial derivatives, on the other hand however, every market operator knows that the data they calibrate on is not arbitrage free because of market imperfections. Beyond these discrepancies in the data, daily model recalibration and the non-convexity of most current calibration methods only add further instability to the derivative pricing, hedging and risk-management process by exposing these computations to purely numerical noise. One of the crucial filters standing between those two sets of prices (market data and computed derivative prices) is the model calibration algorithm.

Recent developments in interest rates modelling have led to a form of technological asymmetry on this topic. The theoretical performance of models such as the Libor Market Model of Interest Rates by Brace et al. (1997) or the affine models (see El Karoui & Lacoste (1992) or Duffie & Kan (1996)) allows a very flexible modelling and pricing of the basic interest rate options (caps and swaptions) at-the-money. However, due to the inefficiency and instability of the calibration procedure, only a small part of the market covariance information that could theoretically be accounted for in the model is actually exploited. To be precise, the most common techniques (see for example Longstaff, Santa-Clara & Schwartz (2000)) perform a completely implicit fit on the caplet variances while only a partial fit is made on the correlation information available in swaptions: the limitations of these methods make it necessary to substitute a statistical estimate to the market information on the forward Libors correlation matrix because the numerical complexity and instability of the calibration process makes it impossible to calibrate a full market covariance matrix. As a direct consequence, these calibration algorithms fail in one of their primary mission: they are very poor market risk visualization tools. The forward rates covariance matrix plays an increasingly important role in exotic interest rate derivatives modelling and there is a need for a calibration algorithm that allows the retrieval of a maximum amount of covariance information in the market. This in turn will improve the stability and robustness of the corresponding hedging strategies by reducing the need for purely numerical hedge portfolio rebalancing. Our objective here is to provide a set of approximations that leads to a simple, meaningful swaptions pricing formula under the Gaussian affine model or the Libor market model and to exploit it to build a fast and robust calibration algorithm.

In the Libor market model, we write swaps as baskets of forwards. As already observed by Rebonato (1998) among others, the weights in this decomposition are empirically very stable. In this part, we first show that this key empirical fact is indeed *accurately reproduced by the model*. We then show that the drift term coming from the change of measure between the forward and the swap martingale measures can be neglected in the computation of the swaption price, thus allowing the forward rate basket option to be priced using the lognormal approximations first detailed in Huynh (1994) and Musiela & Rutkowski (1997).

As we also know from El Karoui & Lacoste (1992) that in the affine Gaussian model, swaptions can be priced as Bond Puts, we notice that in the two models considered here, caplets and swaptions are options on a basket of lognormal assets. In the second chapter, we compute a pricing approximation for these options. We use diffusion approximation techniques similar to those in Fournié et al. (1997) to justify and improve the basket option pricing approximation in Huynh (1994) by accounting for the stochastic nature of the basket volatility. Using results from El Karoui, Jeanblanc-Picqué & Shreve (1998), we get a very intuitive interpretation of the first order price correction as the expected value of the tracking error obtained when hedging with the approximate volatility computed above. Besides its simplicity, this swaption pricing approximation technique has the decisive advantage of expressing the resulting equivalent market variance as a linear function of the forward rates covariance matrix. This means that the calibration problem can be reduced to that of finding a positive semidefinite matrix subject to a set of linear market constraints. Since the pathbreaking work of Nesterov & Nemirovskii (1994) and Vandenberghe & Boyd (1996), techniques derived from interior point methods in linear programming solve this problem very efficiently.

The basket option representation was used in El Karoui & Lacoste (1992) where swaptions were written as Bond Put options in the Linear Gauss Markov affine model. Rebonato (1998) and Rebonato (1999) detail their decomposition as baskets of forwards in the Libor market model. In parallel results, Brace & Womersley (2000) used the order zero lognormal approximation and semidefinite programming to study the impact of the model dimension on Bermudan swaptions pricing. They rely on simulation results dating back to Huynh (1994), Musiela & Rutkowski (1997) or more recently Brace et al. (1999) in an equity framework to justify the lognormal volatility approximation of the swap process. A big step in the same direction had also been made by Rebonato (1999) where the calibration problem was reparametrized on a hypersphere. However, because it did not recognize the convexity of the problem, this last method could not solve the key numerical issue. In recent works, Singleton & Umantsev (2001) studied the effect of zero-coupon dynamics degeneracy on swaption pricing in an affine term structure model while Ju (2002) use a Taylor expansion of the characteristic function to derive basket and Asian option approximations.

This part is organized around three key contributions:

- In chapter one, we detail the basket decomposition of swaps and recall some important results on the market model of interest rates. We show that the weight's volatility and the contribution of the forward vs. swap martingale measure change can be neglected when pricing swaptions in that model.
- In chapter two, we justify the classical lognormal basket option pricing approximation and compute additional terms in the price expansion. We also study the implications in terms of hedging and the method's precision in practice.
- In a third chapter, we explicit the general calibration problem formulation and discuss its numerical performance versus the classical methods. We specifically focus on the rank issue and its implications in derivatives pricing. We show how the calibration result can be stabilized in the spirit of Cont (2001) to reduce hedging transaction costs.

The results detailed here should greatly improve the amount of covariance information that can retrieved from market prices. On the other hand, their simple geometric nature should help visualize the structure itself of that information. Finally, the possibility of stabilizing the solution to the calibration problem given a specific convex objective will vastly improve the day-to-day stability of the calibration procedure, avoiding some of the infamous P&L swings that were only the result of numerical instability and non-convexity.

Chapter 2

Interest Rate Market dynamics

2.1 Zero-coupon bonds and the H.J.M. framework

2.1.1 Zero-coupon bonds

In a first key difference with the standard derivative framework à la Black & Scholes (1973), all interest rate models have as their fundamental underlying an infinite dimensional variable describing the structure of the interest rate curve at any given date. Historically, because all the interest rate trading activity (hence the information) was first concentrated on bonds, the fundamental underlying of choice was the set of discount factors for various maturities. We note these discount factors $B(t,T)$ and they represent the *price* in t (today) of one euro paid at time T.

$$B(t,T) = \text{ price in } t \text{ of 1 euro paid at time } T$$

At any date t, the price of these discount factors is, in general, not directly quoted by the market but can be inferred from that of the various coupon bonds. In that spirit, the discount factors $B(t,T)$ are often called zero-coupon bonds (Z.C.).

2.1.2 Arbitrage free dynamics

Suppose now that in addition to the various available discount factors, we can invest in a savings account that is continuously compounded with the short rate r_s. We note β_T the value at time T of one euro invested in this account at t. We have

$$\beta_T = \exp\left(\int_t^T r_s ds\right)$$

To preclude arbitrage between this savings account and an investment in the Z.C. we have to impose that at time t, an investment of $B(t,T)$ in a Z.C. of maturity T and an investment of this same amount $B(t,T)$ in the continuous savings account produce the same payoff of one euro. As the short rate is stochastic, this absence of arbitrage condition between two portfolios can be written in its dual form:

$$B(t,T) = E_t^{\mathbf{Q}}\left[\exp\left(-\int_t^T r_s ds\right)\right] \tag{2.1}$$

where $\mathbf{Q}$ is some risk-neutral probability measure (see Harrison & Kreps (1979) and Harrison & Pliska (1981)).

We can associate to those discount factor *prices* a continuous zero coupon *rate* that we note $R(t,T)$, such that:

$$B(t,T) = \exp\left(-(T-t)R(t,T)\right)$$

To complete this instrument set, we need to describe the prices of all the forward contracts, i.e. the price at time t of contracts that begin at a certain date $T_1 > t$ in the future. In particular, we can describe the forward zero coupon contracts $B_t(T_1,T_2)$ as the value in t of the amount that has to be paid in T_1 to be guaranteed one euro in T_2. By absence of arbitrage, this can be computed as:

$$B_t(T_1,T_2) = \frac{B(t,T_2)}{B(t,T_1)}$$

To these forward prices, we can associate a forward zero coupon rate $F_t(T_1,T_2)$ such that:

$$B_t(T_1,T_2) = \exp\left(-(T-t)F_t(T_1,T_2)\right)$$

or again

$$F_t(T_1,T_2) = -\frac{(\ln B(t,T_2) - \ln B(t,T_1))}{T_2 - T_1}$$

To describe the dynamics of the interest rate curve it is sometimes useful to introduce the instantaneous version of the forward rates $F_t(T_1,T_2)$ obtained by letting $T_2 - T_1$ go to zero:

$$\begin{aligned} f(t,T) &= \lim_{u\to 0} -\frac{1}{u}\left(\ln B(t,T+u) - \ln B(t,T)\right) \\ &= \frac{\partial \ln B(t,T+u)}{\partial u}|_{u=0} \end{aligned} \tag{2.2}$$

The other rates and prices can be computed from these elementary rates using:

$$B(t,T) = \exp\left(-\int_t^T f(t,s)ds\right) \tag{2.3}$$

In all the quantities described above, the products have some fixed maturity T in the future. This means for example that a zero coupon $B(t,T)$ with maturity five years today will have two years from now a residual maturity of three years. If we want to study the empirical dynamics of the interest rate curve and try to maintain as much stationarity as possible, we need to use variables that remain consistent over time. For this reason, the model dynamics are most often described using rates of constant time to maturity $r(t,\theta)$ with

$$r(t,\theta) = f(t,t+\theta) \text{ for } \theta \geq 0 \tag{2.4}$$

Hence, in what follows, we will use the Musiela parametrization of the Heath et al. (1992) setup and the fundamental rate $r(t,\theta)$ will be the continuously compounded instantaneous forward rate at time t, with duration θ. The dynamics of $r(t,\theta)$ are sometimes called *sliding* maturity (constant time to maturity) while the dynamics of $f(t,T)$ are called *converging*. We will see that the converging dynamics will represent the natural underlying of options and are quoted by the market as forward Rate Agreements, while the sliding rates are supposed to be somewhat more stationary and are used to perform statistical analysis.

Notation 1 *To avoid any confusion, Roman letters will be used for maturity dates (in converging dynamics) and Greek ones for time to maturity (in sliding dynamics).*

We suppose that the zero coupon follow a diffusion process driven by a d dimensional Brownian motion $W = \{W_t, t \geq 0\}$ and because of the arbitrage argument developed in (2.1) above, we know that the drift term of this diffusion must be equal to r_s and we can write the zero coupon dynamics as:

$$\frac{dB(s,T)}{B(s,T)} = r_s ds + \sigma^B(s, T-s)dW_s \tag{2.5}$$

where for all $\theta \geq 0$ the zero-coupon bond volatility process $\{\sigma^B(t,\theta); \theta \geq 0\}$ is $\mathcal{F}_t$-adapted with values in $\mathbb{R}^d$. We assume that the function $\theta \longmapsto \sigma^B(t,\theta)$ is absolutely continuous and the derivative $\tau(t,\theta) = \partial/\partial\theta(\sigma^B(t,\theta))$ is bounded on $\mathbb{R}^2 \times \Omega$. All these processes are defined on the probability space $(\Omega, \{\mathcal{F}_t; t \geq 0\}, \mathbf{Q})$ where the filtration $\{\mathcal{F}_t; t \geq 0\}$ is the $\mathbf{Q}$-augmentation of the natural filtration generated by the d dimensional Brownian motion $W = \{W_t, t \geq 0\}$.

Remark 2 *This construction is the essence of the Heath et al. (1992) framework: because of the no arbitrage conditions between the Z.C. and the savings account, the yield curve dynamics are entirely parametrized by the curve today $B(t,T)$ and the zero coupon volatility function $\sigma^B(s, T-s)$.*

The zero-coupon at time t with maturity T can also be written:

$$B(t,T) = \exp\left(-\int_0^{T-t} r(t,\theta)d\theta\right)$$

And the adapted process $\{r(t,\theta); t, \theta \geq 0\}$ satisfies:

$$dr(s,\theta) = \frac{\partial}{\partial\theta}\left[\left(r(s,\theta) + \frac{1}{2}\left|\sigma^B(s,\theta)\right|^2\right)dt + \sigma^B(s,\theta)dW_s\right] \tag{2.6}$$

The short rate process $r(t,0)$ satisfies:

$$dr_s = \left(\frac{\partial}{\partial\theta}r(s,\theta)\right)_{\theta=0} ds + \left(\frac{\partial}{\partial\theta}\sigma^B(s,\theta)\right)_{\theta=0} dW_s \tag{2.7}$$

and is not Markov in general. The absence of arbitrage condition between all zero-coupons and the savings account then amounts to impose to the process:

$$\frac{B(t,T)}{\beta_t} = B(0,T)\exp\left(-\int_0^t \sigma^B(s,T-s)dW_s - \frac{1}{2}\int_0^t \left|\sigma^B(s,T-s)\right|^2 ds\right) \tag{2.8}$$

to be a martingale under some measure $\mathbf{Q}$ for all $T > 0$.

2.1.3 The Gaussian H.J.M. model

One of the simplest assumptions that can be made to further specify the dynamics above is to impose that the volatility $\sigma^B(t,\theta)$ be deterministic. Because the pricing and calibration of these models is relatively well understood, we only present a very succinct description of these dynamics here and refer the reader to El Karoui & Lacoste (1992), Musiela & Rutkowski (1997) and the work on affine models detailed in Duffie & Kan (1996) for a more complete analysis. In this case the forward rates are Gaussian and the zero coupon follow lognormal dynamics:

$$\frac{dB(s,T)}{B(s,T)} = r_s ds - \sigma^B(s, T-s)dW_s \tag{2.9}$$

We also suppose that the deterministic volatility function $\sigma^B(x,y)$ is twice differentiable with respect to the second variable. We can get the dynamics of the short rate from (2.7):

$$r_t = r(0,t) + \int_0^t \sigma^B(s,T-s)\partial_2\sigma^B(s,T-s)ds + \int_0^t \partial_2\sigma^B(s,T-s)dW_s$$

where we have note ∂_2 the derivative with respect to the second variable. This specification of the volatility is not sufficient to guarantee that the short rate will be Markovian, which can lead to computational difficulties in the simulation and pricing. One way to solve this problem (see for example Musiela & Rutkowski (1997)) is to assume that the volatility is separable and exponentially decreasing with the time to maturity, setting for example:

$$\partial_2\sigma^B(s,T-s) = \sum_{i=1}^{d} \sigma_i(s)\exp\left(-\lambda_i(u-s)\right)$$

with $\sigma_i(s)$ in $\mathbb{R}^d$. This can be seen as a multidimensional generalization of the original Vasicek (1977) model. One of the key characteristics of this class of models (see Duffie & Kan (1996)) is that the instantaneous forward rates $r(t,\theta)$ can all be obtained as an affine function of the short rate r_t:

$$r(t,\theta) = \alpha(t,\theta)r_t + \beta(t,\theta)$$

Remark 3 *In the dynamics above, we have included the correlation directly in the volatility definition by letting $\sigma_i(s)$ be vector valued instead of assuming that the Brownian motions are correlated. This is somewhat different from the usual specification in Musiela & Rutkowski (1997) for example and it supposes that the mean-reverting factors driving the curve movements are not instantaneous forwards but linear combinations of them. The reasons for this unorthodox setup will be made clear when we discuss the calibration method.*

2.2 Libor rates, swap rates and the Libor market model

2.2.1 Libor rates and swaps

Although initially driven by bonds, interest rate markets have seen a significant shift in activity and market depth in the 80's with the introduction of swaps. This shift has seen swap and Libor rates defined below replace the traditional zero coupon rates as reference instruments.

Libor rate

To define these fundamental derivatives we start by remarking that although the market quotes zero coupon prices via bond prices, there is no direct trading in the instantaneous forwards defined above. Instead, the market quotes linearly compounded short rates for maturities starting at three months in general. These rates are often called Libor rates from London Interbank Offered Rates and we note them $L_\delta(t,0)$ with:

$$\frac{1}{1+\delta L_\delta(t,0)} = B(t,t+\delta)$$

where δ less than one year (3 or 6 months in practice). We note $L_\delta(t,\theta)$ the forward Libor rate associated with the forward zero coupon $B_t(t+\theta,t+\theta+\delta)$. Because of the absence of arbitrage, this is given by:

$$\frac{1}{1+\delta L_\delta(t,\theta)} = \frac{B(t,t+\delta+\theta)}{B(t,t+\theta)}$$

or again:

$$B(t,t+\delta+\theta)\delta L_\delta(t,\theta) = B(t,t+\theta) - B(t,t+\delta+\theta)$$

Notation 4 *To clearly distinguish between the sliding and converging dynamics of the Libor rates, we will note* $K(t,T) = L(t,T-t)$ *the converging Libor.*

Unless clearly specified otherwise, we will pick a certain underlying maturity δ (for ex. 3 months) for the forward and we will write $L_\delta(t,\theta)$ instead of $L(t,\theta)$.

Swap rate

A swap rate is then defined as the fixed rate that zeroes the present value of a set of periodical exchanges of fixed against floating coupons on a Libor rate of given maturity at future dates T_i^{fixed} and $T_i^{floating}$. This means that the swap rate $swap(t,T,T_n)$ (which will also be called swap in what follows) is computed as:

$$swap(t,T,T_n)PV^{fixed} = PV^{floating}$$

which is, with i_T such that $T_{i_T} = T$:

$$\begin{aligned} & swap(t,T,T_n)\left(\sum_{i=i_T}^{n} coverage(T_i^{fixed},T_{i+1}^{fixed})B(t,T_{i+1}^{fixed})\right) \\ = \quad & PV\left(\sum_{i=i_T}^{n} coverage(T_i^{fixed},T_{i+1}^{fixed})L_\delta(T_i^{floating},0)\right) \end{aligned}$$

or again, because we know the PV of the floating leg is equal to $B(t,T^{floating}) - B(t,T_{n+1}^{floating})$:

$$swap(t,T,T_n) = \frac{B(t,T^{floating}) - B(t,T_{n+1}^{floating})}{Level(t,T^{fixed},T_n^{fixed})}$$

where, with $coverage(T_i,T_{i+1})$ the coverage (time interval) between T_{i-1} and T_i computed with the appropriate basis (different for the floating and fixed legs), $B(t,T_i^{float})$ the discount factor with maturity T_i^{float} we have defined $Level(t,T^{fixed},T_n^{fixed})$ as the average of the discount factors for the fixed calendar of the swap weighted by their associated coverage:

$$Level(t,T^{fixed},T_n^{fixed}) = \sum_{i=i_T}^{n} coverage(T_i^{fixed},T_{i+1}^{fixed})B(t,T_i^{fixed})$$

which is the sum value of the fixed coupons paid (or fixed leg of the swap). Here $(T_i^{float},...)$ is the calendar for the floating leg of the swap and $(T_i^{fixed},...)$ is the calendar for the fixed leg (this notation is there to highlight the fact that they don't match in general). In a representation that will be critically important in the pricing approximations that follow, we remark that we can write the swaps as baskets of forward Libors (see for ex. Rebonato (1998)).

Lemma 5 *We can write the swap as a basket of forwards:*

$$swap(t,T,T_n) = \sum_{i=i_T}^{n} \omega_i(t)K(t,T_i^{float}) \tag{2.10}$$

with the weights given by:

$$\omega_i(t) = \frac{coverage(T_i^{float}, T_{i+1}^{float}) B(t, T_{i+1}^{float})}{Level(t, T^{fixed}, T_n^{fixed})} \tag{2.11}$$

with $0 \leq \omega_i(t) \leq 1$.

Proof. This is because we can write

$$swap(t, T, T_n) = \frac{\sum_{i=i_T}^{n} B(t, T_i^{floating}) - B(t, T_{i+1}^{floating})}{Level(t, T^{fixed}, T_n^{fixed})}$$

or again

$$swap(t, T, T_n) = \frac{\sum_{i=i_T}^{n} coverage(T_i^{float}, T_{i+1}^{float}) B(t, T_{i+1}^{float}) K(t, T_i^{float})}{Level(t, T^{fixed}, T_n^{fixed})}$$

which is the desired representation. As the corresponding forward Libor rates are positive, we have $B(t, T_{i+1}) \leq B(t, T_i) \leq B(t, T_{i-1})$ for $i \in [i_T + 1, N - 1]$ hence

$$0 \leq \frac{cvg(i, b) B(t, T_i)}{level(t, T, T_N)} \leq 1$$

which means $0 \leq \omega_i(t) \leq 1$, i.e. the weights are positive and bounded by one. ■

In practice, the weights $\omega_i(t)$ prove to have very little variance compared to their respective FRA. Besides Rebonato (1998), p.18, this has been studied by Hamy (1999) of which we report here, with the author's permission, a sample of summary statistics. Here is for example a study of the ratio of the $vol(FRA)/vol(weights)$ ratio in various markets, computed using the standard quadratic variation estimator with exponentially decaying weights (market data courtesy of BNP-Paribas London):

Currency	USD	USD	GBP	GBP	EUR	EUR
swap	$2Y$	$5Y$	$2Y$	$5Y$	$2Y$	$5Y$
Min ratio	712	842	885	981	148	333
Max ratio	7629	7927	6575	3473	5006	4322
Variance	.023	.020	.017	.007	.005	.004

Sample ratio of volatility between weights and corresponding forwards.

where *Min ratio* and *Max ratio* are the minimum (resp. maximum) volatility ratio among the weights of a particular swap. We see that in this sample, the volatility of the weights is always several orders of magnitude lower than the volatility of the corresponding forward. This approximation of swaps as baskets of forwards with constant coefficients is the key factor behind the swaption pricing methods that we detail here.

2.2.2 The Libor market model

As Libor rates and swaps were gaining importance as the fundamental variables on which the market activity was concentrated, a set of options was created on these market rates: the caps and swaptions. Adapting the common practice taken from equity markets and the Black & Scholes (1973) framework, market operators looked for a model that would set the dynamics of the Libors or the swaps as lognormal processes. Intuitively, the lognormal assumption on prices can be justified as the effect of a central limit theorem on returns because the prices are seen as driven by a sequence of independent shocks on returns. That same reasoning cannot be applied to justify the lognormality of Libor or swap rates, which are rates of return themselves. The key justification behind this assumption must then probably be found in the legibility and familiarity of the pricing formulas that are obtained: the market quotes the options on Libors and swaps in terms of their Black (1976) volatility by habit, it then naturally tries to model the dynamics of these rates as lognormal.

Everything works fine when one looks at these prices and processes individually, however some major difficulties arise when one tries to define yield curve dynamics that jointly reproduce the lognormality of Libors and swaps. In fact, it is not possible to find arbitrage free dynamics à la Heath et al. (1992) that make both swaps and Libors lognormal under the appropriate forward measures (see Musiela & Rutkowski (1997) or Jamshidian (1997) for an extensive discussion of this). Here we choose to adopt the Heath et al. (1992) model structure defined in Brace et al. (1997) (see also Miltersen et al. (1995), Miltersen et al. (1997) or Sandmann & Sondermann (1997)) where the Libor rates are specified as lognormal under the appropriate forward measures but we will see in a last section that for the purpose of pricing options on swaps, one can in fact approximate the swap by a lognormal diffusion in this setup. Hence in a very reassuring conclusion, as observed empirically in Brace et al. (1999), we notice that it is in fact possible to specify Heath et al. (1992) dynamics that are reasonably close to the market practice, i.e. lognormal on forwards and close to lognormal on swaps. In particular, we verify that the key assumption in this approximation, namely the stability of the weights $\omega_i(t),$ is indeed accurately reproduced by the Libor market model.

To build the model, we start from the key assumption that for a given maturity δ (for ex. 3 months) the associated forward Libor rate process $\{L(t,\theta); t \geq 0\}$ defined by

$$1 + \delta L(t,\theta) = \exp\left(\int_{\theta}^{\theta+\delta} r(t,\nu)d\nu\right) \tag{2.12}$$

has a log-normal volatility structure:

$$dL(t,\theta) = (...)dt + L(t,\theta)\gamma(t,\theta)dW_t \tag{2.13}$$

where the deterministic function $\gamma : \mathbb{R}_+^2 \longmapsto \mathbb{R}_+^d$ is bounded by some $\bar{\gamma} \in \mathbb{R}_+$ and piecewise continuous.

Although the spirit of the market model is to directly define the evolution of forward rates that are actually quoted by the market, its dynamics are still shaped by the classical no-arbitrage conditions between zero-coupon bonds as detailed in Heath et al. (1992) and as for all Heath et al. (1992) based models, these dynamics are fully specified by the definition of the volatility structure and the forward curve today. With that in mind, we recall here the Libor Market Model setup defined in Brace et al. (1997) to derive the appropriate zero-coupon volatility expression. Using the Ito formula combined with 2.6 we get as in Brace et al. (1997):

$$\begin{aligned} dL(t,\theta) = &\left(\frac{\partial L(t,\theta)}{\partial\theta} + \frac{(1+\delta L(t,\theta))}{\delta}\sigma^B(t,\theta+\delta)(\sigma^B(t,\theta+\delta) - \sigma^B(t,\theta))\right)dt \\ &+ \frac{1}{\delta}\left(1+\delta L(t,\theta)\right)(\sigma^B(t,\theta+\delta) - \sigma^B(t,\theta))dW_t \end{aligned}$$

Again, in order to avoid possible confusion, we will call this the *sliding* dynamics of the forward Libors, referring to the fact that the time to maturity of the Libor considered does not vary. Then to get the right volatility structure we have to impose in (2.5):

$$\sigma^B(t,\theta+\delta)-\sigma^B(t,\theta)=\frac{\delta L(t,\theta)}{1+\delta L(t,\theta)}\gamma(t,\theta) \tag{2.14}$$

The Libor process becomes:

$$dL(t,\theta)=\left(\frac{\partial}{\partial\theta}L(t,\theta)+\gamma(t,\theta)\sigma^B(t,\theta+\delta)L(t,\theta)\right)dt+L(t,\theta)\gamma(t,\theta)dW_t$$

As in Musiela & Rutkowski (1997), we set $\sigma^B(t,\theta)=0$ for all $\theta\in[0,\delta[$ and we get, together with the recurrence relation 2.14 and for $\theta\geq\delta$:

$$\sigma^B(t,\theta)=\sum_{k=1}^{\lfloor\delta^{-1}\theta\rfloor}\frac{\delta L(t,\theta-k\delta)}{1+\delta L(t,\theta-k\delta)}\gamma(t,\theta-k\delta) \tag{2.15}$$

which is stochastic.

Remark 6 *This assumption made in Brace et al. (1997) to set $\sigma^B(t,\theta)=0$ for all $\theta\in[0,\delta[$ is not very intuitive at first sight as one would expect the volatility of short discount factors to be set quite high. In practice however, this volatility has no impact on the changes of forward measure that will be made because these only involve differences of zero-coupon volatilities. Furthermore, this zero coupon volatility is very small compared to the dominant forward Libor volatility and we can neglect it in because we focus on the pricing of options on forwards. Hence, although incoherent with econometric intuition, this assumption has a minimal impact on option pricing.*

Finally the $\delta-$Libor process can also be written as:

$$\begin{aligned}dL(t,\theta)=&\left(\frac{\partial L(t,\theta)}{\partial\theta}+\gamma(t,\theta)\sigma^B(t,\theta)L(t,\theta)+\frac{\delta L(t,\theta)^2}{1+\delta L(t,\theta)}\left|\gamma(t,\theta)\right|^2\right)dt\\&+L(t,\theta)\gamma(t,\theta)dW_t\end{aligned}$$

with the volatility of the zero coupon defined above and the value of the forward curve today, we have fully specified the yield curve dynamics.

2.3 Interest rate options: caps and swaptions

The shift in activity from bonds towards swaps has naturally been accompanied by the apparition of options on swap and Libor rates. These options, namely caps and swaptions, are now the most liquid recipients of volatility information in interest rate markets.

2.3.1 Caps

Let us note again $\beta(t)$, the value of the savings account. In a forward cap on principal 1 settled in arrears at times T_j, $j=1,...,n$, the cash-flows are $(L(T_{j-1},0)-K)^+\delta$ paid at time T_j. The price of the cap at time t is then computed as:

$$\begin{aligned}cap_t&=\sum_1^n E_t\left[\frac{\beta_t}{\beta_{T_j}}\left(L(T_{j-1},0)-K\right)^+\delta\right] \qquad (2.16)\\&=\sum_1^n B(t,T_j)E_t^{T_j}\left[\left(L(T_{j-1},0)-K\right)^+\delta\right]\end{aligned}$$

The cap can be seen as a sum of caplets where each individual caplet price computed as that of a Call on the corresponding forward Libor:

$$E_t\left[\frac{\beta_t}{\beta_{T_j}}\left(L(T_{j-1},0)-K\right)^+\delta\right]$$

2.3.2 Swaptions

To simplify the notations, in what follows we will consider that the calendars described above for the floating and the fixed legs of the swap are set by $T_i^{float}=i\delta$ and $T_i^{fixed}=ib\delta$, in the common case where the fixed coverage is a multiple of the floating coverage (for ex. quarterly floating leg, annual fixed leg). For simplicity, we will note the coverage function for the fixed leg of the swap as a function of the floating dates, allowing the floating dates to be used as reference in the entire swap definition. From now on $(T_i)_{i\in[1,N]}=(T_i^{float})_{i\in[1,N]}$ and we define the coverage function for the fixed leg as $cvg(i,b)\delta=1_{\{i\bmod b=0\}}b\delta$. We set $i_T=\lfloor\delta^{-1}T\rfloor$. We are now in position to write the forward swap rate as:

$$swap(t,T,T_N)=\frac{B(t,T)-B(t,T_{N+1})}{Level(t,T,T_N)}$$

with

$$Level(t,T,T_N)=\sum_{i=i_T}^{N}\delta cvg(i,b)B(t,T_{i+1})$$

The price of a payer swaption with maturity T and strike k, written on this swap is then given at time $t\leq T$ by:

$$swaption_t=B(t,T)E_t^{Q_T}\left[\sum_{i=i_T}^{N}\frac{\beta(T)}{\beta(T_{i+1})}cvg(i,b)\delta\left(swap(T,T,T_N)-k\right)^+\right] \tag{2.17}$$

The expression above computes the price of the swaption as the sum of the corresponding swaplet prices. We first notice that we can think of a caplet as an option on a particular one period swap, hence caplet and swaption prices can be computed in the same fashion. In the two sections that follow, we will also show how to rewrite this pricing expression to describe the swaption (and the caplet) as a basket option.

2.4 Cap and swaption prices in the Gaussian H.J.M. model

We suppose now that the dynamics of the zero-coupon prices are given as in (2.9) by a lognormal process:

$$\frac{dB(s,T)}{B(s,T)}=r_sds+\sigma^B(s,T-s)dW_s$$

where $\sigma^B(s,T-s)\in\mathbb{R}^d$ is deterministic and bounded. We will now compute the price of caps and swaptions as **options on a basket of zero coupon bonds**.

2.4.1 Caps

The cap price can be computed as:

$$cap_t=\sum_{j=1}^{n}E_t\left[\frac{\beta_t}{\beta_{T_j}}\delta\left(L(T_{j-1},0)-k\right)^+\right]$$

or again

$$cap_t = \sum_{j=1}^{n} B(t,T_j) E_t^{T_j} \left[\delta \left(L(T_{j-1},0) - k \right)^+ \right]$$

where E^{T_j} is the expectation under the forward martingale measure $\mathbf{Q}_{Tj}$ defined by:

$$\begin{aligned} \frac{d\mathbf{Q}_{Tj}}{d\mathbf{Q}} &= \beta_t [B(0,T)\beta_{T_j}]^{-1} \\ &= \varepsilon_{Tj}(\sigma^B(\cdot, T_j - \cdot)) \end{aligned}$$

where we have noted $\varepsilon_{T_j}(\cdot)$ the lognormal martingale defined by ($\sigma^B(s, T_j - s)$ is deterministic):

$$\varepsilon_{Tj}(\sigma^B(\cdot, T_j - \cdot)) = \exp\left(\int_t^{T_j} \sigma^B(s, T_j - s) dW_s - \frac{1}{2} \int_t^{T_j} \left\| \sigma^B(s, T_j - s) \right\|^2 ds \right)$$

As before, we define the forward Libor process (or FRA) as the underlying $K(t,T) = L(t, T-t)$ of the caplet paid at time $T + \delta$. In the Gaussian H.J.M. framework we can write the cap as an option on a bond, with:

$$L(t,0) = \frac{1}{\delta} \left(B(t, t+\delta)^{-1} - 1 \right)$$

the cap price can be computed as:

$$\begin{aligned} cap_t &= \sum_{j=1}^{n} B(t,T_{j-1}) E_t^{T_{j-1}} \left[\delta B(T_{j-1},T_j) \left(L(T_{j-1},0) - k \right)^+ \right] \\ &= \sum_{j=1}^{n} B(t,T_{j-1}) E_t^{T_{j-1}} \left[B(T_{j-1},T_j) \left(B(T_{j-1},T_j)^{-1} - 1 - \delta k \right)^+ \right] \end{aligned}$$

or finally

$$cap_t = \sum_{j=1}^{n} B(t,T_{j-1}) E_t^{T_{j-1}} \left[\left(1 - B(T_{j-1},T_j) - \delta k B(T_{j-1},T_j) \right)^+ \right]$$

This sets the cap as a sum of Puts on zero coupon bonds with coupon equal to $1 + k\delta$ and each caplet is a Put on a zero-coupon bond.

Remark 7 *If we look at the definition of the forwards*

$$L(t,0) = \frac{1}{\delta} \left(B(t, t+\delta)^{-1} - 1 \right)$$

we notice that the rate $K(t,T)$ have shifted lognormal dynamics under $P_{T+\delta}$.

2.4.2 Swaptions

With the price of a payer swaption with maturity T and strike k, written on this swap is then given at time $t \leq T$ by:

$$swaption_t = B(t,T) E_t^{Q_T} \left[\sum_{i=i_T}^{N} \frac{\beta(T)}{\beta(T_{i+1})} cvg(i,b) \delta \left(swap(T,T,T_N) - k \right)^+ \right]$$

we can write:

$$
\begin{aligned}
& E_t^{Q_T}\left[\sum_{i=i_T}^{N}\frac{\beta(T)}{\beta(T_{i+1})}cvg(i,b)\delta\left(swap(T,T,T_N)-k\right)^+\right] \\
= \; & E_t^{Q_T}\left[\sum_{i=i_T}^{N}Level(T,T,T_N)\delta\left(\frac{1-B(T,T_{N+1})}{Level(T,T,T_N)}-k\right)^+\right] \\
= \; & E_t^{Q_T}\left[\sum_{i=i_T}^{N}\delta\left(1-B(T,T_{N+1})-kLevel(t,T,T_N)\right)^+\right]
\end{aligned}
$$

hence finally:

$$
swaption_t = B(t,T)E_t^{Q_T}\left[\sum_{i=i_T}^{N}\delta\left(1-B(T,T_{N+1})-\sum_{i=i_T}^{N}k\delta cvg(i,b)B(T,T_{i+1})\right)^+\right]
$$

Hence in the Gaussian H.J.M. model, swaptions can be seen as a options on *a basket of lognormal zero coupon bonds*.

2.5 Caps and swaptions in the Libor market model

As above in (2.14), we now specify the Heath et al. (1992) volatility so that the Libor rates have a lognormal volatility structure:

$$
dL(t,\theta) = (...)dt + L(t,\theta)\gamma(t,\theta)dW_t
$$

we will now show how to compute the price of caps and swaptions under these assumptions.

2.5.1 Caps and the forward martingale measure

Again, with the cap price computed as:

$$
cap_t = \sum_{j=1}^{n} E_t\left[\frac{\beta_t}{\beta_{T_j}}\delta\left(L(T_{j-1},0)-K\right)^+\right]
$$

which can be written

$$
cap_t = \sum_{j=1}^{n} B(t,T_j)E_t^{T_j}\left[\left(L(T_{j-1},0)-K\right)^+\delta\right]
$$

where E^{T_j} is the expectation under the forward martingale measure defined by:

$$
\begin{aligned}
\frac{d\mathbf{Q}_{Tj}}{d\mathbf{Q}} &= \beta_t[B(0,T)\beta_{T_j}]^{-1} \\
&= \varepsilon_T(\sigma^B(\cdot,T_j-\cdot))
\end{aligned}
$$

where we have noted $\varepsilon_{T_j}(\cdot)$ the exponential martingale defined by:

$$
\varepsilon_{T_j}(\sigma^B(\cdot,T_j-\cdot)) = \exp\left(\int_t^{T_j}\sigma^B(s,T_j-s)dW_s - \frac{1}{2}\int_t^{T_j}\left\|\sigma^B(s,T_j-s)\right\|^2 ds\right)
$$

Let us now define the forward Libor process (or FRA), the underlying $K(t,T) = L(t,T-t)$ of the caplet paid at time $T+\delta$, which is given in the Libor market model setup in (2.13) by:

$$dK(t,T) = \gamma(t,T-t)K(t,T)\left[\sigma^B(t,T-t+\delta)dt + dW_t\right]$$

or again:

$$dK(t,T) = \gamma(t,T-t)K(t,T)dW_t^{T+\delta} \tag{2.18}$$

hence $K(t,T)$ is lognormally distributed under $\mathbf{Q}_{T+\delta}$. The pricing of caplets can be done using the Black (1976) formula with variance V_T such that:

$$V_T = \int_t^T \|\gamma(s,T-s)\|^2 \, ds$$

Let us note that the caplet variance used in the Black (1976) pricing formula is a linear form in the covariance matrix Γ_s. Recovering the same kind of result in the swaption pricing approximation will be the key to the calibration algorithm design.

2.5.2 Swaptions and the forward swap martingale measure

Again, with the price of a payer swaption with maturity T and strike k, written on the above swap is then given at time $t \leq T$ by:

$$Ps(t) = B(t,T)E_t^{Q_T}\left[\sum_{i=i_T}^{N} \frac{\beta(T)}{\beta(T_{i+1})} cvg(i,b)\delta\left(swap(T,T,T_N) - k\right)^+\right] \tag{2.19}$$

The expression above computes the price of the swaption as the sum of the corresponding swaplet prices, however its format is not the most appropriate for our pricing purposes. Therefore, using again a change of equivalent probability measure, we will now find another expression that is more suitable for our analysis.

Definition 8 *As in Musiela & Rutkowski (1997) or Jamshidian (1997), we can define the forward swap martingale probability measure $\mathbf{Q}^{LVL}$ equivalent to $\mathbf{Q}^T$, with:*

$$\begin{aligned}\frac{d\mathbf{Q}^{LVL}}{d\mathbf{Q}^T}\Big|_t &= \frac{\sum_{i=i_T}^{N} cvg(i,b)\beta(T)/\beta(T_{i+1})}{E_t^{Q_T}\left[\sum_{i=i_T}^{N} cvg(i,b)\beta(T)/\beta(T_{i+1})\right]} \\ &= B(t,T)\beta(T)\sum_{i=i_T}^{N} \frac{\delta cvg(i,b)\beta^{-1}(T_{i+1})}{Level(t,T,T_N)}\end{aligned}$$

this equivalent probability measure corresponds to the choice of the ratio of the level payment over the savings account as a numeraire and the above relative bond prices are $\mathbf{Q}^T$–local martingale.

The change of measure is identified with an exponential (local) $\mathbf{Q}^T$–martingale and we define the process h_t such that:

$$\varepsilon_{T_N}(h_\bullet) = B(t,T)\beta(T)\frac{\sum_{i=i_T}^{N} \delta cvg(i,b)\beta^{-1}(T_{i+1})}{Level(t,T,T_N)}$$

which imposes:

$$h_t = \sum_{i=i_T}^{N} \frac{\delta cvg(i,b)B(t,T_{i+1})}{Level(t,T,T_N)} \left(\sum_{j=i_T}^{i} \frac{\delta K(t,T_j)}{1+\delta K(t,T_j)} \gamma(t,T_j - t) \right) \tag{2.20}$$

and because the volatility is bounded, we verify that $\varepsilon_{T_N}(h_\bullet)$ is in fact a martingale. Again as in Musiela & Rutkowski (1997) we can apply Girsanov's theorem to show that the process:

$$dW_t^{LVL} = dW_t^T + \sum_{i=i_T}^{N} \left(\frac{\delta cvg(i,b)B(t,T_{i+1})}{level(t,T,T_N)} \sum_{j=1}^{i} \frac{\delta K(t,T_j)}{1+\delta K(t,T_j)} \gamma(t,T_j - t) \right) dt \tag{2.21}$$

is a $\mathbf{Q}^{LVL}$-Brownian motion.

Proposition 9 *We can rewrite the swaption price as:*

$$swaption_t = Level(t,T,T_N) E_t^{Q_{LVL}} \left[(swap(T,T,T_N) - k)^+ \right] \tag{2.22}$$

where the swap rate is a martingale under the new probability measure $\mathbf{Q}^{LVL}$.

Proof. From the definition of the swaption price

$$swaption_t = B(t,T) E_t^{Q_T} \left[\sum_{i=i_T}^{N} \frac{\beta(T)}{\beta(T_{i+1})} cvg(i,b) \delta \left(swap(T,T,T_N) - k \right)^+ \right]$$

we get:

$$swaption_t = Level(t,T,T_N) E_t^{Q_T} \left[\varepsilon_{T_N}(h_\bullet) \left(swap(T,T,T_N) - k \right)^+ \right]$$

which is also, by construction of the probability measure $\mathbf{Q}^{LVL}$:

$$swaption_t = Level(t,T,T_N) E_t^{Q_{LVL}} \left[(swap(T,T,T_N) - k)^+ \right]$$

Because the swap is defined by:

$$swap(t,T,T_N) = \frac{B(t,T) - B(t,T_{N+1})}{Level(t,T,T_N)}$$

as the ratio of a difference of zero-coupon prices over the level payment, by construction the swap will be a (local) martingale under the new probability measure $\mathbf{Q}^{LVL}$ (below, we will see that the swap rate is in fact a $\mathbf{Q}^{LVL}-$martingale). ■

This change of measure first detailed by Jamshidian (1997), allows to price the swaption as a classical Call option on a swap, under an appropriate measure.

2.5.3 Swap dynamics

In a previous section, we have seen that the swaption could be written as an option on a basket of zero-coupon bonds. Here we will try to show that the same swaption can also be seen as an option on a basket of forward Libor rates. In that spirit, we now study the dynamics of the swap rate under the new $\mathbf{Q}^{LVL}$ probability, looking first for an appropriate representation of the volatility function by expressing the swap rate volatility under $\mathbf{Q}^{LVL}$ in its "basket of forwards" decomposition:

$$\begin{aligned} swap(t,T,T_N) &= \sum_{i=i_T}^{N} \left(\frac{\delta B(t,T_{i+1})}{level(t,T,T_N)} \right) K(t,T_i) \\ &= \sum_{i=i_T}^{N} \omega_i(t) K(t,T_i) \end{aligned}$$

We start by detailing the weights' dynamics. Again, we note

$$\sigma^B(t,T_i-t) = \sum_{j=i_T}^{i-1} \frac{\delta K(s,T_j)}{1+\delta K(s,T_j)} \gamma(s,T_j-s)$$

the forward zero coupon volatility defined in (2.5).

Lemma 10 *The weights $\omega_k(s)$ in the swap decomposition follow:*

$$d\omega_k(s) = \omega_k(s) \sum_{i=i_T}^{N} \omega_i(s) \left(\sigma^B(s,T_{k+1}-s) - \sigma^B(s,T_{i+1}-s) \right)$$

Proof. The weights $\omega_i(t)$ are defined by:

$$\omega_i(t) = \frac{\delta cvg(i,b) B(t,T_{i+1})}{level(t,T,T_N)}$$

as the ratio of a zero coupon bond on the level payment. By construction of $\mathbf{Q}^{LVL}$, these weights $\omega_i(t)$ have to be $\mathbf{Q}^{LVL}-$martingales (they are also positive bounded). Using the definition of the forward zero-coupon dynamics we get:

$$\begin{aligned} d\left(\frac{B(s,T_k)}{Level(t,T,T_N)} \right) &= (...)ds + \frac{B(t,T_k)}{Level(t,T,T_N)} \sigma^B(s,T_k-s) dW_s^T \\ & -\frac{B(t,T_k)}{Level(t,T,T_N)} \sum_{i=i_T}^{N} \frac{\delta cvg(i,b) B(t,T_{i+1})}{Level(t,T,T_N)} \sigma^B(s,T_{i+1}-s) dW_s^T \end{aligned}$$

where W_s^T is a $\mathbf{Q}^T$-Brownian motion. ■

We can then use this result to decompose the swap volatility.

Lemma 11 *We can decompose the swap volatility as the sum of the weights volatility term and a term that mimics a basket volatility (the volatility of a basket of with constant coefficients):*

$$dswap(s,T,T_N) = \left(b_{weights}(s) + b_{basket}(s) \right) dW_s^{LVL} \tag{2.23}$$

where the weight's contribution is given by:

$$b_{weights}(s) = \sum_{k=i_T}^{N} \omega_k(s) K(s,T_k) \left(\sigma^B(s,T_{k+1}-s) - \sum_{i=i_T}^{N} \omega_i(s) \sigma^B(s,T_{i+1}-s) \right)$$

and the basket volatility term is:

$$b_{basket}(s) = \sum_{i=i_T}^{N} \omega_i(s) K(s, T_i)\gamma(s, T_i - s)$$

i.e. a $\omega_i(s)$ weighted average of the forward volatilities.

Proof. We can compute the weight's contribution in the swap volatility as:

$$b_{weights}(s) = \sum_{k=i_T}^{N} K(s, T_k)\omega_k(s) \sum_{i=i_T}^{N} \omega_i(s) \left(\sigma^B(s, T_{k+1} - s) - \sigma^B(s, T_{i+1} - s)\right)$$

where as above, $\sigma^B(t, T_i - t)$ is the zero-coupon volatility. We can then get the contribution of the forwards to the volatility as:

$$\begin{aligned} b_{basket}(s) &= \sum_{i=i_T}^{N} \frac{\delta B(s, T_{i+1}) K(s, T_i)\gamma(s, T_i - s)}{Level(s, T, T_N)} \\ &= \sum_{i=i_T}^{N} \omega_i(s) K(s, T_i)\gamma(s, T_i - s) \end{aligned}$$

hence the desired result. ■

The empirical stability of the weights $\omega_i(t)$ discussed in a previous section is the key finding at the origin of the swaption pricing approximations that will follow and one of our goals below will be to show that this stability is *accurately predicted by the model*.

2.5.4 The forwards under the forward swap measure

We study here the dynamics of the forward Libors under the forward swap measure. For purely technical purposes, we start by bounding under $\mathbf{Q}^{LVL}$ the variance of the forward rates $K(s, T_k)$, we will then be able to bound the contribution of the weights to the total swap variance.

Lemma 12 *With $m > 1$, we can bound the L^2 norm of $K(u, T_k)$ under $\mathbf{Q}^{LVL}$ by:*

$$E[K(s, T_k)^m] \le K(t, T_k)^m M_m^m(s) \tag{2.24}$$

where $M_m(s) = \exp\left((s - t)\left(m\bar{\gamma}^2/2 + m\bar{\gamma}^2\delta(N - i_T)\right)\right)$.

Proof. Using (2.21) we can write:

$$K(s, T_k) = K(t, T_k) \exp\left(\int_t^s \gamma(u, T_k - u) dW_u^{LVL} + \int_t^s \alpha(u, T_k)\gamma(u, T_k - u) du\right)$$

where

$$\alpha(s, T_k) = -\sum_{i=i_T}^{N} \omega_i(s) \left(\sum_{j=i_T}^{i} \phi_j(s)\gamma(s, T_j - s)\right) + \sum_{i=i_T}^{k} \phi_i(s)\gamma(s, T_i - s)$$

with $\phi_i(t) = \delta K(s,T_i)/(1+\delta K(s,T_i))$. The corresponding forward Libor rates are positive and we have $0 \leq \phi_i(t) \leq 1$ and as in Brace et al. (1997) remark 2.3, we can bound the forwards by a lognormal process:

$$K(s,T_k) \leq K(t,T_k)\exp\left(\int_t^s \gamma(u,T_k-u)dW_u^{LVL} + \int_t^s \bar{\alpha}(u,T_k)du\right) \quad \text{for } s \in [t,T]$$

where we can use a convexity inequality on the norm $\|.\|^2$ to obtain:

$$\left\|\sum_{i=i_T}^{N} \omega_i(s)\left(\sum_{j=i_T}^{i} \phi_j(s)\gamma(s,T_j-s)\gamma(s,T_k-s)\right)\right\|^2 \leq \delta^2(N-i_T)^2\bar{\gamma}^4$$

because $\left\|\sum_{i=i_T}^{k} \phi_i(t)\gamma(s,T_i-s)\gamma(s,T_k-s)\right\|^2 \leq \delta^2(k-i_T)^2\bar{\gamma}^4$, hence $\bar{\alpha}(s,T_k) = \delta(N-i_T)\bar{\gamma}^2$, which shows the desired result. ■

We can now use this bound to study the dynamics of the weights $\omega_i(t)$ in the swap decomposition.

2.5.5 Swaps as baskets of forwards

For simplicity, in what follows we will suppose that $T_i^{float} = T_i^{fixed}$ and hence $b = 1$. The swaption pricing formula that will be derived in the next chapter relies on two fundamental approximations:

- The weights $\omega_i(s)$ for $s \in [t,T]$ (which are $\mathbf{Q}^{LVL}$-martingales) will be approximated by their value today $\omega_i(t)$.
- We will neglect the change of measure between the forward martingale measures $\mathbf{Q}^{T}$,..., $\mathbf{Q}^{T_{N+1}}$ and the forward swap martingale measure $\mathbf{Q}^{LVL}$.

This is possible here because the weights in (2.10) are positive, monotone and sum to one. In this section, we quantify the error created by these approximations. Because the payoff of the Call options under consideration are Lipschitz, we approximate the swap and forward Libor dynamics in L^2 under the $\mathbf{Q}^{LVL}$ swap martingale measure. First, let us recall that because they are defined as the ratio of a zero-coupon bond and the level payment, the weights $\omega_i(u)$ must be martingales under the $\mathbf{Q}^{LVL}$ probability. We also recall the swap dynamics found above:

$$dswap(s,T,T_N) = \sum_{i=i_T}^{N} \omega_i(s)K(s,T_i)\left(\gamma(s,T_i-s)+\eta(s,T_i)\right)dW_s^{LVL}$$

where

$$\eta(s,T_i) = \left(\sigma^B(s,T_i-s) - \sum_{j=i_T}^{N}\omega_j(s)\sigma^B(s,T_j-s)\right)$$

is the volatility contribution coming from the weights.

Remark 13 *If the forward rate curve is flat ($K(s,T_i) = K(s,T_j)$ for $i,j = i_T,...,N$) we have:*

$$\begin{aligned}\sum_{i=i_T}^{N}\omega_i(s)K(s,T_i)\eta(s,T_i) &= K(s,T_i)\sum_{i=i_T}^{N}\omega_i(s)\left(\sigma^B(s,T_i-s) - \sum_{j=i_T}^{N}\omega_j(s)\sigma^B(s,T_j-s)\right) \\ &= 0 \end{aligned} \tag{2.25}$$

In light of this fact, we will study the size of the weights' contribution to the swap volatility in terms of the slope of the forward rate curve within the maturity range of the swap's floating leg. In the development below, we will rewrite the weight's part in the swap's volatility :

$$\begin{aligned} & E^{LVL}\left[\left\|\sum_{i=i_T}^{N}\omega_i(s)K(s,T_i)\eta(s,T_i)\right\|^2\right] \\ = \quad & E^{LVL}\left[\left\|\sum_{i=i_T}^{N}\omega_i(s)\left(K(s,T_i)-swap(s,T,T_N)\right)\eta(s,T_i)\right\|^2\right] \end{aligned}$$

This sets the weight's contribution as the average product of a difference of forwards with a difference of ZC bond volatilities. We can naturally expect this later term to be negligible relative to the basket volatility term in (2.23). We note $\|\cdot\|_n = \left(E^{LVL}\left[\|\cdot\|^n\right]\right)^{\frac{1}{n}}$, the L^n norm and we first show that the weights $\omega_i(s)$ are bounded.

Lemma 14 *The weights $\omega_i(s)$ defined in (2.10) are bounded above with:*

$$\omega_i(s) \leq \frac{1}{N-i_T} + \delta swap(s,T,T_N)$$

and satisfy $\|\omega_i(s)\|_n \leq \omega_i(t)$ for $s \in [t,T]$.

Proof. Because the weights $\omega_i(s)$ satisfy $\sum_{i=i_T}^{N}\omega_i(t) = 1$, $0 \leq \omega_i(t) \leq 1$ and are decreasing with i because the forward rates are always positive. With:

$$|\omega_j(s) - \omega_i(s)| \leq \delta swap(s,T,T_N) \text{ for } i,j = i_T,...,N$$

we get:

$$\omega_i(s) \leq \frac{1}{N-i_T} + \delta swap(s,T,T_N) \text{ for } s \in [t,T]$$

and $\|\omega_i(s)\|_n \leq \|\omega_i(s)\|_1 = \omega_i(t)$ because the weights are positive $\mathbf{Q}^{LVL}-$martingales. ■

This result provides a bound on the variance contribution of the weights inside the swap rate volatility.

Lemma 15 *The L^2 norm of the weight's contribution in the swap volatility (2.23) is bounded by:*

$$\begin{aligned} & E^{LVL}\left[\left\|\sum_{i=i_T}^{N}\omega_i(s)K(s,T_i)\eta(s,T_i)\right\|^2\right] \\ \leq \quad & \max_j \|(K(s,T_j)-swap(s,T,T_N))\|_8^2 \, M_4^2\bar{\gamma}^2\delta^2 max_{j\in[i_T,N]}K(t,T_j)^2(N-i_T)^2 \end{aligned} \tag{2.26}$$

Proof. Let us note again $swap(s,T,T_N) = \sum_{i=i_T}^{N}\omega_i(s)K(s,T_i)$, the swap rate, which we see here as the average level of the forward rate curve between T and T_N. The squared L^2 norm of the weights' contribution is bounded above by:

$$E^{LVL}\left[\sum_{i=i_T}^{N}\omega_i(s)\left\|\left(K(s,T_i)-swap(s,T,T_N)\right)\eta(s,T_i)\right\|^2\right]$$

using a convexity inequality with $\sum_{i=i_T}^{N}\omega_i(t) = 1$, $0 \leq \omega_i(t) \leq 1$. To bound $\eta(s,T_k)$ in this expression, we use the definition of $\sigma^B(s,T_k-s)$ in (2.15) and the fact that the forwards $K(s,T_j)$ are always positive, we get:

$$E^{LVL}\left[\|\eta(s,T_i)\|^4\right] \leq E^{LVL}\left[\left\|\sum_{i=i_T}^{N}\omega_i(s)\left(\sum_{j=i}^{k}\delta K(s,T_j)\gamma(s,T_j-s)\right)\right\|^4\right]$$

with the convention $\sum_{j=i}^{k} = -\sum_{j=k}^{i}$ if $i > k$. If we recall that $\gamma(s,T_k-s) : \mathbb{R}_+^2 \to \mathbb{R}_+^d$ is a bounded input parameter with $E\left[\|\gamma(s,T_k-s)\|^2\right] \leq \bar{\gamma}^2$, we use (2.24) to get:

$$E^{LVL}\left[\|\eta(s,T_k)\|^4\right] \leq M_4^4\bar{\gamma}^4\delta^4 max_{j\in[i_T,N]}K(t,T_j)^4(N-i_T)^4$$

With these bounds we can rewrite the original inequality, using two successive Cauchy inequalities:

$$\begin{aligned}
& E^{LVL}\left[\sum_{i=i_T}^{N}\omega_i(s)\left\|(K(s,T_i)-swap(s,T,T_N))\right\|^2\left\|\eta(s,T_i)\right\|^2\right] \\
\leq & \sum_{i=i_T}^{N}\|\omega_i(s)\|_4\left\|(K(s,T_i)-swap(s,T,T_N))\right\|_8^2\left\|\eta(s,T_i)\right\|_4^2 \\
\leq & \max_j\left\|(K(s,T_j)-swap(s,T,T_N))\right\|_8^2 M_4^2\bar{\gamma}^2\delta^2 max_{j\in[i_T,N]}K(t,T_j)^2(N-i_T)^2
\end{aligned}$$

Which gives the desired result. ∎

With $\delta K(t,T_k) \simeq 10^{-2}$ and $(K(t,T_i)-K(t,T_j))^2 \simeq 10^{-3}$ in practice, we notice that the contribution of the weights to the swap volatility is several orders of magnitude below that of the basket volatility and we neglect it in the swaptions pricing approximations that follow. Beyond that, the squared norm of the difference between the particular swap rate under consideration and one of its forward rates can be interpreted as the covariance of a spread. As the forward rate covariance has in practice a "spread" factor with a variance that is one order of magnitude below that of the "level" factor, we can again expect this term to be very small. Before detailing the key approximation result, we introduce some preliminary notations.

Notation 16 *We define $K^{LVL}(s,T_i)$ such that:*

$$dK^{LVL}(s,T_i) = K^{LVL}(s,T_i)\gamma(s,T_i-s)dW_s^{LVL}$$

with $K^{LVL}(t,T_i) = K(t,T_i)$. We also define the following residual volatilities:

$$\xi_k(s) = K^{LVL}(s,T_k)\gamma(s,T_k-s) - \gamma^w(s)$$

with $\gamma^w(s) = \sum_{i=i_T}^{N}\omega_i(t)K^{LVL}(s,T_i)\gamma(s,T_k-s)$.

We now approximate the swap rate with a basket of martingales with volatilities matching the forward rate volatilities $\gamma(s,T_k-s)$ and initial value $K(t,T_i)$, the weights in this decomposition being equal to $\omega_i(t)$.

Proposition 17 *We can replace the swap process by a basket Y_s of lognormal martingales weighted by constant coefficients, with:*

$$\begin{aligned}
& E\left[\left(\sup_{t\leq s\leq T}\left(swap(s,T,T_N)-Y_s\right)\right)^2\right] \\
\leq\ & 3\max_{j\in[i_T,N]}\|\xi_j(s)\|_4^2+3\left(K^{LVL}(t,T_k)\,(N-i_T)\,\delta\bar{\gamma}^2\right)^2\exp\left((T-t)\left(\delta\bar{\gamma}^2\,(N-i_T)+\bar{\gamma}^2/2\right)\right) \\
& +3\max_{j\in[i_T,N]}\|(K(s,T_j)-swap(s,T,T_N))\|_8^2\,M_4^2\bar{\gamma}^2\delta^2 max_{j\in[i_T,N]}K(t,T_j)^2(N-i_T)^2
\end{aligned}$$

where

$$dY_s=\sum_{i=i_T}^{N}\omega_i(t)K^{LVL}(s,T_i)\gamma(s,T_i-s)dW_s^{LVL}$$

with $Y_t = swap(t,T,T_N)$.

Proof. With the swap rate dynamics computed as in (2.23), we get:

$$\begin{aligned}
d(swap(s,T,T_N)-Y_s) &= \sum_{k=i_T}^{N}\left(\omega_k(s)-\omega_k(t)\right)K^{LVL}(s,T_k)\gamma(s,T_k-s)dW_s^{LVL} \\
&\quad+\sum_{k=i_T}^{N}\omega_k(s)\left(K(s,T_k)-K^{LVL}(s,T_k)\right)\gamma(s,T_k-s)dW_s^{LVL} \\
&\quad+\sum_{k=i_T}^{N}\omega_k(s)K(s,T_k)\eta(s,T_k)dW_s^{LVL}
\end{aligned}$$

Using the result in (2.26) we can bound the norm of the last term in this decomposition. If we look at the first term and note $\Delta_{k,s} = K(s,T_k) - K^{LVL}(s,T_k)$ with $\Delta_{k,t} = 0$ we have:

$$\begin{aligned}
d\Delta_{k,s} &= \Delta_{k,s}\left(\sum_{i=i_T}^{N}\omega_i(s)\left(\sigma^B(s,T_k-s)-\sigma^B(s,T_i-s)\right)\gamma(s,T_k-s)\right) \\
&\quad+K^{LVL}(s,T_k)\left(\sum_{i=i_T}^{N}\omega_i(s)\left(\sigma^B(s,T_k-s)-\sigma^B(s,T_i-s)\right)\gamma(s,T_k-s)\right)ds \\
&\quad+\Delta_{k,s}\gamma(s,T_k-s)dW_s^{LVL}
\end{aligned}$$

hence:

$$\Delta_{k,T}=K^{LVL}(T,T_k)\int_t^T\left(\mu_{k,s}\exp\left(\int_t^s\mu_{k,u}du\right)ds\right)$$

where

$$\mu_{k,s}=\sum_{i=i_T}^{N}\omega_i(s)\left(\sigma^B(s,T_k-s)-\sigma^B(s,T_i-s)\right)\gamma(s,T_k-s)$$

With $\|\mu_{k,s}\|_2 \leq (N-i_T)\,\delta\bar{\gamma}^2$ we can bound the norm of $\Delta_{k,T}$ by:

$$\|\Delta_{k,T}\|_2\leq K(t,T_k)\,(N-i_T)\,\delta\bar{\gamma}^2\exp\left((T-t)\left(\delta\bar{\gamma}^2\,(N-i_T)+\bar{\gamma}^2/2\right)\right)$$

Focusing on the second term, as in (2.25) with this time $\sum_{i=i_T}^{N} \omega_i(s) - \omega_i(t) = 0$ and $\xi_k(s) = K^{LVL}(s, T_k)\gamma(s, T_k - s) - \gamma^w(s)$, we can write:

$$\left\| \sum_{k=i_T}^{N} (\omega_k(s) - \omega_k(t)) \, K^{LVL}(s, T_k)\gamma(s, T_k - s) \right\|_2^2 \leq \max_{j \in [i_T, N]} \|\xi_j(s)\|_4^2$$

The bound obtained is a function of the norm of the residual volatilities $\|\xi_i(s)\|_4^2$ and of the spread term $\|(K(s, T_i) - swap(s, T, T_N))\|_8^2$. We conclude using Doob's inequality. ■

The term $\|\xi_i(s)\|_4^2$ being equivalent to the variance contribution of the second factor in the covariance matrix and $\|(K(s, T_i) - swap(s, T, T_N))\|_8^2$ being a spread of rates, we can neglect both terms compared to the central volatility term $\gamma^w(s)$ and we will approximate the swaption by an option on the basket Y_s. We can notice that because we approximate one martingale by another, the error is in fact uniformly bounded in L^2. Because of these properties, in the Libor Market Model swaption price approximations that follow, we will be treating *swaptions as options on a basket of lognormal forwards*. In fact, to summarize this first chapter, in *both* the Gaussian H.J.M. model and the Libor Market Model, we have written the swaption price as that of a basket option on lognormal processes.

Chapter 3

Basket pricing

Basket options, i.e. options on a basket of goods, have become a pervasive instrument in financial engineering. Besides the swaptions described in the previous chapter, this class of instruments includes index options and exchange options in the equity markets, or yield curve options and spread options in fixed income markets. In these key markets, baskets provide raw information about the correlation between instruments, which is central to the pricing of more complex derivatives. In this work, we detail an efficient pricing approximation technique that leads to very natural closed-form basket pricing formulas with excellent precision results.

In a first section, we show how to quickly recover the classical ”noise addition in decibels” order zero lognormal approximation studied by Huynh (1994), Musiela & Rutkowski (1997) and Brace et al. (1999) when the underlying instruments follow a Black & Scholes (1973) like lognormal diffusion.

In a second section we approximate the price of a basket using the stochastic expansion techniques exploited by Fournié et al. (1997) and Lebuchoux & Musiela (1999) or Fouque, Papanicolaou & Sircar (2000) on other stochastic volatility problems. This provides a theoretical justification for the classical price approximation and allows us to compute additional terms better accounting for the stochastic nature of the basket volatility. The order zero term in this expansion matches the classical lognormal approximation while the order one correction can be interpreted as a first order approximation of the hedging tracking error, as defined in El Karoui et al. (1998).

Finally, we test the quality of the basket pricing approximation on swaptions by comparing the approximate prices obtained with Monte-Carlo simulations.

3.1 Basket price approximation

We suppose that the market is composed of n risky assets $S_t^i, i = 1, \ldots, n$ plus one riskless asset M_t. We assume that these processes are defined on a probability space $(\Omega, \mathcal{F}, \mathbf{Q})$ and are adapted to the natural filtration $\{\mathcal{F}_t, 0 \leq t \leq T\}$. We suppose that there exists a forward martingale measure $\mathbf{Q}$ as defined in El Karoui, Geman & Rochet (1995) (the notation $\mathbf{Q}$ is left voluntarily non specific for our purposes here because it can either be associated with the forward market of maturity T and constructed by taking the savings account as a numeraire or it could be the level payment induced measure as in the swaption pricing formulas treated in the first chapter). In this market, the dynamics of the forwards F_t^i are given by:

$$dF_s^i = F_s^i \sigma_s^i dW_s \quad \text{and} \quad M_s = 1 \;\; \text{for } s \in [t, T]$$

where W_t is a d-dimensional **Q**-Brownian motion adapted to the filtration $\{\mathcal{F}_t, 0 \leq t \leq T\}$ and $\sigma_s = \left(\sigma_s^i\right)_{i=1,\ldots,n} \in \mathbb{R}^{n \times d}$ is the volatility matrix and we note $\Gamma_s \in \mathbb{R}^{n \times n}$ the corresponding covariance matrix defined as $(\Gamma_s)_{i,j} =< \sigma_s^i, \sigma_s^j >$.

Remark 18 *We use the notations S_t^i and F_t^i to keep our discussion generic and highlight the fact that the product range here includes equity baskets in the classical Black & Scholes (1973) framework, but the link with the previous chapter and swaption pricing can easily be made explicit by setting $F_t^i = K(t, T_i)$ and $\sigma_s^i = \gamma(s, T_i - s)$.*

We study the pricing of an option on a basket of forwards given by

$$F_t^\omega = \sum_{i=1}^n \omega_i F_t^i$$

where $\omega = (\omega_i)_{i=1,\ldots,n} \in \mathbb{R}^n$. The terminal payoff of this option at maturity T is computed as:

$$h\left(F_T^\omega\right) = \left(\sum_{i=1}^n \omega_i F_T^i - k\right)^+$$

for a strike price k. The key observation at the origin of the following approximations is that the basket process dynamics are close to lognormal. The simple formula for basket prices that we will get is specifically based on a deterministic approximation of the lognormal basket volatility in:

$$dF_s^\omega = F_s^\omega \left(\sum_{i=1}^n \widehat{\omega}_{i,s} \sigma_s^i\right) dW_s \tag{3.1}$$

where we have defined:

$$\widehat{\omega}_{i,s} = \frac{\omega_i F_s^i}{\sum_{i=1}^n \omega_i F_s^i}$$

Here, we compute a first simple approximation using Wiener chaos expansion. We then look for an extra term using an approximation of the volatility dynamics and small noise expansion. This first has its origin in the electrical engineering literature as a classic problem in signal processing where it represents, for example, the addition of noise in decibels (see Schwartz & Yeh (1981) among others). The same approximations were then used in finance by Huynh (1994), Musiela & Rutkowski (1997) for equity baskets or Brace et al. (1999) for swaptions.

3.1.1 The classical approximation

We try here to give a straightforward justification of the classical order zero approximation. Intuitively, the forward basket follows a lognormal diffusion with a (mildly) stochastic volatility. We look here for the "best" possible lognormal approximation to the basket dynamics in (3.1). As already pointed out by Lacoste (1996), the direct expansion of (3.1) in the successive chaos of a lognormal martingale, although it seems to be a very natural approximation technique, does not lead to workable approximations in practice. This is because the key orthogonal decomposition property: $L^2(\Omega, G, \mathbf{Q}) = \oplus_{n=0}^\infty \mathcal{H}_n$ (where G is the $\sigma-$field generated by $\left\{\int \sigma_s^i dW_s\right\}_{i=1,\ldots,n}$) is lost when the volatilities σ_s^i of the fundamental martingales are stochastic. In fact, as the lognormal martingales subspace is not convex, there exists no natural metric to describe the "lognormality" of a process.

Let us first detail the dynamics of the weights $\widehat{\omega}_{i,s}$, we have:

$$\begin{aligned}\frac{d\widehat{\omega}_{i,s}}{\widehat{\omega}_{i,s}} &= \left(Tr\left(\Gamma_s\widehat{\omega}_s\widehat{\omega}_s^T\right)ds - \sigma_s^i\sum_{j=1}^n\widehat{\omega}_{j,s}\sigma_s^j\right)ds \\ &\quad + \left(\sigma_s^i - \sum_{j=1}^n\widehat{\omega}_{j,s}\sigma_s^j\right)dW_s\end{aligned}$$

where we have noted $\Gamma_s = \sigma_s\sigma_s^T$ and $\widehat{\omega}_s = (\widehat{\omega}_{j,s})_{j=1,...,n}$. This is again:

$$\frac{d\widehat{\omega}_{i,s}}{\widehat{\omega}_{i,s}} = \left(\sum_{j=1}^n\widehat{\omega}_{j,s}\left(\sigma_s^i - \sigma_s^j\right)\right)\left(dW_s + \sum_{j=1}^n\widehat{\omega}_{j,s}\sigma_s^j ds\right)$$

because $\sum_{j=1}^n\widehat{\omega}_{j,s} = 1$ for $s \in [t,T]$. If the relative volatility variations $\left\|\sigma_s^i - \sigma_s^j\right\| / \|\Gamma_s\|$ are small, then the contribution of the weights volatility given by the above dynamics to the total basket volatility can be neglected relative to that of the forwards. This means that we can compute $E[\widehat{\omega}_{i,s}] = \widehat{\omega}_{i,t}$ and get the first order chaos expansion of $\int\sum_{i=1}^n\widehat{\omega}_{i,s}\sigma_s^i dW_s$ using the Taylor Stroock formula as in Nualart (1995) p.33:

$$\int\sum_{i=1}^n\widehat{\omega}_{i,s}\sigma_s^i dW_s \simeq \int\sum_{i=1}^n\widehat{\omega}_{i,t}\sigma_s^i dW_s$$

which gives the following lognormal basket dynamics:

$$dF_s^\omega = F_s^\omega\left(\sum_{i=1}^n\widehat{\omega}_{i,t}\sigma_s^i\right)dW_s$$

This method helps to understand the origin of the classical approximation of a basket of lognormal processes as another lognormal process by moments matching. It has however two severe shortcomings. First, it approximates the process itself, hence it is by definition a suboptimal approximation of the price of any nonlinear security. Secondly, it does not easily allow for the computation of subsequent terms in the approximation. The approximation technique detailed in the next section corrects both deficiencies.

3.1.2 Diffusion approximation

We can now look for an extra term that better accounts for the (mildly) stochastic nature of the lognormal basket volatility and improves the pricing approximation outside of the money. The approximation above simply expresses the fact that if all the forward volatility vectors were equal then the basket diffusion would then be exactly lognormal. It is then quite natural to look for an extra term by developing the above approximation around the central first-order volatility vector $\sum_{j=1}^n\widehat{\omega}_{i,t}\sigma_s^j$. We first define the residual volatility ξ_s^i as the difference between the original volatility σ_s^i and the central basket volatility $\sum_{j=1}^n\widehat{\omega}_{j,t}\sigma_s^j$:

Notation 19 *We set for $s \in [t,T]$:*

$$\xi_s^i = \sigma_s^i - \sum_{j=1}^n\widehat{\omega}_{j,t}\sigma_s^j \quad \textit{for } i = 1,...,n$$

and we note:

$$\sigma_s^\omega = \sum_{j=1}^{n} \widehat{\omega}_{j,t} \sigma_s^j$$

notice that σ_s^ω *is* $\mathcal{F}_t - measurable.$

We can then write the dynamics of the basket F_s^ω in terms of $\widehat{\omega}_{i,s}$ and the residual volatilities ξ_s^i. From

$$\begin{cases} dF_s^\omega = F_s^\omega \left(\sigma_s^\omega + \sum_{j=1}^n \widehat{\omega}_{j,s} \xi_s^j \right) dW_s \\ \\ d\widehat{\omega}_{i,s} = \widehat{\omega}_{i,s} \left(\sum_{j=1}^n \widehat{\omega}_{j,s} (\xi_s^i - \xi_s^j) \right) \left(dW_s + \sigma_s^\omega ds + \sum_{j=1}^n \widehat{\omega}_{j,s} \xi_s^j ds \right) \end{cases} \tag{3.2}$$

Remember that for $s \in [t,T]$ we have $\widehat{\omega}_{j,s} \geq 0$ with $\sum_{j=1}^n \widehat{\omega}_{j,s} = 1$, hence σ_s^ω is a convex combination of the σ_s^j and $\sum_{j=1}^n \widehat{\omega}_{j,s} \xi_s^j$ is a convex combination of the residual volatilities ξ_s^j with $\sum_{j=1}^n \widehat{\omega}_{j,t} \xi_t^j = 0$. As this last term tends to be very small, we will now compute the small noise expansion of the basket Call price around such small values of $\sum_{j=1}^n \widehat{\omega}_{j,s} \xi_s^j$. We first write

$$\begin{cases} dF_s^{\omega,\varepsilon} = F_s^{\omega,\varepsilon} \left(\sigma_s^\omega + \varepsilon \sum_{j=1}^n \widehat{\omega}_{j,s} \xi_s^j \right) dW_s \\ \\ d\widehat{\omega}_{i,s}^\varepsilon = \widehat{\omega}_{i,s}^\varepsilon \left(\xi_s^i - \varepsilon \sum_{j=1}^n \widehat{\omega}_{j,s} \xi_s^j \right) \left(dW_s + \sigma_s^\omega ds + \varepsilon \sum_{j=1}^n \widehat{\omega}_{j,s} \xi_s^j ds \right) \end{cases}$$

and develop around small values of $\varepsilon > 0$. From now on, we implicitly set

$$\varepsilon = \left\| \sum_{j=1}^n \widehat{\omega}_{j,s} \xi_s^j \right\|$$

and substitute $\widehat{\omega}_{j,s}/\varepsilon$ to $\widehat{\omega}_{j,s}$. As in Fournié et al. (1997) and Lebuchoux & Musiela (1999), we want to evaluate the following expectation:

$$C^\varepsilon = E\left[\left(F_T^{\omega,\varepsilon} - k \right)^+ | \left(F_t^\omega, \widehat{\omega}_t \right) \right]$$

and develop its Taylor series in ε around 0:

$$C^\varepsilon = C^0 + C^{(1)} \varepsilon + C^{(2)} \frac{\varepsilon^2}{2} + o(\varepsilon^2)$$

We can now get the order zero term as the classical basket approximation, which corresponds to that in Huynh (1994), Musiela & Rutkowski (1997) or Brace & Womersley (2000).

Proposition 20 *The first term C^0 is given by the Black & Scholes (1973) formula. In this simple approximation, the basket call price is given by:*

$$C^0 = BS(T, F_t^\omega, V_T) = F_t^\omega N(h(V_T)) - \kappa N \left(h(V_T) - \sqrt{V_T} \right) \tag{3.3}$$

where

$$h(V_T) = \frac{\left(\ln\left(\frac{F_t^\omega}{\kappa} \right) + \frac{1}{2} V_T \right)}{\sqrt{V_T}}$$

with the variance computed as:

$$V_T = \int_t^T \|\sigma_s^\omega\|^2 \, ds$$

which is again

$$V_T = \int_t^T Tr\left(\Omega_t \Gamma_s\right) ds \textit{ with } \Omega_t = \widehat{\omega}_t \widehat{\omega}_t^T$$

this later format will be useful in the calibration program design.

Proof. Because for $s \in [t,T]$ we have $\widehat{\omega}_{j,s} \geq 0$ with $\sum_{j=1}^n \widehat{\omega}_{j,s} = 1$, as in Lebuchoux & Musiela (1999) or Fouque et al. (2000) we can compute C^0 by solving the limiting P.D.E.:

$$\begin{cases} \frac{\partial C^0}{\partial s} + \|\sigma_s^\omega\|^2 \frac{x^2}{2} \frac{\partial^2 C^0}{\partial x^2} = 0 \\ C^0 = (x-K)^+ \text{ for } s = T \end{cases}$$

hence the above result. Finally

$$\begin{aligned} Tr\left(\Omega_t \Gamma_s\right) &= \sum_{i=1}^n \sum_{j=1}^n \widehat{\omega}_{i,t} \widehat{\omega}_{j,t} \left\langle \sigma_s^j, \sigma_s^i \right\rangle \\ &= \|\sigma_s^\omega\|^2 \end{aligned}$$

allows us to rewrite the variance as the inner product of $\Omega_t \, \Gamma_s$. ■

We have recovered the classical order zero approximation, we can now look for an extra term by solving for $C^{(1)}$.

Lemma 21 *Suppose that the underlying dynamics are described by (3.2). The first order term $C^{(1)}(s,x,y)$ can be computed by solving:*

$$\begin{aligned} 0 &= \frac{\partial C^{(1)}}{\partial s} + \|\sigma_s^\omega\|^2 \frac{x^2}{2} \frac{\partial^2 C^{(1)}}{\partial x^2} + \sum_{j=1}^n \left\langle \xi_s^j, \sigma_s^\omega \right\rangle x y_j \frac{\partial^2 C^{(1)}}{\partial x \partial y_j} \qquad (3.4) \\ &\quad + \sum_{j=1}^n \left\|\xi_s^j\right\|^2 \frac{y_j^2}{2} \frac{\partial^2 C^{(1)}}{\partial y_j^2} + \sum_{j=1}^n \left\langle \xi_s^j, \sigma_s^\omega \right\rangle y_j \frac{\partial C^{(1)}}{\partial y_j} + \sum_{j=1}^n \left\langle \xi_s^j, \sigma_s^\omega \right\rangle y_j x^2 \frac{\partial^2 C^0}{\partial x^2} \\ 0 &= C^{(1)} \textit{ for } s = T \end{aligned}$$

with $C^0 = BS(s,x,V_s)$ given by the Black & Scholes (1973) formula as in (3.3).

Proof. Let us first detail explicitly the P.D.E. followed by the price process. With the dynamics given by:

$$\begin{cases} dF_s^{\omega,\varepsilon} = F_s^{\omega,\varepsilon} \left(\sigma_s^\omega + \varepsilon \sum_{j=1}^n \widehat{\omega}_{j,s} \xi_s^j\right) dW_s \\ d\widehat{\omega}_{i,s}^\varepsilon = \widehat{\omega}_{i,s}^\varepsilon \left(\xi_s^i - \varepsilon \sum_{j=1}^n \widehat{\omega}_{j,s}^\varepsilon \xi_s^j\right) \left(dW_s + \sigma_s^\omega ds + \varepsilon \sum_{j=1}^n \widehat{\omega}_{j,s} \xi_s^j ds\right) \end{cases}$$

as in Karatzas & Shreve (1991) we get for

$$C^\varepsilon = E\left[\left(F_T^{\omega,\varepsilon} - k\right)^+ \mid \left(F_t^\omega, \widehat{\omega}_t\right)\right]$$

the corresponding P.D.E. :

$$\begin{cases} L_0^\varepsilon C^\varepsilon = 0 \\ C^\varepsilon = (x-k)^+ \text{ for } s = T \end{cases}$$

where L_0^ε is given by (with x and y_i associated to $F_s^{\omega,\varepsilon}$ and $\widehat{\omega}_{i,s}$ respectively):

$$\begin{aligned} L_0^\varepsilon = & \frac{\partial C^\varepsilon}{\partial s} + \left\| \sigma_s^\omega + \varepsilon \sum_{j=1}^n y_j \xi_s^j \right\|^2 \frac{x^2}{2} \frac{\partial^2 C^\varepsilon}{\partial x^2} \\ & + \sum_{j=1}^n \left(\left\langle \xi_s^j, \sigma_s^\omega \right\rangle + \varepsilon \sum_{k=1}^n y_k \left\langle \xi_s^j - \sigma_s^\omega, \xi_s^k \right\rangle - \varepsilon^2 \left\| \sum_{k=1}^n y_k \xi_s^k \right\|^2 \right) x y_j \frac{\partial^2 C^\varepsilon}{\partial x \partial y_j} \\ & + \sum_{j=1}^n \left\| \xi_s^j - \varepsilon \sum_{k=1}^n y_k \xi_s^k \right\|^2 \frac{y_j^2}{2} \frac{\partial^2 C^\varepsilon}{\partial y_j^2} \\ & + \sum_{j=1}^n \left(\left\langle \xi_s^j, \sigma_s^\omega \right\rangle + \varepsilon \sum_{k=1}^n y_k \left\langle \xi_s^j - \sigma_s^\omega, \xi_s^k \right\rangle - \varepsilon^2 \left\| \sum_{k=1}^n y_k \xi_s^k \right\|^2 \right) y_j \frac{\partial C^\varepsilon}{\partial y_j} \end{aligned}$$

as in Fournié et al. (1997) and Lebuchoux & Musiela (1999) we can differentiate this P.D.E. with respect to ε to get:

$$\begin{aligned} 0 &= L_0^\varepsilon C^{(1),\varepsilon} + \left(2 \sum_{j=1}^n y_j \left\langle \xi_s^j, \sigma_s^\omega \right\rangle + 2\varepsilon \left\| \sum_{k=1}^n y_k \xi_s^k \right\|^2 \right) \frac{x^2}{2} \frac{\partial^2 C^\varepsilon}{\partial x^2} \\ & \quad + \sum_{j=1}^n \left(\sum_{k=1}^n \left\langle \xi_s^j - \sigma_s^\omega, \xi_s^j \right\rangle - 2\varepsilon \left\| \sum_{k=1}^n y_k \xi_s^k \right\|^2 \right) x y_j \frac{\partial^2 C^\varepsilon}{\partial x \partial y_j} \\ & \quad + \sum_{j=1}^n \left(-2 \sum_{k=1}^n y_k \left\langle \xi_s^j, \xi_s^k \right\rangle + 2\varepsilon \left\| \sum_{k=1}^n y_k \xi_s^k \right\|^2 \right) \frac{y_j^2}{2} \frac{\partial^2 C^\varepsilon}{\partial y_j^2} \\ & \quad + \sum_{j=1}^n \left(\sum_{k=1}^n y_k \left\langle \xi_s^j - \sigma_s^\omega, \xi_s^k \right\rangle - 2\varepsilon \left\| \sum_{k=1}^n y_k \xi_s^k \right\|^2 \right) y_j \frac{\partial C^\varepsilon}{\partial y_j} \\ 0 &= C^{(1),\varepsilon} \text{ for } s = T \end{aligned}$$

and again as in Lebuchoux & Musiela (1999) or Fouque et al. (2000) we take the limit as $\varepsilon \to \infty$ and compute $C^{(1)}$ as the solution to:

$$\begin{cases} L_0^0 C^{(1)} + \left(\sum_{j=1}^n y_j \left\langle \xi_s^j, \sigma_s^\omega \right\rangle \right) x^2 \frac{\partial^2 C^0}{\partial x^2} = 0 \\ C^\varepsilon = 0 \text{ for } s = T \end{cases}$$

which is again, with $C^0 = BS(T, F_t^\omega, V_T)$ given by (3.3):

$$\begin{aligned} 0 &= \frac{\partial C^{(1)}}{\partial s} + \|\sigma_s^\omega\|^2 \frac{x^2}{2} \frac{\partial^2 C^{(1)}}{\partial x^2} + \sum_{j=1}^n \left\langle \xi_s^j, \sigma_s^\omega \right\rangle x y_j \frac{\partial^2 C^{(1)}}{\partial x \partial y_j} \\ & \quad + \sum_{j=1}^n \|\xi_s^j\|^2 \frac{y_j^2}{2} \frac{\partial^2 C^{(1)}}{\partial y_j^2} + \sum_{j=1}^n \left\langle \xi_s^j, \sigma_s^\omega \right\rangle y_j \frac{\partial C^{(1)}}{\partial y_j} + \sum_{j=1}^n \left\langle \xi_s^j, \sigma_s^\omega \right\rangle y_j x^2 \frac{\partial^2 C^0}{\partial x^2} \\ 0 &= C^{(1)} \text{ for } s = T \end{aligned}$$

which is the desired result. ■

We can now compute a closed-form solution to the equation verified by $C^{(1)}$ using its Feynman-Kac representation.

Proposition 22 *Suppose that the underlying dynamics are described by (3.2). The derivative* $C^{(1)}\left(t, F_t^\omega, (\widehat{\omega}_{j,t})_{j=1,\ldots,n}\right)$ *can be computed as:*

$$\begin{aligned} C^{(1)} &= F_t^\omega \int_t^T \sum_{j=1}^n \widehat{\omega}_{j,t} \frac{\left\langle \xi_s^j, \sigma_s^\omega \right\rangle}{\sqrt{V_{t,T}}} \exp\left(2\int_t^s \left\langle \xi_u^j, \sigma_u^\omega \right\rangle du\right) \\ & n\left(\frac{\ln \frac{F_t^\omega}{K} + \int_t^s \left\langle \xi_u^j, \sigma_u^\omega \right\rangle du + \frac{1}{2} V_{t,T}}{\sqrt{V_{t,T}}}\right) ds \end{aligned} \tag{3.5}$$

Proof. The limiting diffusions are given by:

$$\begin{aligned} F_s^{\omega,0} &= F_t^\omega \exp\left(\int_t^s \sigma_u^\omega dW_u - \frac{1}{2}\int_t^s \|\sigma_u^\omega\|^2 du\right) \\ \widehat{\omega}_{j,s}^0 &= \widehat{\omega}_{j,t} \exp\left(\int_t^s \tilde{\sigma}_u^j dW_u + \int_t^s \left(\left\langle \xi_u^j, \sigma_u^\omega \right\rangle - \frac{1}{2}\left\|\tilde{\sigma}_u^j\right\|^2\right) du\right) \end{aligned}$$

and because $C^{(1)}$ solves the P.D.E. (3.4) in the above lemma, with

$$\frac{\partial^2 C_s^0}{\partial x^2} = \frac{n(h(x, V_{s,T}))}{x\sqrt{V_{s,T}}}$$

where we have noted

$$n(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right)$$

we can write the Feynman-Kac representation of the solution to (3.4) with terminal condition zero as:

$$C^{(1)} = \int_t^T E\left[\sum_{j=1}^n \left\langle \xi_s^j, \sigma_s^\omega \right\rangle \widehat{\omega}_{j,s}^0 F_s^{\omega,0} \frac{n(h(V_{s,T}, F_s^{\omega,0}))}{\sqrt{V_{s,T}}}\right] ds$$

where

$$h(u,v) = \frac{\left(\ln\left(\frac{v}{\kappa}\right) + \frac{1}{2}u\right)}{\sqrt{u}} \text{ with } V_{s,T} = \int_s^T \|\sigma_u^\omega\|^2 du$$

Hence we can directly compute $C^{(1)}$ as:

$$\begin{aligned} C^{(1)} = F_t^\omega \int_t^T \sum_{j=1}^n \widehat{\omega}_{j,t} \left\langle \xi_s^j, \sigma_s^\omega \right\rangle \exp\left(\int_t^s -\frac{1}{2}\left\|\xi_u^j - \sigma_u^\omega\right\|^2 du\right) \\ E\left[\frac{\exp\left(\int_t^s \left(\sigma_u^\omega + \xi_u^j\right) dW_u\right)}{\sqrt{V_{s,T}}} n\left(\frac{\ln \frac{F_t^\omega}{K} + \int_t^s \sigma_u^\omega dW_u - \frac{1}{2}V_{t,s} + \frac{1}{2}V_{s,T}}{\sqrt{V_{s,T}}}\right)\right] ds \end{aligned}$$

which is, using the Cameron-Martin formula:

$$C^{(1)} = F_t^\omega \int_t^T \sum_{j=1}^n \widehat{\omega}_{j,t} \frac{\left\langle \xi_s^j, \sigma_s^\omega \right\rangle \exp\left(2\int_t^s \left\langle \xi_u^j, \sigma_u^\omega \right\rangle du\right)}{\sqrt{V_{s,T}}}$$
$$E\left[n\left(\frac{\ln\frac{F_t^\omega}{K} + \int_t^s \sigma_u^\omega dW_u + \int_t^s \left\langle \xi_u^j, \sigma_u^\omega \right\rangle du + \frac{1}{2}V_{t,T}}{\sqrt{V_{s,T}}} \right)\right] ds$$

and because for $g = N(a, b^2)$:

$$E[n(g)] = \frac{1}{\sqrt{b^2+1}} n\left(\frac{a}{\sqrt{b^2+1}}\right)$$

we get:

$$\begin{aligned} C^{(1)} &= F_t^\omega \int_t^T \sum_{j=1}^n \widehat{\omega}_{j,t} \frac{\left\langle \xi_s^j, \sigma_s^\omega \right\rangle}{\sqrt{(V_{t,s}+V_{s,T})}} \exp\left(2\int_t^s \left\langle \xi_u^j, \sigma_u^\omega \right\rangle du\right) \\ & n\left(\frac{\ln\frac{F_t^\omega}{K} + \int_t^s \left\langle \xi_u^j, \sigma_u^\omega \right\rangle du + \frac{1}{2}V_{t,T}}{\sqrt{(V_{t,s}+V_{s,T})}} \right) ds \end{aligned}$$

which is the desired result. ■

We will show below that this result can be interpreted as a correction accounting for the mis-specification of the volatility induced.

3.1.3 Robustness interpretation

The basket dynamics are essentially that of an almost lognormal process with a mildly stochastic volatility. By approximating these dynamics with a true lognormal process, we will make a small ”tracking error” in the computation of the replicating portfolio by computing the delta using an incorrect specification of the volatility. As in El Karoui et al. (1998), we can compute this tracking error almost explicitly. Suppose that $\Pi_{\sigma_s^\omega,s}$ is the value at time s of a self-financing delta hedging portfolio computed using the approximate volatility σ_s^ω. As the volatility in this delta computation is only approximately equal to the volatility driving the underlying assets, there will be a small hedging tracking error e_s computed as:

$$e_s = P_{\sigma_s^\omega,s} - \Pi_{\sigma_s^\omega,s}$$

where $P_{\sigma_s^\omega,s}$ is the price of the option at time s, computed using the approximate volatility σ_s^ω. Of course, we know that $P_{\sigma_s^\omega,T} = (F_t^\omega - K)^+$ and we can understand $E[e_s]$ as a price correction accounting for the volatility misspecification. From El Karoui et al. (1998) we know that we can compute this (exact) tracking error explicitly as:

$$e_T = \frac{1}{2}\int_t^T \left(\left\| \sum_{i=1}^n \widehat{\omega}_{i,s}\sigma_s^i \right\|^2 - \left\|\sigma_s^\omega\right\|^2 \right) (F_s^\omega)^2 \frac{\partial^2 C^0(F_s^\omega, V_{t,T})}{\partial x^2} ds \tag{3.6}$$

From the computation of $C^{(1)}$ in the previous part we know:

$$C^{(1)} = \int_t^T E\left[\sum_{j=1}^n \left\langle \xi_s^j, \sigma_s^\omega \right\rangle \widehat{\omega}_{j,s} F_s^\omega \frac{n(h(V_{s,T}, F_s^\omega))}{\sqrt{V_{s,T}}}\right] ds$$

With $\sigma_s^i = \sigma_s^\omega + \xi_s^i$, and because $\sum_{i=1}^n \widehat{\omega}_{i,s} = 1$, we rewrite (3.6) as:

$$\begin{aligned} e_T &= \int_t^T \left(\left\langle \sum_{i=1}^n \widehat{\omega}_{i,s}\sigma_s^i - \sigma_s^\omega, \sigma_s^\omega \right\rangle\right) (F_s^\omega)^2 \frac{\partial^2 C^0(F_s^\omega, V_{t,T})}{\partial x^2} ds \\ &\quad + \frac{1}{2}\int_t^T \left(\left\| \sum_{i=1}^n \widehat{\omega}_{i,s}\sigma_s^i - \sigma_s^\omega \right\|^2\right) (F_s^\omega)^2 \frac{\partial^2 C^0(F_s^\omega, V_{t,T})}{\partial x^2} ds \end{aligned}$$

or again:

$$\begin{aligned} e_T &= \int_t^T \left\langle \sum_{i=1}^n \widehat{\omega}_{i,s}\xi_s^i, \sigma_s^\omega \right\rangle (F_s^\omega)^2 \frac{\partial^2 C^0(F_s^\omega, V_{t,T})}{\partial x^2} ds \\ &\quad + \frac{1}{2}\int_t^T \left(\left\| \sum_{i=1}^n \widehat{\omega}_{i,s}\xi_s^i \right\|^2\right) (F_s^\omega)^2 \frac{\partial^2 C^0(F_s^\omega, V_{t,T})}{\partial x^2} ds \end{aligned}$$

The first order expansion of e_T for small values of ξ_s^i gives:

$$e_T^{(1)} = \int_t^T \left\langle \sum_{i=1}^n \widehat{\omega}_{i,s}\xi_s^i, \sigma_s^\omega \right\rangle (F_s^\omega)^2 \frac{\partial^2 C^0(F_s^\omega, V_{t,T})}{\partial x^2} ds$$

writing the value of the Gamma explicitly, we get:

$$e_T^{(1)} = \int_t^T \sum_{i=1}^n \left\langle \xi_s^i, \sigma_s^\omega \right\rangle \widehat{\omega}_{i,s} F_s^\omega \frac{n(h(V_{s,T}, F_s^\omega))}{\sqrt{V_{s,T}}} ds$$

and finally:

$$C^{(1)} = E\left[e_T^{(1)}\right] \tag{3.7}$$

This means that the first order correction in the basket price approximation can also be interpreted as the expected value of the first order tracking error approximation for small values of the residual volatility ξ_s^i. This validates the price approximation in terms of *both pricing and hedging performance*. To make the link with chapter one explicit, we now write the order zero approximation in the particular case of swaption pricing.

3.1.4 Swaption price approximation

If we go back to the particular swaption pricing problem developed in chapter one, the result above allows us to compute the price of a swaption today if for example in the Libor market model we approximate the forward swap rate by:

$$dswap(s, T, T_N) = \sum_{i=1}^N \widehat{\omega}_i(t)\gamma(s, T_i - s)dW_s^{LVL}$$

where we can compute the weights $\widehat{\omega}_i(t)$

$$\widehat{\omega}_i(t) = \frac{B(t, T_{i+1})\delta K(t, T_i)}{\sum_{i=i_T}^{N} B(t, T_{i+1})\delta K(t, T_i)}$$

which can be written:

$$\widehat{\omega}_i(t) = \omega_i(t)\frac{K(t, T_i)}{swap(t, T, T_N)} \tag{3.8}$$

as the product of the weights in the swap decomposition by the forward/swap ratio.

This means that we can price the corresponding swaption using the lognormal Black formula with the cumulative variance set to V_T. If we use a piecewise constant γ, this last formula has the advantage of computing the market variance of a particular instrument (cap or swaption) as a quadratic form on the γ function and the weights $\widehat{\omega}_i(t)$. Furthermore, as the weights $\widehat{\omega}_i(t)$ are not very volatile (this is especially true when the forward curve is close to flat), the quadratic form used in the variance computation is almost constant over time. We can now conclude this section on the first order approximation of swaption prices as a direct consequence of the development above.

Proposition 23 *Using the above approximations, the price of a payer swaption with maturity T and strike κ, written on a forward swap starting at T with maturity T_N is given at time $t \leq T$ by the Black formula plus a correction term:*

$$swaption_t = Level(t, T, T_N)\left(swap(t, T, T_N)N(h) - \kappa N(h - \sqrt{V_T})\right) + Level(t, T, T_N)C^{(1)} \tag{3.9}$$

with

$$h = \frac{\left(\ln\left(\frac{swap(t,T,T_N)}{\kappa}\right) + \frac{1}{2}V_T\right)}{\sqrt{V_T}}$$

where $swap(t, T, T_N)$ is the market value of the forward swap today and

$$V_T = \int_t^T \|\gamma^\omega(s)\|^2 \, ds$$

with

$$\hat{\omega}_i(t) = \omega_i(t)\frac{K(t, T_i)}{swap(t, T, T_N)} \text{ and } \gamma^\omega(s) = \sum_{i=1}^{N} \hat{\omega}_i(t)\gamma(s, T_i - s)$$

and

$$\begin{aligned} C^{(1)} \;\; = \;\; & \int_t^T \sum_{j=1}^{n} \widehat{\omega}_j(t) \frac{\langle \tilde{\gamma}(s, T_j - s), \gamma^\omega(s)\rangle}{\sqrt{V_{t,T}}} \exp\left(2\int_t^s \langle \tilde{\gamma}(s, T_j - s), \gamma^\omega(s)\rangle \, du\right) \\ & n\left(\frac{\ln \frac{Level(t,T,T_N)}{K} + \int_t^s \langle \tilde{\gamma}(s, T_j - s), \gamma^\omega(s)\rangle \, du + \frac{1}{2}V_{t,T}}{\sqrt{V_{t,T}}}\right) ds \end{aligned}$$

where

$$\tilde{\gamma}(s, T_i - s) = \gamma(s, T_i - s) - \gamma^\omega(s)$$

We will now study the practical precision of this approximation by comparing the price obtained using the formulas above with the price given by Monte-Carlo simulations in both the Libor Market model and in the generic multidimensional Black & Scholes (1973) model.

3.2 Approximation precision

To assess the practical performance of the lognormal swap rate approximation in the pricing of swaptions, we will compare the prices obtained for a large set of key liquid swaptions using Monte-Carlo simulation and the lognormal forward swap approximation. We begin by recalling the key characteristics of the Libor Market Model by Brace et al. (1997) and how the model is discretized for simulation purposes. We then compare the swaption prices obtained by simulation with those provided by the approximated swaption pricing formula above.

3.2.1 Discretization

We have used the classic Euler discretization scheme as developed for example in Sidenius (1998), a less traditional technique can also be found Glasserman & Zhao (2000). We discretize the $\delta-$forward process above over a time interval Δt as:

$$K(t+\Delta t,T)=K(t,T)\left(1+\gamma(t,T-t)\varepsilon\sqrt{\Delta t}+\gamma(t,T-t)\widetilde{\sigma}(t,T-t+\delta)dt\right)$$

where $\varepsilon\in\mathbb{R}^d$ is a vector of independent normalized Gaussian random variables. Now if T_i are the forward calendar dates, we divide each interval $[T_i,T_{i+1}]$ in n_r steps with:

$$t_{jn_r}=T_j \text{ and } \Delta t_i=t_{i+1}-t_i$$

hence, if we note $K_{i,j}=K(t_{i+1},T_j)$, we can compute the forward rate evolution as:

$$\frac{K_{i+1,j}}{K_{i,j}}=\left(1+\gamma(t_i,T_j-t_i)\varepsilon\sqrt{\Delta t_i}+\gamma(t_i,T_j-t_i)\widetilde{\sigma}(t_i,T_j-t_i+\delta)dt\right) \tag{3.10}$$

having set $\widetilde{\sigma}(t_i,u)=0$ for $u<\delta$ to maintain the absence of arbitrage property between Zero Coupon bonds and the Money Market account. The value of this account process is here computed directly from the forwards using:

$$\beta_T=\frac{\left(1+\delta K(T_{j_T},T_{j_T})^{(T-T_j)/\delta}\right)}{\prod_{j=0}^{j_T-1}(1+\delta K(T_j,T_j))}$$

to simplify the procedure.

3.2.2 Numerical results

In the first figure (3.1), we present a plot of the difference between two distinct sets of swaption prices in the Libor Market Model. One is obtained by Monte-Carlo simulation using enough steps to make the 95% confidence margin of error always less than 1bp. The second set of prices is computed using the order zero approximation formula above. We can notice that the absolute error is increasing in the underlying maturity of the swaption and that its sign is not constant. This plot is based on the prices obtained by calibrating the model to EURO swaption prices on November 6 2000. We have used all cap volatilities and the following swaptions: 2Y into 5Y, 5Y into 5Y, 5Y into 2Y, 10Y into 5Y, 7Y into 5Y, 10Y into 2Y, 10Y into 7Y, 2Y into 2Y, 1Y into 9Y (the motivation behind this choice of swaptions is liquidity, all swaptions in the 10Y diagonal or in 2Y, 5Y, 7Y, 10Y are supposed to be more liquid). The absolute error is always less than 4 bp which is very significantly lower than the Bid-Ask spreads.

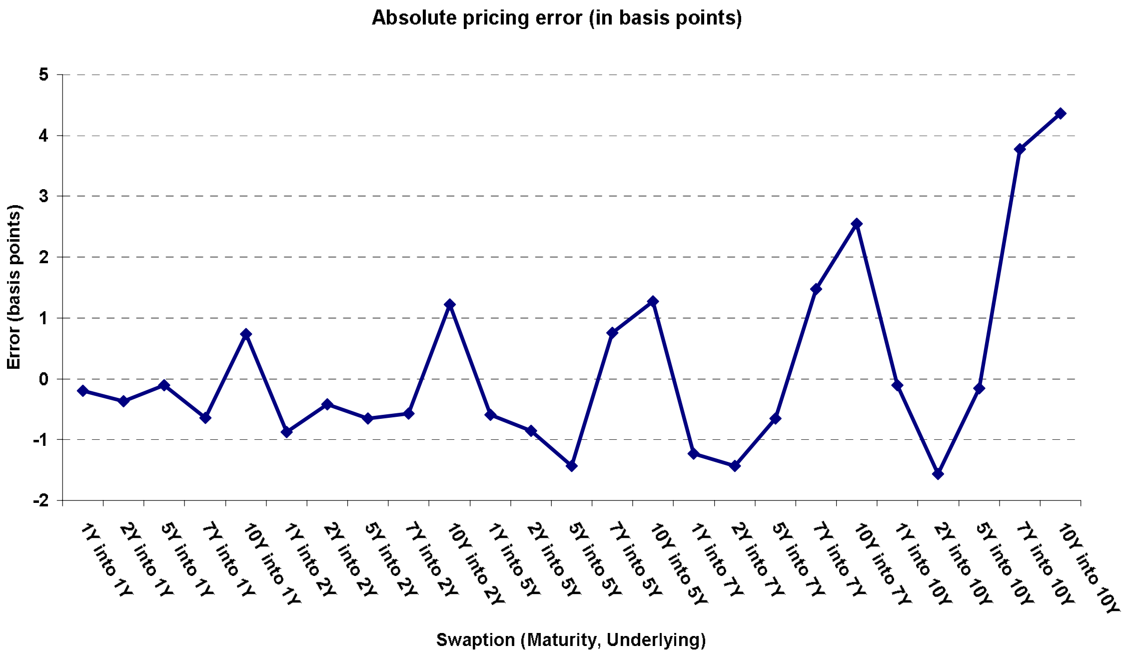

Figure 3.1: Absolute error in the order zero price approximation versus the Libor market model prices estimated using Monte-Carlo simulation, for various ATM Swaptions.

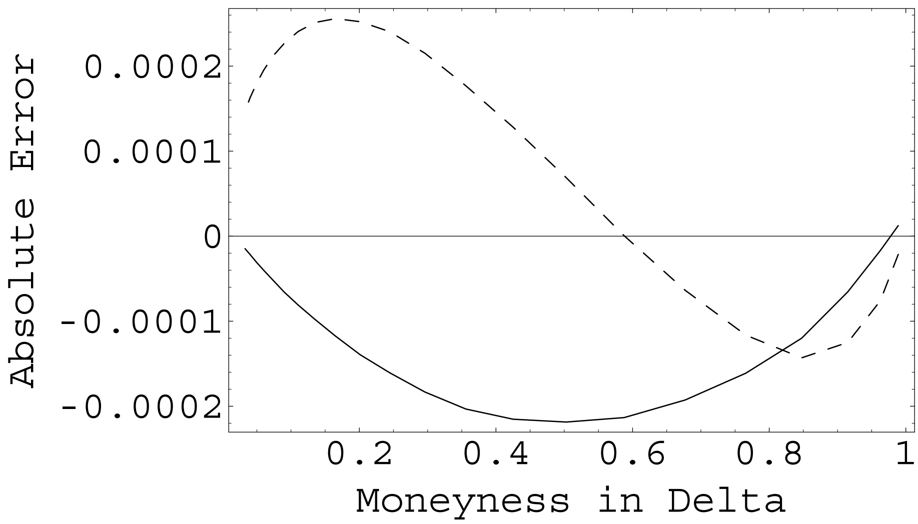

Figure 3.2: Order zero (dashed) and order one (plain) absolute approximation error versus the multidimensional Black-Scholes basket prices obtained by simulation for various strikes.

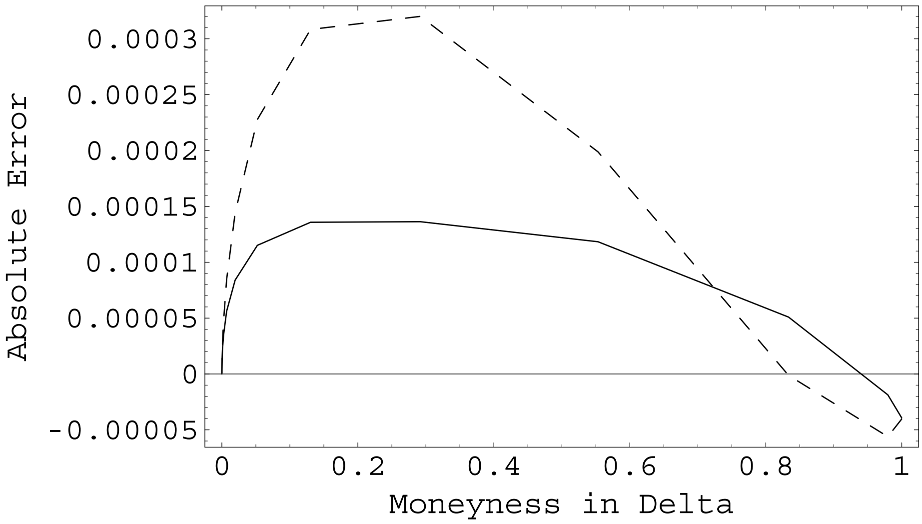

Figure 3.3: Order zero (dashed) and order one (plain) absolute approximation error versus the multidimensional Black-Scholes basket prices obtained by simulation for various strikes. (Diagonal covariance matrix.)

In the second figure (3.2), we plot the error in the basket pricing formula for a basket of assets, having supposed that the forwards are all martingale under the same probability measure (hence we test the precision of the approximations without the error from the forward measures). The reference is given by a Monte-Carlo estimate with 40000 steps.

The numerical values used here are $F_0^i = \{0.7, 0.5, 0.4, 0.4, 0.4\}$, $\omega_i = \{0.2, 0.2, 0.2, 0.2, 0.2\}$, $T = 5$ years, and the covariance matrix is given by:

$$\frac{11}{100}\begin{pmatrix} 0.64 & 0.59 & 0.32 & 0.12 & 0.06 \\ 0.59 & 1 & 0.67 & 0.28 & 0.13 \\ 0.32 & 0.67 & 0.64 & 0.29 & 0.14 \\ 0.12 & 0.28 & 0.29 & 0.36 & 0.11 \\ 0.06 & 0.13 & 0.14 & 0.11 & 0.16 \end{pmatrix}$$

This covariance comes from an historical estimate and has the typical level, spread, convexity eigenvector structure. These values are meant to replicate the pricing of a 5Y into 5Y swaption without the change in measure. We can see that the pricing error is less than 2bp with the order zero approx. and the additional order one term does not provide a significant benefit. In fact, the order zero term reaches an excellent precision near the money, a feature that is constantly observed when the covariance matrix has the structure given above, where the first level eigenvector accounts for around 90% of the volatility and the model is close to univariate (as noted in Brace et al. (1997)). However, we observe (in figure (3.3) below, for example) that the order one approximation does provide a significant precision improvement when the rates are less correlated or the underlying used is not significantly smaller than one (in equity basket options for example). More details and numerical examples are given at the end of this work.

Chapter 4

Market Model Calibration

In this part, we detail the calibration problem and its numerical characteristics. Let us start by observing one of the key features of the swaption pricing approximation formula obtained in the last chapter. We write the market variance as the scalar product of the forward rates covariance matrix and a matrix computed from market data on the swap weights:

$$\begin{aligned} V_T &= \int_t^T \left\| \sum_{i=1}^N \hat{\omega}_i(t)\gamma(s, T_i - s) \right\|^2 ds \\ &= \int_t^T \left(\sum_{i=1}^N \sum_{j=1}^N \hat{\omega}_i(t)\hat{\omega}_j(t) \left\langle \gamma(s, T_i - s), \gamma(s, T_j - s) \right\rangle \right) ds \\ &= \int_t^T \mathbf{Tr}\left(\Omega_t X_s\right) ds \end{aligned}$$

where $\Omega_t, X_t \in \mathbb{R}^{N\times N}, t \in [0, T]$ are positive semidefinite symmetric matrixes defined by:

$$\Omega_t = \hat{\omega}(t)\hat{\omega}(t)^T = \left(\hat{\omega}_i(t)\hat{\omega}_j(t)\right)_{i,j\in[1,N]} \succeq 0$$

and

$$X_s = \left(\langle \gamma(s, T_i - s), \gamma(s, T_j - s)\rangle\right)_{i,j\in[1,N]} \succeq 0$$

the covariance matrix of the forward rates (Gram matrix of the $\gamma(s, T_i - s)$ vectors).

This shows that the cumulative market variance of a particular swaption can be written as a *linear functional* of the forward rates covariance matrix. This fact will be the cornerstone of the calibration algorithm that follows.

The market quotes prices for caps and swaptions in terms of implied Black volatility. We can compute the market cumulative variance for the swaption of maturity T_k as $\sigma^2_{market,k}T_k$, the fundamental parameter in the model calibration. For simplicity, we discretize the volatility function yearly to make it piecewise constant. We are now ready to formulate the calibration problem in an efficient way.

Proposition 24 *The general calibration problem can written as an infinite-dimensional linear matrix inequality (L.M.I.) :*

$$\begin{array}{ll} \textit{Find} & X_s \\ \textit{s.t.} & \mathbf{Tr}\left(\Omega_t \left(\int_t^T X_s ds\right)\right) = \sigma^2_{market,k} T_k \ \ \textit{for } k = (1, ..., M) \\ & X_s \succeq 0 \end{array}$$

in the variable $X_s : \mathbb{R}_+ \to \mathbf{S}_n$, where we have set

$$\Omega_t = \hat{\omega}(t)\hat{\omega}(t)^\intercal = \left(\hat{\omega}_i(t)\hat{\omega}_j(t)\right)_{i,j\in[1,N]} \succeq 0$$

as quoted by the market today.

Proof. From the market option prices, we can get the implied cumulative variances $\sigma^2_{market,k}T_k$ for $k = (1, ..., M)$ by inverting the pricing formulas found in the last section (which are always strictly increasing in variance). With the key parameter being now the p.s.d. matrix valued $X_s : \mathbb{R}_+ \to \mathbf{S}_n$, from the development above the problem of matching the model price with the market price can be reduced to $\mathbf{Tr}\left(\Omega_t\left(\int_t^T X_s ds\right)\right) = \sigma^2_{market,k}T_k$ for $k = (1, ..., M)$. ■

Because the market variance constraints are linear with respect to the underlying variable X_s and the set of positive semidefinite matrixes is a convex cone, we find that the general calibration problem is convex and given a convex objective function, it has a unique global solution. We will further detail the advantages of this formulation later on.

Remark 25 *With $X_s = \left(\langle\gamma(s, T_i - s), \gamma(s, T_j - s)\rangle\right)_{i,j\in[1,N]}$, the volatilities $\gamma(s, T_j - s)$ are extracted by Cholesky decomposition.*

4.1 The calibration constraints

We will discuss the most general calibration problem implementation in the next part on risk-management. For simplicity now and to keep the focus on the problem geometry, we discretize with a δ frequency. We also make the common assumption that although the forward rates volatilities are not stationary, their instantaneous correlation is and hence the volatility function take a quasi-stationary form $\gamma(s, x) = \sigma(s)\eta(x)$ with σ and η such that $\sigma(s) = \sigma(\frac{1}{\delta}\lfloor\delta s\rfloor)$, $\eta(u) = \eta(\frac{1}{\delta}\lfloor\delta u\rfloor)$ and $\sigma(s) = \eta(s) = 0$ when $s \leq 0$. The expression of the market cumulative variance then becomes

$$V_T = \sum_{i=t}^{T} \delta\, \mathbf{Tr}\left(\Omega_t X_i\right)$$

and with Ω_t being quoted by the market today, calibrating the model to the swaptions $k = (1, ..., M)$ can be written as:

$$\begin{array}{ll} \text{Find} & X_i \\ \text{s.t.} & \sum_{i=t}^{T} \delta\, \mathbf{Tr}\left(\Omega_{t,k} X_i\right) = \sigma_k^2 T_k \text{ for } k = 1, ..., M \\ & X_i \succeq 0 \text{ for } i = 0, ..., T \end{array} \tag{4.1}$$

where $\Omega_{t,k}$ is the matrix computed as in (3.9) from the particular $\hat{\omega}_{i,k}(t)$ weights in the underlying swap of the swaption k. We can account for Bid-Ask spreads in the market data by relaxing the constraints as:

$$\begin{array}{ll} \text{Find} & X_i \\ \text{s.t.} & \sigma^2_{Bid,k}T_k \leq \sum_{i=t}^{T} \delta\, \mathbf{Tr}\left(\Omega_{t,k} X_i\right) \leq \sigma^2_{Ask,k}T_k \text{ for } k = 1, ..., M \\ & X_i \succeq 0 \text{ for } i = 0, ..., T \end{array} \tag{4.2}$$

Here again we have

$$X_i = \left(\sigma^2(s)\left\langle\eta(T_i - s), \eta(T_j - s)\right\rangle\right)_{i,j\in[1,N]} \succeq 0$$

keeping in mind that the vectors $\eta(T_i - s)$ creating this matrix "shift" from period to period.

Although there is substantial empirical evidence to contradict the stationarity assumption, the fact that it is the simplest model parametrization which is *coherent with a day-to-day recalibration procedure* makes it an interesting case study and a central reference. Furthermore, we will see in a later part that the non-stationary calibration program has the exact same format as the stationary one except that the matrixes have a block-diagonal structure instead of being dense. For this reason, we will keep the stationary case here as our central example and we refer the reader to the next part on risk-management for the nonstationary program implementation details. With $\gamma(s,x)$ discretized yearly and stationary, we note:

$$\gamma(s,x) = \gamma_{\lfloor x \rfloor}$$

and the model cap cumulative variance is given by:

$$\sigma^2_{market} T = \sum_{i=1}^{\lfloor T \rfloor} \gamma_i^2 \quad \text{for caps}$$

For the swaptions, using the simple approximated pricing formula computed above we get:

$$\sigma^2_{market} T_0 = \mathbf{Tr}(X \Omega_{T_0})$$

where Ω_{T_0} is the sum of the matrix Ω_k for $k = (1, .., . \lfloor T_0 \rfloor)$ with Ω_k the matrix with submatrix $\varphi\varphi^T$ starting at element (k,k) and all other blocks equal to zero. The formula for caps is then quite naturally seen as the scalar product of the matrix Γ with a matrix with 1 on the diagonal up to element $\lfloor T \rfloor$ and zero elsewhere. Hence the calibration problem can be seen as:

$$\begin{array}{ll} \text{find} & X \\ \text{s.t.} & \mathbf{Tr}(X \Omega_{T_k}) = \left(\sigma^2_{market}\right)_k T_k \\ & X \succeq 0 \end{array} \tag{4.3}$$

or again

$$\begin{array}{ll} \text{find} & X \\ \text{s.t.} & \sigma^2_{Bid,k} T_k \leq \mathbf{Tr}(X \Omega_{T_k}) \leq \sigma^2_{Ask,k} T_k \\ & X \succeq 0 \end{array} \tag{4.4}$$

accounting for spreads in the market data. The matrix X is found as a particular solution to this convex problem given a particular convex objective.

4.2 A convex problem

The calibration constraints above define the set of calibrated matrixes as the intersection of the semi definite cone with a polyhedra. As such *the set of market calibrated matrixes is convex*. The calibration problem then becomes that of finding the "best possible" matrix in this set and this is best done by minimizing a convex function (usually a linear functional) over this set. With a linear objective, the class of problems such as the one above are called a Semi Definite Programs (SDP).

In many ways, this represents a shift in paradigm from the usual calibration procedures. Let us describe the key differences between the non-convex procedures such as the factor parametrization based algorithms Rebonato (1999) and the SDP formulation to illustrate why the convex SDP can very significantly improve the stability of the calibration results. As pointed out by Rockafellar (1970), the great divide between "easy" and "hard" does not stand between linear and non-linear

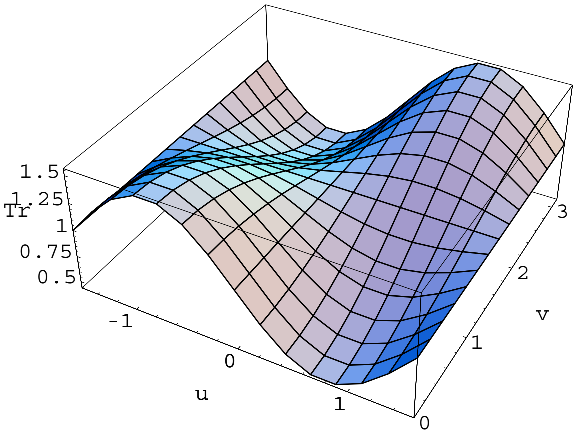

Figure 4.1: Plot of the objective function in the factor parametrized problem

but between convex and non-convex. The SDP program is minimizing a convex functional under linear constraints over a convex set, hence is convex whereas the rotation based methods is not, as we will see below. The consequence of this is that the SDP has an unique global optimum where the rotation algorithm has multiple local ones. Because of that, an insignificant daily variation in the market prices can make the solution of non-convex calibration algorithm shift from local optimum to local optimum provoking subsequent *"jumps" in the P&L and dangerous errors in the Greeks computation*. On the contrary, the same shock will only slightly shift the global optimum of the SDP program. Furthermore, as we will see in the next part, a lot can be said about the direction and amplitude of this shift, as for example in Todd & Yildirim (1999), so the *convex SDP calibration program formulation avoids the P&L variations that were only caused by numerical instability between local optima*.

4.2.1 Non convex example

Let us illustrate this on a simple example in $\mathbb{R}^{2\times 2}$. Suppose that using the factor parametrization based algorithm we are looking to solve the following simple calibration program:

$$\begin{array}{ll} \max & \mathbf{Tr}\left(\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} X\right) \\ \text{s.t.} & \mathbf{Tr}\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} X\right) = 1 \\ & X \succeq 0 \end{array}$$

The problem is then fully parametrized by the following definition of X:

$$X(u,v) = \begin{pmatrix} \cos^2(u) & \cos(v)\cos(u)\sin(u) \\ \cos(v)\cos(u)\sin(u) & \sin^2(u) \end{pmatrix}$$

In the figure (4.1) we plot the objective value as a function of (u, v) where we can see that the function obtained is non-convex and has two distant local maxima.

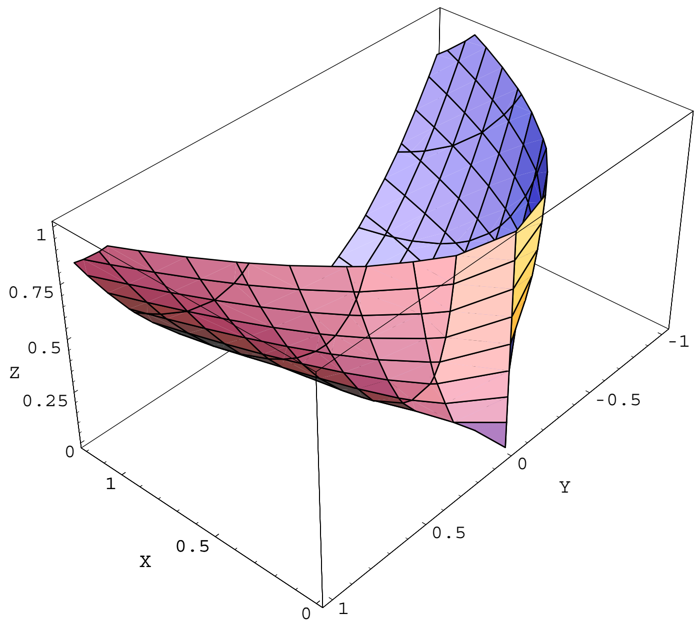

Figure 4.2: The semidefinite cone in dimension 3.

4.2.2 Convex example

The equivalent SDP algorithm on the other hand is optimizing a linear functional under linear constraints within a convex set.

Although the boundaries of this set are defined by a non-linear function, namely:

$$X = \begin{pmatrix} x & y \\ y & z \end{pmatrix} \succeq 0 \Longleftrightarrow \min_i \lambda_i(X) \geq 0$$

where $\lambda_i(X)$ are the eigenvalues of X, the overall problem can be solved very efficiently (see Nesterov & Nemirovskii (1994), Vandenberghe & Boyd (1996)) because both the feasible set and the objective are convex and a barrier function with excellent numerical properties can be computed. This frontier is plotted in figure (4.2) on the (x, y, z) space, together with figure (4.3) showing a cut of the cone, orthogonal to the $(-.15, .1, 1.2)$ direction.

Of course, all these examples are limited by the necessity to obtain a three-dimensional representation and figure (4.2) can also be seen as a rotated second-order (or Lorentz) cone. Another result of the recent progress in convex optimization is that the calibration above can be solved in polynomial time (usually less than a second for a typical problem size), with *an upper bound on the absolute error* (distance to the optimum and constraints). Furthermore, the practical complexity of those methods is also very well understood. In all the numerical tests that follow, the computing time on a standard 500 MHz workstation was close to a second on average.

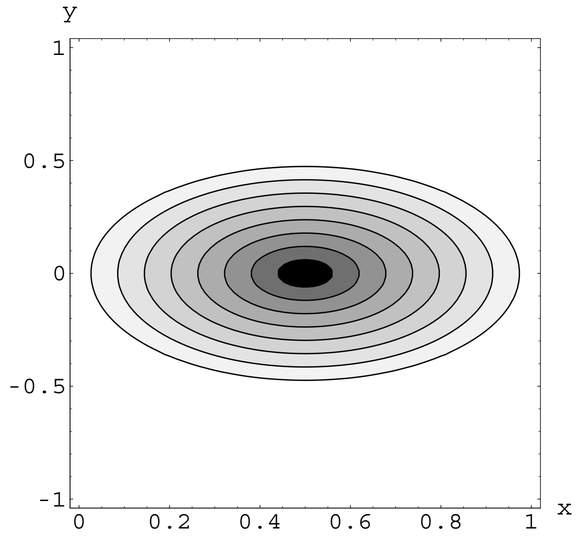

Figure 4.3: Plot of a typical feasible domain, a planar cut of the semidefinite cone.

4.3 Algorithm Implementation

The general form of the problem to be solved is given by:

$$\begin{array}{ll} \min & \mathbf{Tr}(CX) \\ \text{s.t.} & \mathbf{Tr}(X\Omega_{T_i}) = \sigma^2_{market,k} T_k, k = 1, ..., M \\ & X \succeq 0 \end{array}$$

where X is a block-matrix. For a general overview of Semi-definite programming algorithms see Vandenberghe & Boyd (1996), Nesterov & Nemirovskii (1994) or Alizadeh, Haeberly & Overton (1998). The algorithm we used in most of the numerical results that follow can be defined as HMPCAHO, which stands for the Potra-Sheng homogeneous formulation of the Mehrotra type predictor-corrector algorithm using the Alizadeh, Haeberly, Overton search direction Alizadeh et al. (1998). *The homogeneous formulation Potra & Sheng (1995) allows certain detection of primal or dual infeasibility which in our case corresponds to market prices incompatible with a stationary positive semi-definite market covariance matrix for the forward Libors.* The MPCAHO algorithm is the slowest of all formulations but it's reputation is to be the most robust. For a typical problem dimension of 20 (discretized yearly and stationary), the average number of steps is around 15 using an absolute precision goal on the market variance of 10^{-8}. The calibration time is less than a second for a full size implicit market covariance calibration. We followed the implementation structure given in Toh, Todd & Tutuncu (1996), having adapted in C the Mathematica algorithm by Brixius, Potra & Sheng (1996). Some more recent libraries including a more efficient formulation of the SOCP (quadratic: smoothness, euclidean distance ...) and L.P. constraints are becoming available. These include the MATLAB package by Sturm (1999) for symmetric cone programming, which we have extensively used here as well.

4.3.1 Simple objective

The problem with the above methods of smoothing and distance minimization is that they multiply the dimension of the problem, thus slowing down the process. We have found that in practice, setting directly the objective matrix C to a target covariance matrix provided good results. Intuitively, this is explained by the fact that a linear objective will force the solution to be a vertex of the feasible set, i.e. a p.s.d. matrix of low rank (see Fazel et al. (2000) for details). We can also understand this by remarking that the more constrained problem:

$$\begin{array}{ll} \text{max} & \mathbf{Tr}\,(CX) \\ \text{s.t.} & \mathbf{Tr}\,X = n \\ & X \succeq 0 \\ & [X, C] = 0 \end{array}$$

is in fact a simple L.P. in the eigenvalues because C and X can be simultaneously diagonalized. It is explicitly solved by a rank one matrix $e^* \left(e^*\right)^T$ where e^* is the eigenvector associated with the largest eigenvalue of C as an application of the Perron-Frobenius theorem. This shows why this kind of objective tends to lead to low rank matrixes that are close to the target covariance main factor, it has proved to be very efficient in practice.

4.3.2 Applications

In general, the calibration problem gives an entire set of solutions. The choice of the objective matrix C is a function of the ultimate objective of the calibration. We consider different choices of C:

Bounds on other swaptions

One of the most simple choices of objective matrix C is to set it to another swaptions associated matrix Ω_{T_i}. The calibration problem finds the parameters for the Market Model that gives either a minimum or a maximum arbitrage-free price (within the BGM framework) to the considered swaption while matching a certain set of market prices on other caps and swaptions. We will show some examples of this in the next section.

Spread Options. Baskets

The swaption pricing approximation given above can be seen as a pricing formula for baskets and we can set ω to reflect the difference of two forwards $\omega = (0, ..., 0, -a, b, 0, ..., 0)$, the difference of two swap rates $\omega = a\omega^1 - b\omega^2$ or in fact any basket of swaps and forwards (if we look at the shifted price). The calibration program using an objective matrix C computed from the above weights would then produce bounds on the price of a certain Spread Option as implied by the current market price of the calibration instruments.

Spectral distance to a target covariance

Let A be a target covariance matrix (for example, a previous calibration result or an historical estimate). The key point here is the choice of the matrix distance to be used. The norm that is most naturally adapted to semi definite programming settings is the spectral norm, i.e. with $\lambda_k(X)$ the eigenvalues of X.

$$\|X\| = \max_k |\lambda_k(X)|$$

This norm $\|X\|$ can also be computed using Weil's characterization of eigenvalues:

$$\max_k |\lambda_k(X)| = \sup_{\|u\|=1} \|Xu\|$$

with $\|.\|_2$ the euclidean norm.

Lemma 26 *The solution of the calibration problem that is* $\|.\|$ *closest to* A *can then be computed by solving:*

$$\begin{array}{ll} \textit{min} & t \\ \textit{s.t.} & \mathbf{Tr}(X\Omega_{T_i}) = \left(\sigma^2_{market}\right)_i T_i \\ & X - A \preceq tId \\ & X - A \succeq -tId \\ & X \succeq 0 \textit{ and } t \geq 0 \end{array}$$

Proof. As in Vandenberghe & Boyd (1996). We know that $X - A \preceq tId \Longleftrightarrow \max_k \lambda_k(X - A) \leq t$, hence in the above problem we have $\max_k |\lambda_k(X - A)| = \|X - A\| \leq t$. ■

The result given by this kind of norm minimization is then "vectorwise uniformly" close to the targeted covariance matrix.

In general however, one would also like to use the euclidean norm on the set of matrixes:

$$\|X\| = \sqrt{\mathbf{Tr}(X^2)}$$

This is a quadratic function of the underlying matrix and is not immediately a linear objective. We can however design a semidefinite program to solve for this minimization as in Nesterov & Nemirovskii (1994) §6.4.2 or use a symmetric cone formulation as in Nesterov & Todd (1997) or Sturm (1999).

4.4 Multiobjective calibration problem

Here, we extend the general calibration problem to incorporate the minimum rank and smoothness requirements. Those two necessary features of the calibrated matrix arise from very different problems. The rank minimization is a purely technical issue and is a consequence of the limitations of the pricing algorithms. Because most of the time the calibrated matrix will be used to generate paths in a Monte-Carlo simulation, a low rank matrix is necessary there to maintain acceptable computing times. A low rank solution is also critical for the pricing of products with American features where only trees and dynamic programming can be used. In the later case, the rank of the matrix must be kept to at most one or two for computations to be possible. The smoothness constraint is of course not as critical, it only reflects the belief that a market operator's pricing of the variance at one point in the covariance matrix will not be radically different from its input at an adjacent point in the matrix. The smoothness maximization is intended to produce realistic matrixes in that sense.

We then study how it is possible to combine those objectives inside a general calibration problem an we examine the resulting trade-offs. In particular, we show that it is possible to find a solution of rank two to a calibration problem based on actual market data, but that this rank result quickly deteriorates with the addition of a smoothness constraint. We will see that the same result holds for robustness.

4.4.1 Rank Minimization

Because the calibrated model will be used to compute prices of other derivatives using mostly Monte-Carlo techniques or trees, it is highly desirable to get a low rank solution. In general, the matrix solution to the calibration problem will lie on a vertex of the semidefinite cone and hence will be singular but there is no guarantee that the rank will remain below a certain level. The general problem of finding the calibrated matrix of lowest rank can be identified as part of a wider class of combinatorial problems and is hard. In practice, as we will see below, the market constraints always tend to produce results with very rapidly decreasing eigenvalues and some very good convex heuristics can enhance this feature significantly (see for example Fazel et al. (2000)) hence it is always possible to find a good approximated solution of rank two or three. We will see that in practice and in accordance with prior empirical studies, all solutions (even those with a high rank) tend to have only one or two dominant eigenvalues with the rest of the spectrum several orders of magnitude smaller, hence the rank issue is only fueled by the numerical limitations of derivative pricing techniques.

The trace heuristic

In the recent years, a lot of work has been focused on the problem of minimizing the rank of a matrix over a convex set (see Fazel et al. (2000) for a review of recent works). In general, as showed in Davis (1994) or Vandenberghe & Boyd (1996), this problem is NP-Hard, i.e. the key difference with all the other programs detailed in this paper is that there is provably (if $P \neq NP$) no polynomial-time algorithm to find a global solution to the minimum rank problem. In this case however, a lot of very efficient heuristical methods have been developed and we will use here the trace minimization technique detailed in Fazel et al. (2000). The particular instance of this class of problems we focus on here is the minimum rank calibration problem, which can be stated as:

$$\begin{array}{ll} \text{minimize} & \mathbf{rank}(X) \\ \text{subject to} & \sigma^2_{Bid,k}T_k \leq \mathbf{Tr}(\Omega_k X) \leq \sigma^2_{Ask,k}T_k \text{ for } k = 1,...,M \\ & X \succeq 0 \end{array} \quad (4.5)$$

in the variable $X \in \mathbf{S}^n$ with parameters $\Omega_k, C \in \mathbf{S}^n$ and $\sigma^2_{Bid,k}T_k$, $\sigma^2_{Ask,k}T_k \in \mathbb{R}_+$, for $k = 1,...,M$. A very common heuristic for solving the above problem is to substitute to the rank function the scalar product of the matrix X with another matrix $C \succeq 0$:

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}(CX) \\ \text{subject to} & \sigma^2_{Bid,k}T_k \leq \mathbf{Tr}(\Omega_k X) \leq \sigma^2_{Ask,k}T_k \text{ for } k = 1,...,M \\ & X \succeq 0 \end{array} \quad (4.6)$$

As pointed out by Mesbahi & Papavassilopoulos (1997), for some specific feasible set geometries, this produces exact solutions. In the general case, Fazel et al. (2000) show that this heuristic can be seen as the first iterate of an algorithm to find some local minimum of the logarithm of the determinant of X, which always lies at a low rank solution. This method finds a local optimum to the minimum rank problem and in practice, requires a few iterations. All the algorithm steps are trace minimization problems such as the one in (4.6) with the first one usually producing a low rank solution and the next iterations improving it by one or two zero eigenvalues.

Truncating the solution

The heuristic methods produce low rank solutions to the calibration problem but it is often the case that the rank of the solution can be decreased a little bit more by setting to zero all the eigenvalues of the solution that are below a preset precision level. This is a numerically very simple step, the only difficulty being to ensure that the solution remains feasible. As we show below, using the fact that all the matrixes Ω_k are p.s.d., this can in fact be done optimally, i.e. we can maximize the value below which the solution can be truncated and still remain feasible.

Proposition 27 *The semidefinite program:*

$$\begin{array}{ll} \textit{max} & \alpha \\ \textit{subject to} & \sigma^2_{Bid,k}T_k + \alpha\,\mathbf{Tr}\,(\Omega_k) \leq \mathbf{Tr}(\Omega_k X) \leq \sigma^2_{Ask,k}T_k \textit{ for } k = 1,...,M \\ & X \succeq 0 \end{array}$$

in the variables $X \in \mathbf{S}^n$ and $\alpha \in \mathbb{R}$, with parameters $\Omega_k, C \in \mathbf{S}^n$ and $\sigma^2_{Bid,k}T_k$, $\sigma^2_{Ask,k}T_k \in \mathbb{R}_+$, has a unique optimal solution (X,α) with α the greatest value such that X_α remains feasible for the original calibration problem:

$$\begin{array}{ll} \textit{find} & X \\ \textit{subject to} & \sigma^2_{Bid,k}T_k \leq \mathbf{Tr}(\Omega_k X) \leq \sigma^2_{Ask,k}T_k \textit{ for } k = 1,...,M \\ & X \succeq 0 \end{array}$$

where X_α is the truncated matrix obtained by setting to zero all the eigenvalues of the matrix X inferior to α.

Proof. For $A \in S_n, A \succeq 0$, let us first recall the following classical result:

$$\beta\,\mathbf{Tr}(A) = \begin{array}{ll} \text{max} & \mathbf{Tr}(XA) \\ \text{subject to} & \|X\| \leq \beta \\ & X \succeq 0 \end{array}$$

where $\|X\|$ is the spectral norm (maximum eigenvalue) of the variable $X \in S_n$. Let (X^*, α^*) be the solution to the above problem, we can now write the truncated matrix $X^*_{\alpha^*} = X - \Delta X$ with $\Delta X \succeq 0$ and $\|\Delta X\| \leq \alpha^*$. Now because $\mathbf{Tr}(\Omega_k X) \leq \sigma^2_{Ask,k}T_k$ for $k = 1,...,M$ and $\Delta X \succeq 0$ we know that $\mathbf{Tr}(\Omega_k X^*_{\alpha^*}) \leq \sigma^2_{Ask,k}T_k$ for $k = 1,...,M$. Now, using the result above, with $\mathbf{Tr}(\Omega_k \Delta X) \leq \alpha\,\mathbf{Tr}\,(\Omega_k)$ for $k = 1,...,M$, we know that $\sigma^2_{Bid,k}T_k \leq \mathbf{Tr}(\Omega_k X^*_{\alpha^*})$ for $k = 1,...,M$ and the matrix $X^*_{\alpha^*}$ is feasible for the general calibration problem. ■

We can combine the two methods above to get a low rank solution to the calibration problem that can be truncated and is guaranteed to remain feasible. This is solved by the program

$$\begin{array}{ll} \text{min} & -\alpha + \gamma\,\mathbf{Tr}(CX) \\ \text{subject to} & \sigma^2_{Bid,k}T_k + \alpha\,\mathbf{Tr}\,(\Omega_k) \leq \mathbf{Tr}(\Omega_k X) \leq \sigma^2_{Ask,k}T_k, \quad k = 1,...,M \\ & X \succeq 0 \end{array}$$

where the weight γ is an input and is set to optimally balance the two programs. On the same Nov. 6 2000 data set, the above heuristics allow us to find a rank two solution of the calibration problem. The calibration set is again formed by using all caps and the following swaptions: 2Y into 5Y, 5Y into 5Y, 5Y into 2Y, 10Y into 5Y, 7Y into 5Y, 10Y into 2Y, 10Y into 7Y, 2Y into 2Y, 1Y into 9Y. The calibrated forward rates covariance matrix we obtain is showed in the first figure.

The second figure shows its ten biggest eigenvalues in a semilog graph.

The above results above clearly underline the need to introduce smoothness constraints in the calibration procedure. This is detailed in what follows.

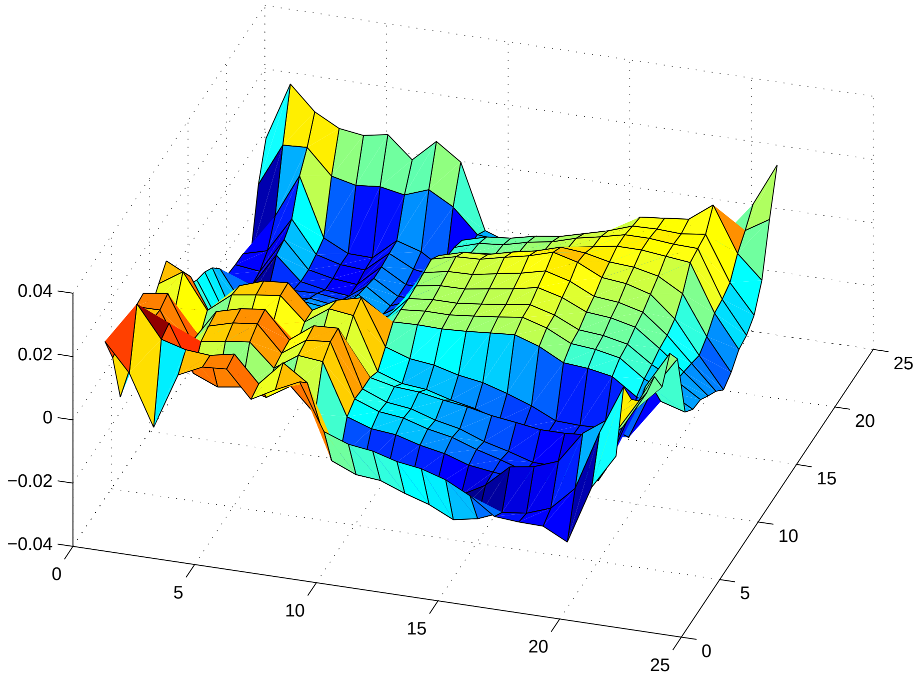

Figure 4.4: Truncated low rank solution matrix.

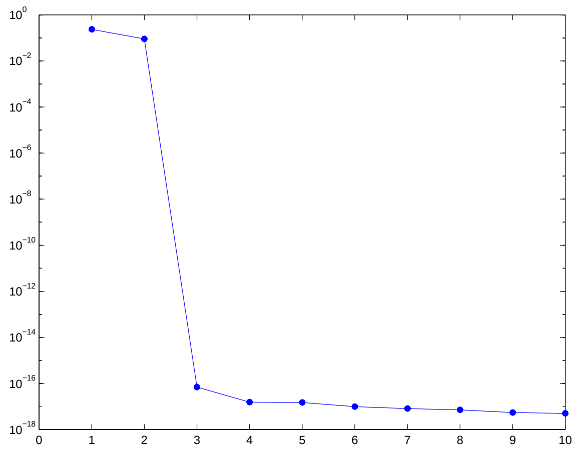

Figure 4.5: Semilog plot of the truncated solution matrix eigenvalues.

4.4.2 Smoothness constraints

As discussed above, it is sometimes desirable to impose smoothness objectives on the calibration problem to reflect the fact that market operators will tend to price similarly the variance of two products with close characteristics. A common way of smoothing the solution is to minimize the surface of the covariance matrix that we approximate here by:

$$S = \sum_{i,j\in[2,n]} \|\Delta_{i,j}X\|^2$$

where

$$\Delta_{i,j}X = \begin{pmatrix} X_{i,j} - X_{i-1,j} \\ X_{i,j} - X_{i,j-1} \end{pmatrix}$$

The calibration program then becomes:

$$\begin{array}{ll} \min & t \\ \text{subject to} & \sum_{i,j\in[2,n]} \|\Delta_{i,j}X\|^2 \leq t \\ & \sigma^2_{Bid,k}T_k \leq \mathbf{Tr}(\Omega_k X) \leq \sigma^2_{Ask,k}T_k, \quad k = 1, ..., M \\ & X \succeq 0 \end{array}$$

in the variables $X \in \mathbf{S}^{\max_i N_i}$ and $t \in \mathbb{R}$, with parameters $\Omega_k, C \in \mathbf{S}^{\max_i N_i}$ and $\sigma^2_{Bid,k}T_k, \sigma^2_{Ask,k}T_k \in \mathbb{R}_+$, which is a symmetric cone program as defined by Nesterov & Todd (1997), among others. These programs can be solved very efficiently by the same numerical methods used to solve semidefinite programs. Some very efficient numerical packages such as the one developed by Sturm (1999) are already available.

4.4.3 The robustness versus low rank trade-off

By combining all the different program requirements above, namely the robustness, smoothness and low rank objectives, we can form a general calibration program as follows:

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}(CX) + \alpha t_1 - \beta t_2 + \gamma t_3 \\ \text{subject to} & \sum_{i,j\in[2,n]} \|\Delta_{i,j}X\|^2 \leq t_1 \\ & \mathbf{Tr}(\Omega_k X) \geq \sigma^2_{Bid,k}T_k + t_2\, \mathbf{Tr}\,(\Omega_k) + t_3 \\ & \mathbf{Tr}(\Omega_k X) \leq \sigma^2_{Ask,k}T_k - t_3, \quad k = 1, ..., M \\ & X \succeq 0 \end{array}$$

which is a symmetric cone program in the variables $X \in \mathbf{S}^n$ and $t_1, t_2, t_3 \in \mathbb{R}$, with parameters $\Omega_k, C \in \mathbf{S}^n$ and $\sigma^2_{Bid,k}T_k, \sigma^2_{Ask,k}T_k \in \mathbb{R}_+$, where we can determine the relative importance of each objective by adjusting the weights $(\alpha, \beta, \gamma) \in \mathbb{R}^3_+$. This last form highlights the trade-offs that exists between the various calibration objectives detailed above. Perhaps the most important one is the trade-off between the rank minimization and the robustness.

This shows how much the rank limitations in the pricing techniques impact the stability of the calibration solution. As these limitations force us to look for a fixed rank solution (in fact, a solution of rank at most two), the calibration problem becomes NP-Hard and we have to choose between two radically different approximate numerical approaches. One is to use the traditional low rank parameterizations of the covariance matrix such as the one detailed in Rebonato (1999). As discussed in the previous section, these approaches are non-convex, which means that there is no guarantee that a global solution can be found in polynomial time and that the solution, if

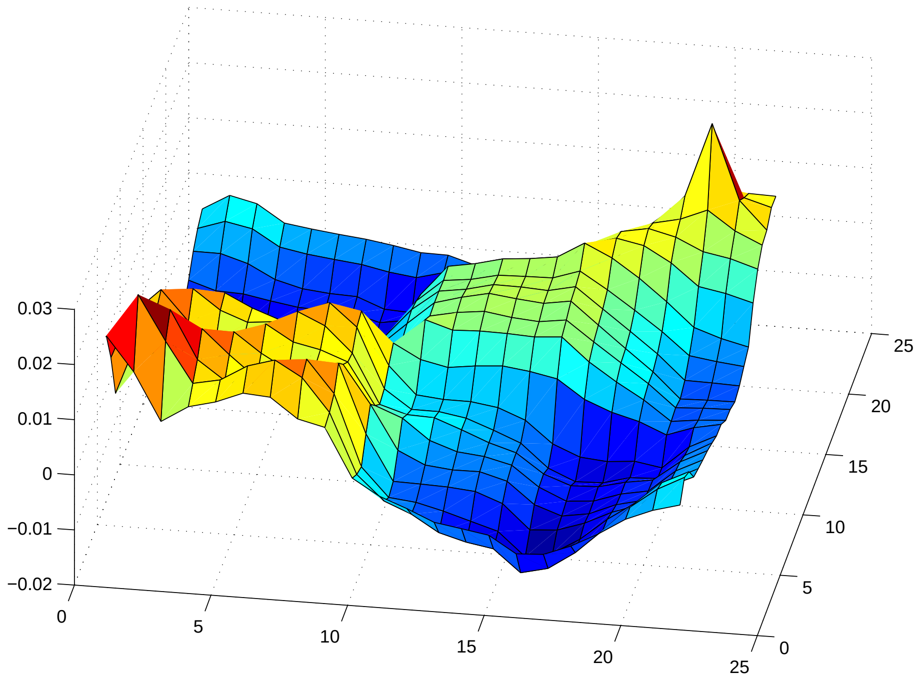

Figure 4.6: Solution to the calibration problem with smoothness constraints

found, can be extremely unstable. On the other hand, we can use the heuristics detailed here which can be seen as convex relaxations of the low rank calibration problem. They are guaranteed to provide a low rank solution, together with sensitivity results in polynomial-time, however, there is no guarantee that a particular rank objective can be reached. It is also important to keep in mind that because the fixed rank calibration problem is NP-Hard, there is provably no algorithm with reasonable (polynomial) complexity that can solve it globally. It appears then that the only way to stabilize the calibration procedure is to find pricing methods that can accommodate covariance matrixes with a rank bigger than simply one or two. Monte-Carlo methods provide an answer to this problem for the pricing of European type derivatives and the remaining fundamental difficulty lies in pricing American derivatives. Recent advances in American Monte-Carlo methods (see Longstaff & Schwartz (1998) among others) and quantization algorithms (see Bally & Pages (2000)) where it is possible to evaluate options in with a model dimension closer to the one we obtain in the stable solutions.

Below we show the impact of the smoothness constraints on the minimum rank heuristic results. The corresponding forward rates covariance matrix is plotted in the first figure.

We have significantly improved the matrix smoothness. The price to pay for this regularity is an increase in the rank of the solution as can be seen in the second figure We observe that the solution has a much higher rank than the one obtained with the rank minimization heuristic alone. This illustrates the central trade-off curve along which the calibration is performed: smoothness versus rank.

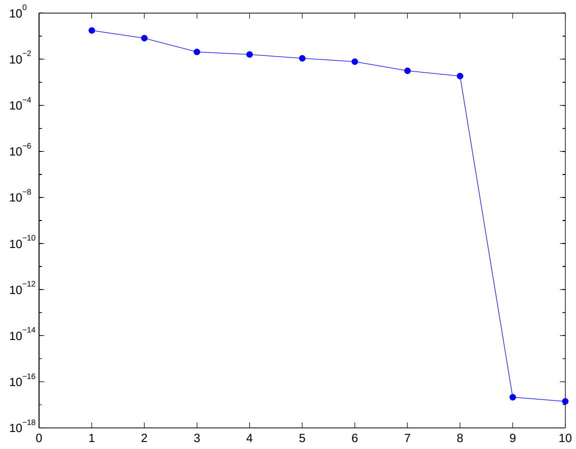

Figure 4.7: Semilog plot of the smooth matrix solution eigenvalues.

4.4.4 Calibration stabilization: a Tikhonov regularization

Along the lines of Cont (2001), we can explore more rigorously the impact of the smoothness constraints introduced above. Intuitively, one can look at the smoothness criterion as the minimization of a quadratic entropy, hence we can expect it to stabilize the calibration solution trough the minimization of the mutual information with a (flat) prior. As suggested by Cont (2001), we can think of the calibration as an ill-posed inverse problem and write the smooth calibration program as a Tikhonov (1963) regularization of the original problem. We are trying to solve the calibration problem:

$$A(X) = \left(\sigma_k^2 T_k\right)_{k=1,\ldots,M} \tag{4.7}$$

where

$$\begin{aligned} A : \mathbf{S}^N &\longrightarrow \mathbb{R}^M \\ X &\longmapsto AX := (\mathbf{Tr}\,(A_i X))_{i=1,\ldots,M} \end{aligned}$$

with the additional constraint that X be semidefinite positive. We can replace this hard constraint with a relaxed one together with an additional constraint on the norm of X and solve:

$$\text{minimize } \left\| A(X) - \left(\sigma_k^2 T_k\right) \right\|^2 + \alpha \left\| X \right\|^2$$

which, if we omit the positivity constraint, has a solution computed as:

$$X = (A^* A + \alpha I)^{-1} A^* b$$

where

$$\begin{aligned} A^* : \mathbb{R}^M &\longrightarrow \mathbf{S}^N \\ y &\longmapsto A^* y := \sum_{i=1}^{M} y_i \Omega_i \end{aligned}$$

and because all the eigenvalues of the operator $A^* A + \alpha I$ are greater than α, the condition number of the problem is improved and we can expect the sensitivity of the calibrated solution to be significantly reduced. Hence we see that the introduction of a smoothness constraint directly improves the stability of the calibration problem.

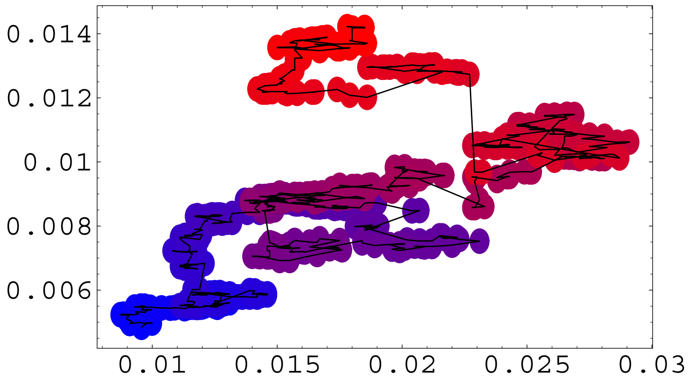

Figure 4.8: Historical evolution of the empirical volatility versus rate ratio over time (Red is the past and blue the present, JPY 1Y into 2Y forward Swap rate in 1998).

4.4.5 A word on stochastic volatility

All the calibration programs detailed up to now have at least one significant drawback: the model used don't allow any degree of freedom in the smile modelling. Is it possible to keep all the numerical advantages of the convex programming calibration techniques above while allowing some flexibility in the smile shape description? The answer here is a partial yes. First, we can remember that as the Gaussian HJM model can be seen as a shifted lognormal model on Libors and by letting the size of the shift vary we obtain a class of dynamics that includes both the Libor Market Model and the Gaussian HJM. We can write the discrete Tenor dynamics (see Musiela & Rutkowski (1997)) of this skewed market model by specifying the forward volatility:

$$dK(s,T_i) = (a_i + K(s,T_i))\,\gamma(s,T_i - s)dW_s^{T+\delta}$$

This setup includes the classical Libor Market model with $a_i = 0$ and the Gaussian HJM with $a_i = 1/\delta$. With $swap(s,T,T_N) = \sum_{i=1}^N \omega_i K(s,T_i)$ and $(a_i + K(s,T_i))$ lognormal, the swap is a shifted sum of lognormal processes and can be approximated by a shifted lognormal process:

$$swap(s,T,T_N) = \sum_{i=1}^{N} \omega_i (a_i + K(s,T_i)) - \sum_{i=1}^{N} \omega_i a_i$$

This means that we can use the swaption pricing formulas above by simply adjusting the strike to account for the shift $\sum_{i=1}^N \omega_i a_i$. Each a_i allows an additional degree of freedom in calibrating the skewness. Finally, to accommodate some adjustment in the smile convexity, we can suppose for example that the volatility follows a Markov process with two states (and for example exponentially distributed interarrival times). This very basic setup models the regime switching in volatility that is often observed in practice (see figure 4.8). If we further suppose that the switch in regime is independent of the dynamics of the underlying rate, we can price the swaption by conditioning on the path of the volatility, computing the expected value of the approximation formulas found above for the two volatility states:

$$P_{swaption} = p_{high} BS\left(V_{t,T}^{high}\right) + (1 - p_{high})\, BS\left(V_{t,T}^{low}\right)$$

We have created some additional flexibility in the smile shape but a simultaneous calibration of the smile and the covariance is impossible. The calibration must then be performed in two steps: first the smile shape and level ($a_i^{high}, V_{t,T}^{high}, a_i^{low}$ and $V_{t,T}^{low}$) is fitted on market data on caplets or swaptions, then the covariance in each regime is calibrated using the ATM vol. levels from the calibrated smile.

Part II

Risk-Management

Introduction

In the previous part we showed how under the market model of interest rates, the forward swap can be successfully approximated by a lognormal diffusion, allowing swaptions to be priced consistently with the market practice of using the Black (1976) pricing formula. The variance used in this approximation was computed as a linear form on the covariance matrix of the forward rates, which cast the calibration problem as that of finding a positive semidefinite matrix verifying a set of linear constraints, in other words, solving a semidefinite program. Recent optimization techniques solve these problems with very low practical complexity and excellent stability.

Here, we exploit some natural by-products of the calibration procedure which provide an excellent description of the solution sensitivity to a change in the market constraints. In fact, because the algorithms used to solve the calibration problem jointly solve the problem and its dual, *the sensitivity of the optimal objective value is readily available as the dual solution to the calibration program*. Further analysis allows us to deduce in the same way the sensitivity of the calibrated covariance matrix itself. We can, for example, compute the sensitivity of a particular swaption with respect to another swaption price, or estimate the sensitivity of the entire calibrated matrix with respect to a given change in market conditions. This completes the set of fast and stable algorithms for calibrating the market model of interest rates by another set of equally fast and stable methods for the model risk-management.

We then show how the calibration problem can be designed so that its solution is optimally robust to a given change in market conditions. We detail various problem formulation for different market movement models. Finally, we show how these same optimization techniques can be used to efficiently solve the problem of managing the Gamma exposure of a basket portfolio, as it was posed by Douady (1995).

The lognormal approximation for basket pricing dates back to Huynh (1994) and Musiela & Rutkowski (1997). Brace et al. (1999) also tested the validity of its application to swaption pricing. Two papers, Rebonato (1998) and Rebonato (1999), highlight the importance of jointly calibrating the volatility and the correlation matrix. They also detail some of the most common non-convex calibration techniques, based on the parametrization of the forward rates covariance factors on a hypersphere. The idea of exploiting the dual solution to the approximate calibration problem to improve hedging can be traced back to the work by Avellaneda, Levy & Paras (1995) and Avellaneda & Paras (1996) on the equity market.

This part is organized around two key contributions:

- In a first chapter, we show how the dual solution to the calibration program directly provides *all* the sensitivities of the calibrated covariance to small changes in market conditions. We also show how to make the calibration optimally robust to these changes.

- In a second chapter, we write a primal-dual pair of semidefinite programs giving the upper (or lower) bound on swaption prices that has a direct interpretation as a hedging program à la Avellaneda & Paras (1996). We then revisit a related result of Romagnoli & Vargiolu (2000) in the light of semidefinite programming.

The results we obtain here underline the key advantages of applying symmetric cone programming methods to the calibration problem: besides their radical numerical performance, they naturally provide some central results on the sensitivity and risk-management of the solution. They also avoid the numerical errors in the sensitivity computations that were caused by the instability of the classical non-convex calibration solution. This should greatly improve the pricing and hedging process of exotic interest rate derivatives by reducing the part of hedge portfolio rebalancing that was only caused by the calibration program instability.

Chapter 5

Risk Management and Sensitivity analysis

5.1 The generic calibration program

We start by recalling the basic results from part one. We know that for option pricing purposes, we can approximate the forward swap by a lognormal process defined by:

$$\frac{dswap(s, T_0, T_N)}{swap(s, T_0, T_N)} = \sum_{i=1}^{N} \widehat{\omega}_{i,s} \gamma(s, T_j - s) dW_s \tag{5.1}$$

where $\widehat{\omega}_i$ is defined by:

$$\widehat{\omega}_{i,t} = \frac{\omega_i(t) K(t, T_i)}{\sum_{j=1}^{n} \omega_j(t) K(t, T_j)}$$

which can be computed from the market data today. We can use the first order basket pricing approximation in Huynh (1994) and compute the price of a payer swaption starting at T with maturity at T_N and strike κ using the Black (1976) pricing formula:

$$Level(t, T, T_N) \left(swap(t, T, T_N) N(h) - \kappa N \left(h - \sqrt{V_T} \right) \right) \tag{5.2}$$

where

$$h = \frac{\left(\ln \left(\frac{swap(t,T,T_N)}{\kappa} \right) + \frac{1}{2} V_T \right)}{\sqrt{V_T}}$$

with $swap(t, T, T_N)$, the value of the forward swap today and

$$\begin{aligned} V_T &= \int_t^T \left\| \sum_{i=1}^{N} \hat{\omega}_i(t) \gamma(s, T_i - s) \right\|^2 ds \\ &= \int_t^T \mathbf{Tr} \left(\Omega_t \Gamma_s \right) ds \end{aligned} \tag{5.3}$$

which is a linear form on the forward rates covariance. Having constructed Ω_t and $\Gamma_s \in \mathbf{S}^N$ such that:

$$\Omega_t = \hat{\omega} \hat{\omega}^{\mathsf{T}} = (\hat{\omega}_i \hat{\omega}_j)_{i,j \in [1,N]} \succeq 0$$

and

$$\Gamma_s = \left(\langle \gamma(s,T_i - s), \gamma(s,T_j - s)\rangle\right)_{i,j\in[1,N]} \succeq 0$$

which is the covariance matrix of the forward rates (or the Gram matrix of the $\gamma(s,T_i - s)$ volatility function defined above). We have here neglected a term coming from the volatility of the weights, because as $\delta K(t,T_i) << 1$, we observe that this last term can be discarded without significant precision loss and that the *swaption can be priced as a basket* with constant coefficients.

Our objective here is to show how Lagrangian duality will allow us to use the primal and dual solutions produced by the solver to collect information on the sensitivity of the calibrated solution to a change in the given market prices. For simplicity, we discretize the volatility function yearly and make it piecewise constant. We start by quickly recalling the practical implementation of the calibration program using the swaption pricing approximation detailed above. This is done by discretizing in s the covariance matrix Γ_s. We note $\mathbf{S}^n$ the set of symmetric matrixes of size $n \times n$. We suppose that the calibration data set is made of m swaptions with option maturity T_{S_k} written on swaps of maturity $T_{N_k} - T_{S_k}$, with market volatility given by σ_k.

5.1.1 A simple example

Let $S = \max_{k=1,...,m} S_k$ and M be the maximum number of periods covered by all the input instruments. In the very simple case where the volatility of the forwards is of the form $\gamma(s,T-s) = \gamma(T-s)$ with γ piecewise constant over intervals of length δ, the calibration problem becomes:

$$\begin{array}{ll} \text{find} & X \\ \text{s.t.} & \mathbf{Tr}(\Omega_k X) = \sigma_k^2 T_{S_k} \text{ for } k = 1,...,m \\ & X \succeq 0 \end{array} \tag{5.4}$$

which is a semidefinite feasibility problem in the covariance matrix $X \in \mathbf{S}^M$ ($X \succeq 0$ meaning X p.s.d.). As above, $\sigma_k^2 T_k \in \mathbb{R}_+$ is the Black (1976) cumulative variance of swaption k written on $swap(t,T_{S_k},T_{N_k})$ and $\Omega_k = \sum_{j=1}^{S_k} \delta\varphi_{k,j}$ with $\varphi_{k,j} \in \mathbf{S}^M$ the rank one matrix with submatrix $\hat{\omega}_k \hat{\omega}_k^T$ starting at element (j,j) and all other blocks equal to zero. Note that $\hat{\omega}_k$ is here the vector of weights associated to swaption k with $\hat{\omega}_k = (\hat{\omega}_{i,k})_{i=S_k,...,N_k}$.

5.1.2 The general case

Here we show that for general volatilities $\gamma(s,T-s)$, the format of the calibration problem remains similar to that of the simple example above, except that X will be block-diagonal. In the general non-stationary case where γ is of the form $\gamma(s,T-s)$ and piecewise constant on intervals of size δ, the expression of the market cumulative variance becomes

$$\sigma_k^2 T_{S_k} = \sum_{i=0}^{T_{S_k}} \delta Tr\left(\Omega_{k,i} X_i\right)$$

where $\Omega_{k,i} \in \mathbf{S}^{M-i}$ is a block-matrix with submatrix $\hat{\omega}_k \hat{\omega}_k^T$ starting at element $(S_k - i, S_k - i)$ and all other blocks equal to zero if $S_k - i \geq 0$ and is zero otherwise. Calibrating the model to the swaptions $k = (1,...,m)$ can then be written as the following semidefinite feasibility problem.:

$$\begin{array}{ll} \text{find} & X_i \quad i = i,...,T_M \\ \text{s.t.} & \sum_{i=0}^{T_{S_k}} \delta_i\, \mathbf{Tr}\left(\Omega_{i,k} X_i\right) = \sigma_k^2 T_k \text{ for } k = 1,...,m \\ & X_i \succeq 0 \text{ for } i = 0,...,T_S \end{array}$$

and the variables here are the matrixes $X_i \in \mathbf{S}^{M-i}$. We can write this general problem in the exact same format used in the simple stationary case. Let X be the block matrix

$$X = \begin{bmatrix} X_1 & 0 & . & 0 \\ 0 & . & . & . \\ . & . & . & 0 \\ 0 & . & 0 & X_{T_M} \end{bmatrix}$$

then the calibration program can be written in the same format as (5.4):

$$\begin{array}{ll} \text{find} & X \\ \text{s.t.} & \mathbf{Tr}(\bar{\Omega}_k X) = \sigma_k^2 T_k \text{ for } k = 1, ..., m \\ & X \succeq 0, X \text{ band-diagonal} \end{array} \tag{5.5}$$

except that $\bar{\Omega}_k$ and $X \in \mathbf{S}^{M-i}$ are here ”block-diagonal”. We can also replace the equality constraints in (5.4) with Bid-Ask spreads. The new calibration problem is then written as the L.M.I.:

$$\begin{array}{ll} \text{find} & X \\ \text{s.t.} & \sigma^2_{Bid,k} T_{S_k} \leq \mathbf{Tr}(\Omega_k X) \leq \sigma^2_{Ask,k} T_{S_k} \text{ for } k = 1, ..., m \\ & X \succeq 0 \end{array}$$

in the variable $X \in \mathbf{S}^M$, with parameters Ω_k, $\sigma^2_{Bid,k}$, $\sigma^2_{Ask,k}$, T_{S_k}. Let us note that we can rewrite this problem as the L.M.I.:

$$\begin{array}{ll} \text{find} & X \\ \text{s.t.} & \mathbf{Tr}\left(\begin{bmatrix} \Omega_k & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} X & 0 & 0 \\ 0 & U_1 & 0 \\ 0 & 0 & U_2 \end{bmatrix} \right) = \sigma^2_{Ask,k} T_{S_k} \\ & \mathbf{Tr}\left(\begin{bmatrix} \Omega_k & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -I \end{bmatrix} \begin{bmatrix} X & 0 & 0 \\ 0 & U_1 & 0 \\ 0 & 0 & U_2 \end{bmatrix} \right) = \sigma^2_{Bid,k} T_{S_k} \text{ for } k = 1, ..., m \\ & X, U_1, U_2 \succeq 0 \end{array}$$

which is a standard form L.M.I. and can be summarized as

$$\begin{array}{ll} \text{find} & \tilde{X} \\ \text{s.t.} & \mathbf{Tr}(\tilde{\Omega}_{Ask,k} \tilde{X}) = \sigma^2_{Ask,k} T_{S_k} \\ & \mathbf{Tr}(\tilde{\Omega}_{Bid,k} \tilde{X}) = \sigma^2_{Bid,k} T_{S_k} \text{ for } k = 1, ..., m \\ & \tilde{X} \succeq 0, \tilde{X} \text{ block-diagonal} \end{array} \tag{5.6}$$

with $\tilde{X}, \tilde{\Omega}_k \in \mathbf{S}^{3M}$. Because of these transformations and to simplify the analysis, we will always handle the stationary case with equality constraints in the following section, knowing that all results can be directly extended to the general case (non-stationary covariance with Bid-Ask constraints) by embedding them in a larger standard form semidefinite program. Furthermore, the common practice of *daily model recalibration* makes the calibration of a model *as close to stationary as possible* a central requirement to ensure the risk-management’s coherence.

5.2 Semidefinite duality

We very briefly summarize here the duality theory for semidefinite programming. We refer the reader to Nesterov & Nemirovskii (1994) orVandenberghe & Boyd (1996) for a complete analysis. As we have seen in the previous section, the calibration problem can be written as a standard form primal semidefinite program:

$$\begin{array}{ll} \text{maximize} & \mathbf{Tr}(CX) \\ \text{s.t.} & \mathbf{Tr}(\Omega_k X) = \sigma_k^2 T_{S_k} \text{ for } k = 1, ..., m \\ & X \succeq 0 \end{array} \tag{5.7}$$

in the variable $X \in \mathbf{S}^M$ with parameters $\Omega_k, C \in \mathbf{S}^M$ and $\sigma_k^2 T_{S_k} \in \mathbb{R}_+$, where, for example, we can set C as a target covariance matrix. For $X \succeq 0$, $y \in \mathbb{R}^m$, we form the following Lagrangian:

$$\begin{aligned} L(X, y) &= -\mathbf{Tr}(CX) + \sum_{k=1}^{m} y_k \left(\mathbf{Tr}(\Omega_k X) - \sigma_k^2 T_{S_k}\right) \\ &= \mathbf{Tr}\left(\sum_{k=1}^{m} (y_k \Omega_k - C) X\right) - \sum_{k=1}^{m} y_k \sigma_k^2 T_{S_k} \end{aligned}$$

and because the semidefinite cone is self-dual, we find that $L(X, y)$ is bounded below in $X \succeq 0$ iff:

$$0 \preceq \sum_{k=1}^{m} y_k \Omega_k - C$$

hence the dual semidefinite problem becomes:

$$\begin{array}{ll} \text{minimize} & -\sum_{k=1}^{m} y_k \sigma_k^2 T_{S_k} \\ \text{s.t.} & 0 \preceq \left(\sum_{k=1}^{m} y_k \Omega_k - C\right) \end{array} \tag{5.8}$$

All modern solvers (see for example Sturm (1999)) produce both primal and dual solutions to this problem as well as a certificate of optimality for the solution in the form of the associated duality gap:

$$\mu = \mathbf{Tr}\left(X \left(\sum_{k=1}^{m} y_k \Omega_k - C\right)\right)$$

which is an upper bound on the absolute error. We now show how this dual solution can be used for risk-management purposes.

5.3 Sensitivity Analysis

Let us suppose that we have solved both the primal and the dual calibration problems above with market constraints σ_i^2 and let us note X^{opt} and y^{opt} the optimal solutions. Suppose also that the market price constraints in the original calibration problem are modified by a small amount $u \in \mathbb{R}^m$. The new calibration problem becomes the following semidefinite program:

$$\begin{array}{ll} \text{maximize} & \mathbf{Tr}(CX) \\ \text{s.t.} & \mathbf{Tr}(\Omega_k X) = \sigma_k^2 T_k + u_k \text{ for } k = 1, ..., m \\ & X \succeq 0 \end{array}$$

in the variable $X \in \mathbf{S}^M$ with parameters $\Omega_k, C \in \mathbf{S}^M$ and $\sigma_k^2 T_k \in \mathbb{R}_+$ for $k = 1, ..., m$ and if we note $p^{opt}(u)$ the optimal solution to the revised problem and assume that it is differentiable, we get the sensitivity to a change in market condition as:

$$\frac{\partial p^{opt}(0)}{\partial u_k} = -y_k^{opt} \tag{5.9}$$

where y^{opt} is the optimal solution to the dual problem. As we will see below, this has various interpretations depending on the objective function.

5.3.1 Sensitivity of price bounds

In the spirit of El Karoui & Quenez (1991) and El Karoui et al. (1998), let us suppose that the market rate dynamics are exactly that of the Libor Market Model above. In this case, the "real" prices of caps and swaptions are given by an unknown covariance matrix V^r. We observe a set of cap/swaption prices given by their Black-Scholes implied volatility σ_k. Given those prices, we are interested in computing the maximum and minimum price of another swaption. To do this, we can set the objective matrix C to be an instrument matrix Ω_0 and solve the calibration problem. We obtain the upper (resp. lower) bound on another swaption price as a semidefinite program, solving for the covariance matrix that maximizes (resp. minimizes) the objective:

$$\begin{array}{rll} p_{up} = & \max & \mathbf{Tr}(\Omega_0 X) \\ & \text{s.t.} & \mathbf{Tr}(\Omega_k X) = \sigma_k^2 T_k + u_k \text{ for } k = 1, ..., m \\ & & X \succeq 0 \end{array}$$

and

$$\begin{array}{rll} p_{down} = & \min & \mathbf{Tr}(\Omega_0 X) \\ & \text{s.t.} & \mathbf{Tr}(\Omega_k X) = \sigma_k^2 T_k + u_k \text{ for } k = 1, ..., m \\ & & X \succeq 0 \end{array}$$

in the variable $X \in \mathbf{S}^M$ with parameters $\Omega_k, C \in \mathbf{S}^M$ and $\sigma_k^2 T_k \in \mathbb{R}_+$, for $u_k = 0$ and $k = 1, ..., m$.

In this case we can directly compute the sensitivity of the product's price bounds to a change in the cap/swaption price data. As above, let us suppose that the dual solutions to the Max. (resp. Min.) problems are given by y_{up} and y_{down} and that the market prices have been modified by some small quantities $(u_i) \in \mathbb{R}^M$. We can compute (at least formally here) the sensitivity of the upper and lower bounds as:

$$\frac{\partial p_{up}^{opt}(0)}{\partial u_k} = -y_{up,k}^{opt} \text{ and } \frac{\partial p_{low}^{opt}(0)}{\partial u_k} = -y_{low,k}^{opt} \tag{5.10}$$

respectively. In fact, as we will now show, the dual solution allows to compute the sensitivity of the optimal covariance matrix itself to a change in market conditions. Besides giving a rigorous computation of the above sensitivities, we will also be able to compute the exposure of a given product to *actual market scenarios*.

Remark 28 *Intuitively, we expect the dual solution to represent the coefficients of a hedging portfolio. As we detail in a the next section, this is indeed the case and the y_k^{opt} represent the optimal number of swaptions k to hold the superreplicating portfolio defined in Avellaneda & Paras (1996).*

5.3.2 Solution sensitivity

We now study the variation in the matrix solution itself, given a small change in the market conditions.

The Newton step

Here we study the impact of a change in market conditions on the solution itself. Let us suppose that we have solved the general calibration problem:

$$\begin{array}{ll} \text{maximize} & \mathbf{Tr}(CX) \\ \text{s.t.} & \mathbf{Tr}(\Omega_k X) = \sigma_k^2 T_k + u_k \text{ for } k = 1, ..., m \\ & X \succeq 0 \end{array} \tag{5.11}$$

$X \in \mathbf{S}^M$ with parameters $\Omega_k, C \in \mathbf{S}^M$ and $\sigma_k^2 T_k \in \mathbb{R}_+$, for $u_k = 0$ and $k = 1$. Here C is, for example, an historical estimate of the covariance matrix).

Notation 29 *We introduce here a set of standard semidefinite programming notations. Let us call X^{opt} and y^{opt} the primal and dual solutions to the above problem. We note*

$$Z^{opt} = \left(C - \sum_{k=1}^{M} y_k^{opt} \Omega_k \right)$$

the dual solution matrix. As in the paper by Alizadeh et al. (1998), we also set the symmetric Kronecker product:

$$(P \circledast Q)\, K := \frac{1}{2} \left(PKQ^T + QKP^T \right)$$

As in (4.7), we note A the linear operator defined by:

$$\begin{aligned} A : \mathbf{S}^M &\longrightarrow \mathbb{R}^m \\ X &\longmapsto AX := \left(\mathbf{Tr}\left(A_i X \right) \right)_{i=1,\ldots,m} \end{aligned}$$

and its dual

$$\begin{aligned} A^* : \mathbb{R}^m &\longrightarrow \mathbf{S}^M \\ y &\longmapsto A^* y := \sum_{i=1}^{m} y_i \Omega_i \end{aligned}$$

We will now use the results in Todd & Yildirim (1999) to compute the impact ΔX on the solution of a small change in the market price data $(u_k)_{k=1,..,m}$, i.e. given u small enough we will compute the next Newton step ΔX. Each solver implements one particular search direction to compute this step and some of the most common ones are the A.H.O. search direction based on the work by Alizadeh et al. (1998), the H.K.M. direction by Helmberg, Rendl, Vanderbei & Wolkowicz (1996), Kojima, Shindoh & Hara (1997) and Monteiro (1997) and finally the N.T. direction detailed in Nesterov & Todd (1997) and Nesterov & Todd (1998). Depending on the choice of the search direction, we define a matrix P such that:

- $P = I$ for the A.H.O. direction
- $P = Z^{opt}$ for the H.K.M. direction
- $P = \left(X^{opt\frac{1}{2}} \left(X^{opt\frac{1}{2}} Z^{opt} X^{opt\frac{1}{2}} \right)^{\frac{-1}{2}} X^{opt\frac{1}{2}} \right)^{-1}$ for the N.T. direction

we can then define the linear operators:

$$E = Z^{opt} \circledast P \ \text{ and } \ F = PX^{opt} \circledast I$$

and their adjoints

$$E^* = Z^{opt} \circledast P \ \text{ and } \ X^{opt}P \circledast I$$

and provided the strict feasibility and nonsingularity conditions in §3 of Todd & Yildirim (1999) hold, we can compute the Newton step ΔX as:

$$\Delta X = E^{-1}FA^* \left[\left(AE^{-1}FA^* \right)^{-1} u \right] \tag{5.12}$$

and we know that this will lead to a feasible point $X^{opt} + \Delta X \succeq 0$ iff the market variation movement u is such that:

$$\left\| \left(X^{opt}\right)^{-\frac{1}{2}} \left(E^{-1}FA^* \left[\left(AE^{-1}FA^* \right)^{-1} u \right] \right) \left(X^{opt}\right)^{-\frac{1}{2}} \right\|_2 \leq 1 \tag{5.13}$$

We remark that if $A, B \in \mathbf{S}^M$ commute, with eigenvalues $\alpha, \beta \in \mathbb{R}^M$ and common eigenvectors v_i for $i = 1, ..., M$, then $A \circledast B$ has eigenvalues $(\alpha_i \beta_j + \alpha_j \beta_i)$ for $i, j = 1, ..., M$ and eigenvectors $v_i v_i^T$ if $i = j$ and $(v_i v_j^T + v_j v_i^T)$ if $i \neq j$ for $i, j = 1, ..., M$. The matrix in (5.12) produces a *direct method for updating* X which we can now use to compute price sensitivities for any given portfolio. This illustrates how a semidefinite programming based calibration allows to test various *realistic scenarios* at a minimum numerical cost by improving on the classical non-convex methods that either had to ”bump the market data and recalibrate” the model for every scenario with the risk of jumping from one local optimum to the next, or simulate unrealistic market movements by directly adjusting the covariance matrix. One key remaining question is that of stability: the calibration program in (5.4) has a unique solution, but this optimum can be very unstable and the matrix in (5.12) badly conditioned. In the spirit of the work by Cont (2001) on volatility surfaces, we now look for a way to stabilize the calibration result.

5.4 Robustness

The previous sections were focused on how to compute the impact of a change in market conditions. Here we will focus on how to anticipate those variations and make the calibrated matrix optimally robust to a given set of scenarios. Depending on the way the perturbations are modelled, this problem can remain convex and be solved very efficiently. Let us suppose here that we want to solve the calibration problem on a set of market Bid-Ask spreads data defined by the following L.M.I.:

$$\begin{array}{ll} \text{find} & X \\ \text{s.t.} & \sigma^2_{Bid,k} T_{S_k} \leq \mathbf{Tr}(\Omega_k X) \leq \sigma^2_{Ask,k} T_{S_k} \text{ for } k = 1, ..., m \\ & X \succeq 0 \end{array}$$

in the variable $X \in \mathbf{S}^M$ with parameters $\Omega_k, C \in \mathbf{S}^M$ and $\sigma^2_{Bid,k} T_{S_k}$, $\sigma^2_{Ask,k} T_{S_k} \in \mathbb{R}_+$ for $k = 1, ..., m$. In the absence of any information on the uncertainty in the market data, we can simply maximize the distance between the solution and the market bounds to ensure that it remains valid in the event of a small change in the market variance input. As the robustness objective is equivalent to a maximization of the distance between the solution and the constraints (or Chebyshev centering), the input of assumptions on the movement structure is equivalent to a choice of norm. Without any particular structural information on the volatility market dynamics, we can use the l_∞ norm and the calibration problem becomes:

$$\begin{array}{ll} \text{maximize} & t \\ \text{s.t.} & \sigma^2_{Bid,k} T_{S_k} + t \leq \mathbf{Tr}(\Omega_k X) \leq \sigma^2_{Ask,k} T_{S_k} - t \;\; \text{for } k = 1, ..., m \\ & X \succeq 0 \end{array}$$

or, using the l_1 norm instead:

$$\begin{array}{ll} \text{maximize} & \sum_{i=1}^{m} t_k \\ \text{s.t.} & \sigma^2_{Bid,k} T_{S_k} + t_k \leq \mathbf{Tr}(\Omega_k X) \leq \sigma^2_{Ask,k} T_{S_k} - t_k \;\; \text{for } k = 1, ..., m \\ & X \succeq 0 \end{array}$$

The problems above optimally center the solution within the Bid-Ask spreads, which makes it robust to a change in market conditions given no particular information on the nature of that change. In the same vein, Ben-Tal, El Ghaoui & Lebret (1998) also show how to design a program that is robust to a change in the matrixes Ω_k. However, because the matrixes Ω_k are computed from ratios of zero-coupon bonds, their variance is negligible compared to that of σ_k^2.

Suppose now that V is a statistical estimate of the daily covariance of the changes in $\sigma_k^2 T_{S_k}$ (the mid-market volatilities in this case) and let us assume that these volatilities have a Gaussian distribution, we adapt the method used by Lobo, Vandenberghe, Boyd & Lebret (1998) for robust L.P. Let η be a given confidence level, we require that each price constraint should hold with a probability exceeding η. We suppose that the matrix V is of full rank. We can then construct the following program:

$$\begin{array}{ll} \text{maximize} & \mathbf{Tr}(CX) \\ \text{s.t.} & \mathbf{Tr}\,(\Omega_k X) - \sigma_k^2 T_{S_k} = v_k \;\; \text{for } k = 1, ..., m \\ & \left\| V^{-\frac{1}{2}} v \right\|_\infty \leq \Phi^{-1}(\mu) \\ & X \succeq 0 \end{array}$$

where $\|\cdot\|_\infty$ is the l_∞ norm and $\Phi(x)$ is given by

$$\Phi(x) = 1 - \frac{1}{\sqrt{2\pi}} \int_{-x}^{x} \exp(-u^2/2) du$$

This ensures that each price constraint will hold with a probability exceeding η. There is no guarantee that this program is feasible and we can solve instead for the best confidence level by forming the following program:

$$\begin{array}{ll} \text{minimize} & t \\ \text{s.t.} & \mathbf{Tr}\,(\Omega_k X) - \sigma_k^2 T_{S_k} = v_k \;\; \text{for } k = 1, ..., m \\ & \left\| V^{-\frac{1}{2}} v \right\|_\infty \leq t \\ & X \succeq 0 \end{array}$$

The optimal confidence level is then $\eta = \Phi(t)$ and "centers" the calibrated matrix with respect to the uncertainty in $\sigma_k^2 T_{S_k}$. This is a symmetric cone program, i.e. a program mixing LP, second-order and semidefinite cone constraints, and can be solved very efficiently using the code by Sturm (1999) for example.

Chapter 6

Hedging

Here, we show how the programs solved above can be used to build superreplicating portfolios, providing an upper and lower hedging prices in the sense of El Karoui & Quenez (1991) and El Karoui & Quenez (1995). An efficient technique for computing those price bounds with general non-convex payoffs on a single asset with univariate dynamics was introduced by Avellaneda et al. (1995) and we adapt it here to the approximate basket pricing problem. Recent work on this topic by Romagnoli & Vargiolu (2000) provided closed-form solutions for the prices of exchange options and options on the geometrical mean of two assets. Gozzi & Vargiolu (2000) applied this technique to caps and Floors.

We start from the approximate pricing PDE to compute the price of a particular option and use it to compute arbitrage bounds on the price of a basket using the method developed by Avellaneda et al. (1995) in the one-dimensional case. We then provide approximate (to within 1-2%) closed-form solutions for these arbitrage bounds based on the semidefinite programs detailed in the last section and show how one can built an optimal hedging portfolio in the sense of Avellaneda & Paras (1996), using the derivative securities from the calibration set.

Finally, we show how these same optimization techniques can also be used to efficiently solve the problem of managing the Gamma exposure of a basket portfolio, as it was posed by Douady (1995).

6.1 The approximate basket pricing PDE

We suppose that the market is composed of n risky assets $S_t^i, i = 1, \ldots, n$ plus one riskless asset M_t. We assume that these processes are defined on a probability space (Ω, F, P) and adapted to a filtration $\{F_t, 0 \leq t \leq T\}$. We suppose that there exists a forward martingale measure Q as defined in El Karoui et al. (1995), which can be either the $Level$ forward market for swaption pricing or the standard forward measure of maturity T in the general basket case. For simplicity we will note F_t^i the forward rates and F_t^ω the forward swap (which in what follows could be any basket of forwards). We note $\gamma^i = \gamma(t, T_j - t)$ and for simplicity, we drop the time dependency the extension of what follows to the general case being straightforward. We recall the simple market definition in Avellaneda & Paras (1996). In this market, the dynamics of the forwards F_t^i are given by:

$$\begin{aligned} & m_t = 1 \\ & dF_s^i = F_s^i \gamma^i dW_t^Q \end{aligned} \tag{6.1}$$

where W_t is a d-dimensional Q-Brownian motion adapted to the filtration $\{F_t\}$ and $\gamma = \left(\gamma^i\right)_{i=1,\ldots,n} \in \mathbb{R}^{n\times d}$ is the volatility matrix.

Definition 30 *A portfolio process* $\{\Delta_t \in \mathbb{R}^n, 0 \leq t \leq T\}$ *is a bounded adapted process representing the quantity of each asset held at time* t*. The value of a portfolio is then given by* $\Pi_t = m_t + \Delta_t F_t^i$*. Given an initial value* Π_0*, a portfolio is said to be self-financing iff its dynamics are given by:*

$$d\Pi_t = \Delta_t dF_t^i \tag{6.2}$$

We study the pricing of an option on a basket of forwards given by

$$F_t^\omega = \sum_{i=1}^n \omega_i F_t^i$$

where $\omega = (\omega_i)_{i=1,\ldots,n} \in \mathbb{R}^n$. The terminal payoff of this option at maturity T is then computed as:

$$h_{\omega,k}\left(F_T^\omega\right) = \left(\sum_{i=1}^n \omega_i F_T^i - k\right)^+$$

for some strike price $K \geq 0$. We recall the basket approximation result

The dynamics of the basket of forwards can be approximated by the process:

$$dF_s^\omega = \sum_{i=1}^n \widehat{\omega}_{i,t} \gamma^i dW_s \tag{6.3}$$

where $\eta_s \in \mathbb{R}^n, t \leq s \leq T$, is computed from the market data today and $\widehat{\omega}_i$ is defined by:

$$\widehat{\omega}_i = \frac{\omega_i F_t^i}{\sum_{i=1}^n \omega_i F_t^i}$$

which can also be computed from the market data today.

6.2 Quasi-static arbitrage bounds

The approximation formula above allows the pricing of baskets using the Black & Scholes (1973) formula with a variance that is a linear function constructed from the variance parameter $\sigma^T\sigma$ and a matrix that can be computed from the market data today. Because a set of price calibration constraints then simply becomes a set of linear constraints on the covariance matrix, the set of market calibrated matrixes is the feasible set of a Linear Matrix Inequality. It then becomes possible to investigate upper and lower bound on the variance of a particular basket given the price of a set of other baskets and vanilla options as well as design hedging portfolio using the dual program. In this section, we show that this produces dynamic arbitrage bounds similar to that found in El Karoui & Quenez (1991) and Avellaneda et al. (1995) and construct an optimal mixed dynamic - static hedge (hence the term quasi-static), where a static hedging portfolio is constructed using the calibration instruments and the residual risk is hedged dynamically.

6.2.1 Approximate bounds

As in the last section, let us quickly recall how we can maximize the price of a particular swaption given the market data $\sigma_k^2 T_k$ for $k = 1, ..., M$ by solving a semidefinite program. Suppose we have a set of market prices represented by volatilities $\sigma_i, i = 1, \ldots, M$ for basket options with coefficients

$\omega_k \in \mathbb{R}^{N_k}, k = 1, \ldots, M$, we can compute an upper bound on the price of another basket ω_0 by solving the following program:

$$\begin{array}{ll} \text{maximize} & \sigma_{\max}^2 T = \mathbf{Tr}(\Omega_0 X) \\ \text{s.t.} & \mathbf{Tr}(\Omega_k X) = \sigma_k^2 T_k \text{ for } k = 1, ..., M \\ & X \succeq 0 \end{array} \tag{6.4}$$

Having constructed Ω_t and $\Gamma_s \in \mathbf{S}^N$ such that:

$$\Omega_t = \widehat{\omega}\widehat{\omega}^T = (\widehat{\omega}_{i,0}\widehat{\omega}_{j,0})_{i,j\in[1,N]} \succeq 0$$

and

$$X = \left(\gamma^{iT}\gamma^j\right)_{i,j\in[1,N]} \succeq 0$$

As we have seen in the first section, we can then compute an upper bound on the price of the option of strike κ using the Black & Scholes (1973) formula:

$$F_t^\omega N(h) - KN\left(d - \sigma_{\max}\sqrt{T}\right)$$

where

$$d = \frac{\left(\ln\left(\frac{F_t^\omega}{\kappa}\right) + \frac{1}{2}\sigma_{\max}^2 T\right)}{\sigma_{\max}\sqrt{T}}$$

When the volatilities γ^i are constant, this appears directly in the Black-Scholes-Barenblatt equation.

Proposition 31 *Suppose that the volatilities γ^i are constant. The price obtained by solving the program in (6.4) is an approximate arbitrage upper bound on the basket Call price given the market data (σ_i, ω_i) for $i = 1, \ldots, k$.*

Proof. For a given volatility parameter γ, we know (see for example Karatzas & Shreve (1991)) that the approximate Call price follows the PDE:

$$\begin{cases} \frac{\partial C(x,t)}{\partial t} + \frac{1}{2}\left\|\sum_{i=1}^n \widehat{\omega}_{i,0}\gamma^i\right\| x\frac{\partial^2 C(x,t)}{\partial x^2} = 0 \\ \\ C(x,T) = (x-k)^+ \end{cases}$$

As above, we can rewrite:

$$\left\|\sum_{i=1}^n \widehat{\omega}_{i,0}\gamma^i\right\| = Tr(\Omega_0 X)$$

and we know from Avellaneda et al. (1995) that the upper bound $C(x,t)$ on the price of a basket Call must verify the Black-Scholes-Barenblatt equation:

$$\begin{cases} \frac{\partial C(x,t)}{\partial t} + \frac{1}{2}\max_{Tr(\Omega_k X)=\sigma_k^2 T_k} \mathbf{Tr}(\Omega_0 X) x\frac{\partial^2 C(x,t)}{\partial x^2} = 0 \\ \\ C(x,T) = (x-k)^+ \end{cases}$$

which is again

$$\begin{cases} \frac{\partial C(x,t)}{\partial t} + \frac{1}{2}\sigma_{\max} x\frac{\partial^2 C(x,t)}{\partial x^2} = 0 \\ \\ C(x,T) = (x-k)^+ \end{cases}$$

where $\sigma_{\max}$ is the optimum value of the program in (6.4). ■

As in El Karoui et al. (1998) and Avellaneda et al. (1995), a superreplicating strategy is then obtained by maintaining a self-financed dynamic portfolio composed of cash and delta shares, where the delta is computed using the volatility found by solving the semidefinite program above. Hence, we set $\Delta_t = \frac{\partial C^{\sigma_{\max}}(S_t,t)}{\partial x}$ to be the amount of basket underlying in the replicating portfolio at time t. As in Avellaneda & Paras (1996), we can improve that dynamic hedge by adding a static portfolio of derivatives and solve the following (formal) program:

$$\textbf{Price} = \text{Min}\left\{\text{Value of } \textbf{static hedge} + \text{Max}\left(\text{PV of } \textbf{residual liability}\right)\right\}$$

Because of the sub-additivity of the above program with respect to payoffs, we expect this diversification of the volatility risk to bring down the total cost of hedging. Suppose we have a set of market prices $C_i, i = 1, \ldots, M$ for basket options with coefficients $\omega_i \in \mathbb{R}^n, i = 1, \ldots, M$ and payoffs $h_{\omega_i,K_i}(F_T) = \left(\sum_j \omega_{i,j} F_T^j - k_j\right)^+$ where $F_T = \left(F_T^i\right)_{i=1,\ldots,n}$.

Remark 32 *It is important to understand that the two parts of the PV equation above are not computed using the same pricing methodology. One, the Value of the* ***static hedge*** *is a static portfolio of instruments quoted by the market today, the other, the PV of the* ***residual liability****, is computed as the value today of a self-financing dynamic portfolio that superreplicates the residual liability. This last computation is subject to much stronger assumptions on the market liquidity than the first one. Despite this apparent inhomogeneity in method, the Price computed above still accurately reflects the behavior of a market operator who prices his products by calibrating on a liquid set of instruments that will also constitute his hedge and looks for the most conservative set of parameters.*

As more and more time dependence and flexibility is allowed in the model, the price obtained becomes closer and closer to the lower static hedging bound. This indicates that although market operators use prices obtained from a dynamic hedging methodology, the observed tendency to push for time inhomogeneity and less parametric models tends to produce prices that are more conservative and closer to static hedging bounds.

We can apply the above program to basket pricing and solve for the optimal mixed static-dynamic hedging portfolio as in Avellaneda & Paras (1996).

Proposition 33 *The approximate optimal hedging portfolio in the sense of Avellaneda & Paras (1996) is composed of a static part with λ_k^{opt} baskets with payoffs $h_{\omega_k,K_k}(F_T)$ and a dynamic hedging strategy with $\Delta_t^k = \frac{\partial BS_0(Tr(\Omega_0 X^{opt}))}{\partial F_{k,0}}$, having computed λ_k^{opt} and X^{opt} from:*

$$\lambda_k^{opt} = -y_k^{opt} \frac{\partial BS_0\left(Tr(\Omega_0 X)\right)/\partial v}{\partial BS_k\left(Tr(\Omega_k X)\right)/\partial v}$$

where X^{opt} and y_k^{opt} are the primal and dual solutions to the semidefinite program:

$$\begin{array}{ll} \textit{maximize} & \sigma_{\max}^2 T = \mathbf{Tr}(\Omega_0 X) \\ \textit{s.t.} & \mathbf{Tr}(\Omega_k X) = \sigma_k^2 T_k \textit{ for } k = 1, \ldots, M \\ & X \succeq 0 \end{array}$$

Proof. We can write the portfolio optimization program as:

$$\inf_{\lambda \in \mathbb{R}^M} \left\{ \sum_{k=1}^m \lambda_k C_k + \left(\sup_P E^P \left[h_{\omega,k}(F_T) - \sum_{k=1}^m \lambda_k h_{\omega_k,K_k}(F_T) \right] \right) \right\} \tag{6.5}$$

where P varies within the set of equivalent martingale measure. We can rewrite the above problem as:

$$\inf_{\lambda}\left\{\sup_{P}\left(E^P\left[h_{\omega,0}(F_T)\right]-\sum_{i=1}^{m}\lambda_k\left(E^P\left[h_{\omega_k,K_k}(F_T)\right]-C_k\right)\right)\right\}$$

where we can recognize the optimum hedging portfolio problem as the dual of the maximum price problem above:

$$\begin{array}{ll}\text{maximize} & E^P\left[h_{\omega,K}(F_T)\right]\\ \text{s.t.} & E^P\left[h_{\omega_i,K_k}(F_T)\right]=C_k \text{ for } k=1,...,m\end{array}$$

We can find an approximate solution by solving the following problem:

$$\begin{array}{ll}\text{maximize} & BS_0(Tr(\Omega_0 X))\\ \text{s.t.} & BS_k\left(Tr(\Omega_k X)\right)=C_k \text{ for } k=1,...,m\\ & X\succeq 0\end{array}$$

and its dual:

$$\inf_{\lambda}\left\{\sup_{X\succeq 0}\left(BS(Tr(\Omega_0 X))-\sum_{k=1}^{m}\lambda_k\left(BS\left(Tr(\Omega_k X)\right)-C_k\right)\right)\right\}$$

where, for simplicity, we have noted $BS_k(v)$ the Black & Scholes (1973) price of basket k as a function of the cumulative variance $\sigma_k^2 T_k$. The primal problem, after we write it in terms of variance, becomes the following semidefinite program:

$$\begin{array}{ll}\text{maximize} & \sigma_{\max}^2 T=\mathbf{Tr}(\Omega_0 X)\\ \text{s.t.} & \mathbf{Tr}(\Omega_k X)=\sigma_k^2 T_k \text{ for } k=1,...,m\\ & X\succeq 0\end{array}$$

If we note $y^{opt}\in\mathbb{R}^m$ the solution to the dual of this last problem:

$$\begin{array}{ll}\text{minimize} & \sum_{k=1}^m y_k\sigma_k^2 T_k\\ \text{s.t.} & 0\preceq\sum_{k=1}^m y_k\Omega_k-\Omega_0\end{array}$$

We know (see Vandenberghe & Boyd (1996) for example) that the Karush-Kuhn-Tucker conditions on the primal-dual semidefinite program pair above can be written:

$$\left\{\begin{array}{l}0\preceq\sum_{k=1}^m y_k\Omega_k-\Omega_0\\ 0=\sum_{k=1}^m y_k\Omega_k X-\Omega_0 X\\ \mathbf{Tr}(\Omega_k X)=\sigma_k^2 T_k \text{ for } k=1,...,m\\ 0\preceq X\end{array}\right.$$

We can compare those to the KKT conditions for the price maximization problem:

$$\left\{\begin{array}{l}Z=\frac{\partial BS_0(\Omega_0 X)}{\partial v}\Omega_0+\sum_{k=1}^m\lambda_k\frac{\partial BS_k(\Omega_k X)}{\partial v}\Omega_k\\ XZ=0\\ BS_k\left(\mathbf{Tr}(\Omega_k X)\right)=C_i \text{ for } k=1,...,m\\ 0\preceq X,Z\end{array}\right.$$

with dual variables $\lambda\in\mathbb{R}^m$ and $Z\in\mathbf{S}^M$. An optimal dual solution for the price maximization problem can then be constructed from y^*, the optimal dual solution of the semidefinite program on the variance as:

$$\lambda_k^{opt}=-y_i^{opt}\frac{\partial BS_0\left(Tr(\Omega_0 X)\right)/\partial v}{\partial BS_k\left(Tr(\Omega_k X)\right)/\partial v}$$

and this gives the composition of the optimal static hedging portfolio in the baskets (ω_k, K_k). ■

6.2.2 The exact problem

The bounds found in the section above are only approximate solutions to the superreplicating problem. Although the relative error in this approximation is known to be about 1-2%, it is interesting to notice that although it does not remain completely tractable, the exact problem shares the same optimization structure as the approximate one. In fact, we will see below that the optimization problem in the exact Black-Scholes-Barenblatt equation retains most of the structure of the approximate one. Let us recall the results in Romagnoli & Vargiolu (2000). If we note $C(F_t, t)$ the superreplicating price of a basket option, then $C(F_t, t)$ is the solution to the following multidimensional Black-Scholes-Barenblatt equation:

$$\begin{cases} \frac{\partial C(x,t)}{\partial t} + \frac{1}{2}\max_{\gamma\in\Lambda} \mathbf{Tr}\left(\frac{\partial^2 C(x,t)}{\partial x^2}\left(\bar{x}\gamma\right)^T\left(\bar{x}\gamma\right)\right) = 0 \\ C(x,T) = \left(\sum_{i=1}^n \omega_i x - k\right)^+ \end{cases}$$

where $\bar{x}$ is the diagonal matrix formed with the components of x.

We can create a superreplicating strategy by dynamically trading in a portfolio composed of $\Delta_t^i = \frac{\partial C}{\partial x_i}(t, F_t^i)$ of each asset in the basket. The volatility optimization problem embedded in the BSB equation above has been studied by Romagnoli & Vargiolu (2000).

We can rewrite it in a format that is closer to that of the approximate problem above and it becomes:

$$\begin{cases} \frac{\partial C(x,t)}{\partial t} + \frac{1}{2}\max_{\Gamma\in\Lambda} \mathbf{Tr}\left(\bar{x}\frac{\partial^2 C(x,t)}{\partial x^2}\bar{x}\Gamma\right) = 0 \\ C(x,T) = \left(\sum_{i=1}^n \omega_i x - k\right)^+ \end{cases}$$

where we have noted $\Gamma = \gamma\gamma^T$ the model covariance matrix. If the set Λ is given by the intersection of the semidefinite cone (the covariance matrix has to be p.s.d.) with a polyhedra (for example approximate price constraints, sign constraints or bounds on the matrix coefficients, ...), then the embedded optimization problem becomes a semidefinite program:

$$\text{maximize}_{\ \Gamma\in\Lambda}\ \mathbf{Tr}\left(\Gamma\bar{x}\frac{\partial^2 C(x,t)}{\partial x^2}\bar{x}\right)$$

on the feasible set Λ. Under the conditions above, the general upper-lower hedging price is then computed by solving:

$$\begin{cases} \frac{\partial C(x,t)}{\partial t} + \frac{1}{2}\max_{\Gamma\in\Lambda} \mathbf{Tr}\left(\bar{x}\frac{\partial^2 C(x,t)}{\partial x^2}\bar{x}\Gamma\right) = 0 \\ C(x,T) = \left(\sum_{i=1}^n \omega_i x - k\right)^+ \end{cases}$$

where the embedded problem is a semidefinite program. This is simply a rearrangement of the equation in Romagnoli & Vargiolu (2000) where we have parametrized the embedded optimization problem using the covariance matrix.

We recover the same optimization problem as in the approximate solution found in the section above, the only difference being here that the solution to the exact general problem might not be equal to a Black-Scholes price. This gives a convex (and much simpler) formulation of the embedded problem in Romagnoli & Vargiolu (2000).

6.3 Optimal Gamma Hedging

We study here the problem of optimally adjusting the Gamma of a portfolio using only options on single assets. This problem is essentially motivated by a difference in liquidity between the vanilla and basket option markets, which makes it impractical to use some baskets in the hedging portfolio. Suppose we have an initial portfolio with a Gamma sensitivity matrix given by Γ in a market with underlying assets $x_i, i = 1, \dots n$. We want to hedge this position with y_i vanilla options on each single asset x_i with Gamma given by γ_i We assume that the portfolio is maintained delta-neutral at all times using the appropriate proportions of each particular stock. A small perturbation of the stock price will induce a change in the portfolio price given by:

$$\Delta P(X + \Delta X) = P(X) + \frac{1}{2}\Delta S^T \Gamma(y) \Delta S$$

where $\Gamma(y) = \Gamma + diag(\gamma)y$, with $diag(\gamma)$ the diagonal matrix with components γ_i. As in Douady (1995), our objective is to minimize in y the maximum possible perturbation given by:

$$\max_{\Delta S \in E} \left| \Delta S^T \Gamma(y) \Delta S \right|$$

where E is the ellipsoid defined by

$$E = \left\{ X \in \mathbb{R}^n | X^T \Sigma X = 1 \right\} \text{ with } \Sigma = (cov(x_i, x_j))_{i,j=1,\dots,n}$$

the covariance matrix of the underlying assets. This is amounts to minimizing the maximum eigenvalue of the matrix $\Sigma\Gamma(y)$.

Proposition 34 *The optimum Gamma hedging portfolio can be solved by the following semidefinite program:*

$$\begin{array}{ll} \textit{minimize} & t \\ \textit{subject to} & -tI \preceq \Sigma\Gamma + \Sigma diag(\gamma) y \preceq tI \end{array}$$

Proof. Because we know that $\Sigma\Gamma + \Sigma diag(\gamma)y \preceq tId \Longleftrightarrow \max_k \lambda_k(\Sigma\Gamma + \Sigma diag(\gamma)y) \leq t$, hence in the above problem we have $\max_k |\lambda_k(\Sigma\Gamma + \Sigma diag(\gamma)y)| = \|\Sigma\Gamma + \Sigma diag(\gamma)y\|_2 \leq t$. ■

This can be solved very efficiently using solvers such as the one by Sturm (1999). We can balance this with constraints on the cost of the hedge. Suppose that there are proportional transaction costs associated with trading in the vanilla option on x_i given by $k_i |y_i|$ for some $k_i \geq 0$. The problem becomes is a symmetric cone program given by:

$$\begin{array}{ll} \text{minimize} & t_0 + \alpha \sum_{i=1}^n k_i t_i \\ \text{subject to} & -t_0 I \preceq \Sigma\Gamma + \Sigma diag(\gamma) y \preceq t_0 I \\ & -t_i \leq y_i \leq t_i \end{array}$$

and can also be solved using the code by Sturm (1999). The parameter α describes the relative importance of minimizing the cost of the hedge compared to minimizing the gamma.

6.4 Static hedging

The objective of this section is to provide an additional piece in the calibration toolbox that this work is meant to create. As the model itself is allowed to be significantly more flexible by letting the variance covariance matrix vary among the full set of semidefinite positive matrixes instead of the rank one or two matrixes, an important number of additional swaptions can be incorporated in the calibration program. In practice, this has the direct effect of reducing the number of times when a swaption cannot be incorporated into the calibration set because it makes the problem infeasible (this would qualify as an arbitrage if the market was following the dynamics given by the LIBOR market model). The objective of the section that follows is to further reduce the frequency of such problems by detecting the swaption prices that are not statically-arbitrage free (due to a data error or a false extrapolated input from the volatility stripping) and eliminate them from the calibration set. The bounds obtained here will of course be much wider than the one discussed trough dynamic arbitrage in the previous chapters, but they still provide some structural insights into the relative price structure of caps and swaptions.

The classical setup that has been formulated by Black-Scholes Black & Scholes (1973) and Merton Merton (1973) computes the unique price of an option as the price today of a dynamic portfolio strategy that perfectly replicates the option payoff at maturity. This "dynamic hedging" approach relies of course heavily on the hypothesis that the markets are complete and frictionless, i.e. that traders can instantaneously adjust their hedging portfolio at no cost and perfectly replicate any contingent claim. This is of course not verified in reality but the Black-Scholes hedging procedure performs sufficiently well in practice to make the (dynamic) arbitrage-free price of a product a fundamental reference (see for example El Karoui et al. (1998)). As we have already seen in a previous section, if the markets are incomplete and not all payoffs are attainable, the arbitrage argument does not provide a unique price but rather upper and lower bounds inside of which the absence of arbitrage is guaranteed. Those bounds correspond to the price of a dynamic strategy that almost surely dominates (resp. is dominated by) the payoff of the contingent claim (see El Karoui & Quenez (1991), El Karoui & Quenez (1995) or Karatzas & Shreve (1998)). Within this price interval, the choice of a particular trading price is defined by further assumptions on the agent preferences (Davis (1994) or Hodges & Neuberger (1989)).

All the pricing and hedging methods above suppose that it is possible to form a dynamic hedging portfolio and provide precise constraints on contingent claims prices given that market are complete and frictionless. In this paper, we are interested in the pricing results that can be obtained with a very minimum set of assumptions on the market setup, hence we will explore the information that can be inferred from the market by static arbitrage instead of dynamic ones. We call *static* arbitrage an arbitrage that can be realized by investing in a portfolio at a particular time and trading at only a few dates in the future (typically two here). The static arbitrage arguments produce pricing and hedging results that are much more robust to market imperfection or in other words produce much *harder price bounds* which are wider than their dynamic counterparts. The objective of this approach is thus more to give a reliable reference tool in the analysis of the relative structure of market prices rather than a pricing methodology. Of course this static arbitrage approach can be seen as a particular case of the dynamic hedging methodology with an additional restriction on the intensity of trading, however the theory behind dynamic arbitrage does not allow the addition of direct restrictions on the trading intensity in a tractable way. The setup we present here is thus adapted to accommodate these restrictions.

The idea of inferring information from option prices dates back at least to Breeden & Litzenberger (1978) and was adapted to a diffusion model $\grave{a}\ la$ Black-Scholes by Dupire (1994), both these papers suppose that a continuum of option prices is observed. More recently Edirisinghe, Naik & Uppal (1993) use a linear programming approach to compute the upper and lower hedging prices of European contingent claims in a dynamic framework. Jackwerth & Rubinstein (1996) solve the problem of computing minimum entropy distributions from market data. Laurent & Leisen (2000) define the necessary and sufficient conditions on the prices of a set of European Options for the absence of dynamic arbitrage. In a paper that is focused on moment constraints under the risk-neutral probability, Bertsimas & Popescu (2002) solve the same problem in a one-dimensional framework and propose an approximate solution based on a relaxation of the infinite dimensional problem[1].

In this section, we directly deduce the conditions for the absence of static arbitrage by adapting the setup in Edirisinghe et al. (1993) and Laurent & Leisen (2000), we then extend these results to options on a basket of assets. Our basic market price information set will be composed of individual assets and some options of different strikes and maturities written on those assets. We use the information inferred from this set of instruments to compute static arbitrage price bounds on other European contingent claims. In the case where a static arbitrage strategy exists, we compute the corresponding portfolio as the solution to the dual problem. We also study the static arbitrage conditions on option prices in the presence of transaction costs and we describe how these results can be extended to some of the classical dynamic, continuous-time models. Our key argument will be that the correct pricing of "butterfly" and "calendar" spreads (see Merton (1973), this will be further detailed below) together with a few additional linear constraints is a necessary and sufficient condition for the absence of static arbitrage between European options.

We start by recalling the classical one-dimensional result and we then extend it to basket options.

6.4.1 Necessary conditions

Let us note S_t the price process of a given asset and let $C(K,T)$ be the price of an European Call option written on this asset with strike K and maturity T. Suppose that we are given the market price of a set of such options:

$$C(K_i,T_i)=p_i \text{ for } i=1,...,M$$

In this section we will derive the two necessary conditions on this set of Call prices that preclude the existence of static arbitrage. Following the ideas in Merton (1973) and Laurent & Leisen (2000), these conditions are derived from the pricing of two basic trading strategies.

The "butterfly" spread

A butterfly spread is a portfolio of three call options with a common maturity date T. If we note K_1,K_2,K_3 the strike prices of the options, if $K_1<K_3$ we can form a butterfly spread by buying one of each options at strike K_1 and K_2 and selling short two options at a strike $K_2=(K_1+K_3)/2$. If S_T is the value of the underlying asset at time T, the payoff at maturity T of the butterfly spread is then given by:

$$(S_T-K_1)^+-2(S_T-K_2)^++(S_T-K_3)^+$$

Because the payoff at maturity of this product is always positive, its price at time zero must also be positive to avoid arbitrage, hence we must have:

$$C(K_1,T)-2C(K_2,T)+C(K_3,T)\geq 0 \text{ for } T>0$$

if the market data is made of a smooth continuum of Call prices, we can rewrite the constraint above as:

$$\frac{\partial^2 C(K,T)}{\partial K^2}\geq 0 \text{ for } T>0 \tag{6.6}$$

in other words, the static arbitrage free pricing of the butterfly spreads imposes that the Call prices be convex with respect to the strike price.

[1] Note that both Laurent & Leisen (2000) and Bertsimas & Popescu (2002) do not mention the lower bound constraint $C(K,T)\geq S_0-B(0,T)K$ because their static market does not always include the forward.

The ”calendar” spread

We form a calendar spread as a portfolio of two options with a common strike K. If we note T_1, T_2 the option maturities, with $T_1 < T_2$, we can form a calendar spread by buying the Call with maturity T_2 and selling short the Call with maturity T_1.

Following the ideas of Merton (1973), let us suppose now that $C(K,T_1) > C(K,T_2)$. We can ”invest” in a calendar spread at time zero and receive a strictly positive cash flow. At time T_1, if the asset price is greater than the strike K, we exercise the Call, if not, we do nothing. We carry all proceedings up to maturity T_2, if for simplicity we suppose that interest rates are zero, this generates a payoff of:

$$(S_{t+1} - K)^+ - 1_{\{S_t > K\}}(S_{t+1} - K) \geq 0$$

and this strategy constitutes an arbitrage (strictly positive cash flow at date zero, zero cash flow at time T_1, positive cash flow at time T_2), we can conclude that we must have:

$$C(K,T_1) < C(K,T_2)$$

or again, with a smooth continuum of Calls:

$$\frac{\partial C(K,T)}{\partial T} \geq 0 \text{ for } K > 0 \tag{6.7}$$

and as the above strategy involves only trading at two dates, we show that the static arbitrage free price of Call options must be increasing with maturity.

Convex conditions

For simplicity we assume again that a smooth continuum of Calls $C(K,T)$ is observed from the market and we only add the trivial condition that Call prices be decreasing with respect to strike,

Proposition 35 *The necessary conditions for the absence of static arbitrage become:*

$C(K,T) \geq S_0 - B(0,T)K$
$C(K,T)$ *nondecreasing in* T
$C(K,T)$ *convex, nonincreasing in* K
C *homogeneous of degree one*

Proof. See Laurent & Leisen (2000), except for the omitted condition $C(K,T) \geq S_0 - B(0,T)K$. ■

Because the set of Call prices verifying these constraints is the intersection of three convex sets (both the set of decreasing and the set of convex functions are convex cones), it must be convex, the cone property is simply a direct consequence of the invariance with respect to a change of numeraire.

6.4.2 Non-parametric bounds on the price of baskets

We can now exploit the two conditions above to find price bounds on European contingent claims that are not priced by the market. From the results above, we know that if a contingent claim's price falls out of these bounds there is an arbitrage and we can form a static portfolio to take advantage of it. In fact, we can extend these results to basket options and give necessary and sufficient conditions for the absence of static arbitrage between baskets. Let us start by showing the result with a single maturity date T.

And as above, the no arbitrage condition between options of different maturity dates is equivalent to the correct pricing of calendar spreads. This allows us to formulate the general result on the absence of static arbitrage between basket Call options.

Proposition 36 *For a fixed maturity $T > 0$, let us note $C(\omega, T, K)$ with $\omega \in \mathbb{R}^n$ and $K \in \mathbb{R}$, the price today of a Call basket option on the assets $S_t^i, i = 1, ..., n$ with payoff:*

$$\left(\sum_{i=1}^{n} \omega_i S_T^i - K\right)^+$$

then there are no static arbitrage opportunities iff the function $C(\omega, T, K) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^+$ verifies:

$$\begin{array}{l} C(\omega, 0, T) = \omega^T S_t \\ C(\omega, K, T) \geq \sum_{i=1}^{n} \omega_i S_0^i - B(0,T)K \\ C(\omega, K, T) \text{ convex in } (\omega, K)\text{, nonincreasing in } K \\ C(\omega, K, T) \text{ nondecreasing in } \omega\text{, nonincreasing in } K \\ C(\omega, K, T) \text{ nondecreasing in } T \\ C(\omega, K, T) \text{ homogeneous of degree one in } (\omega, K) \end{array}$$

where $B(0, T)$ is the price of a zero coupon bond with maturity T, as quoted by the market today.

We now notice that the program above can be solved exactly.

Proposition 37 *The infinite program above can be discretized exactly into the following program:*

$$\begin{array}{rll} v_{red}(D) = & \min & p_0 \\ & \text{s.t.} & \langle g_i, (\omega_j, K_j) - (\omega_i, K_i) \rangle \leq p_j - p_i \text{ for } i, j = 0, ..., k \\ & & g_{i,j} \geq 0, g_{i,n+1} \leq 0 \text{ for } i = 0, ..., k \text{ and } j = 1, ..., n \\ & & \langle g_i, (\omega_i, K_i) \rangle = p_i \text{ for } i, j = 0, ..., k \end{array} \tag{6.8}$$

which is a finite L.P. in the variables $p_0 \in \mathbb{R}_+$ and $g_i \in \mathbb{R}^{n+1}$ for $i = 0, ..., k$.

Proof. We first notice that as a discretization of the infinite program, the finite L.P. will compute a lower bound on the its optimal value. Let us now show that this lower bound is actually attained. If we note $z^* = \left[p_0^*, g_0^{*T}, \ldots, g_k^{*T}\right]^T$ the optimal solution to the LP problem above and if we define:

$$s(x) = \max_{i=0,\ldots,k} \{p_i^* + \langle g_i^*, x - x_i \rangle\}$$

$s(x)$ verifies

$$s(x_i) = p_i, \quad i = 1, \ldots, k$$

and, by construction, $s(x_0)$ attains the lower bound p_0 computed in the finite L.P.. Because $s(x)$ is convex as the pointwise maximum of affine functions and is piecewise affine with gradient g_i, which implies that it also verifies the monotonicity conditions, it is a feasible point of the infinite dimensional problem. Finally the condition $\langle g_i, (\omega_i, K_i) \rangle = p_i$ guarantees that the solution is homogeneous, hence both problems have the same optimal value and $s(x)$ is an optimal solution to the Infinite Linear Program. ■

This provides both an upper and a lower static arbitrage bound on the price of a basket. As in d'Aspremont & El Ghaoui (2002), it can be showed that these bounds are sharp in some cases.

Part III

Numerical performance

Chapter 7

Approximation precision

7.1 Generic Basket pricing

We look here at the approximation error for generic multidimensional Black & Scholes (1973) basket prices. Here, as in figure (3.2), we plot the error in the basket pricing formula for a basket of assets, having supposed that the forwards are all martingale under the same probability measure (hence we test the precision of the approximations without the error from the forward measures). The reference is given by a Monte-Carlo estimate with enough steps to make the confidence interval less than one basis point. The numerical values used here in figure (7.1) are $F_0^i = \{0.7, 0.5, 0.4, 0.4, 0.4\}$, $\omega_i = \{0.2, 0.2, 0.2, 0.2, 0.2\}$, $T = 5$ years, and the covariance matrix is given by:

$$\frac{11}{100}\begin{pmatrix} 0.64 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0.64 & 0 & 0 \\ 0 & 0 & 0 & 0.36 & 0 \\ 0 & 0 & 0 & 0 & 0.16 \end{pmatrix}$$

In the covariance above, we have supposed that the rates are independent. Here again, these values are meant to replicate the pricing of a 5Y into 5Y swaption if we neglect the change of measure and the weight dynamics. We can see that the pricing error is consistent with that found in (3.2). Here however the order one term provides a significant improvement over the order zero price.

We can also look at the error in a standard equity basket (with forwards closer to 100). In figure (7.2) The numerical values used here in figure (7.1) are $F_0^i = \{100, 80, 120, 40, 15\}$, $\omega_i = \{0.4, 0.1, 0.1, 0.2, 0.2\}$, $T = 2$ years, and the covariance matrix is again given by:

$$\frac{11}{100}\begin{pmatrix} 0.64 & 0.59 & 0.32 & 0.12 & 0.06 \\ 0.59 & 1 & 0.67 & 0.28 & 0.13 \\ 0.32 & 0.67 & 0.64 & 0.29 & 0.14 \\ 0.12 & 0.28 & 0.29 & 0.36 & 0.11 \\ 0.06 & 0.13 & 0.14 & 0.11 & 0.16 \end{pmatrix}$$

We have computed the relative pricing error here and we can see that in the pure equity case the order one correction provides a very significant improvement over the order zero price.

7.2 Swaption pricing in the Libor Market model

As opposed to the previous section, here we look at actual Libor market prices obtained by Monte-Carlo (including the exact change of measure and weights dynamics). As in figure (3.1), we present

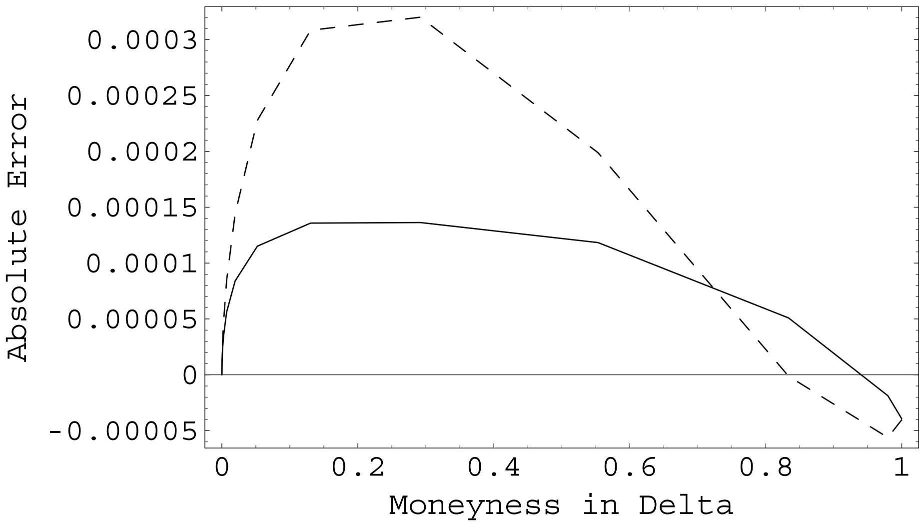

Figure 7.1: Order zero (dashed) and order one (plain) absolute approximation error versus the multidimensional Black-Scholes basket prices obtained by simulation for various strikes.

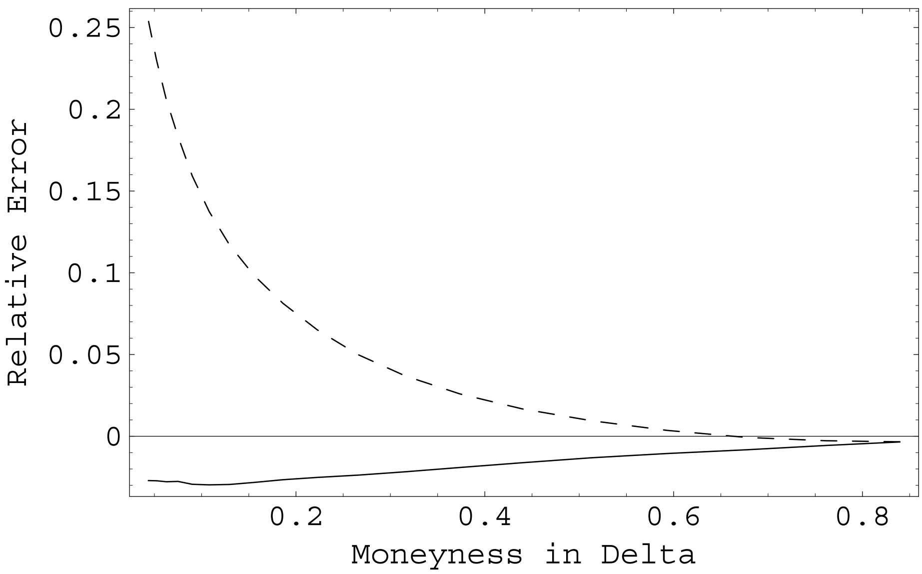

Figure 7.2: An equity basket example: Order zero (dashed) and order one (plain) relative approximation error versus the multidimensional Black-Scholes basket prices obtained by simulation for various strikes.

a plot of the difference between two distinct sets of swaption prices in the Libor Market Model. One is obtained by Monte-Carlo simulation using enough steps to make the 95% confidence margin of error always less than 1bp. The second set of prices is computed using the order zero approximation formula above. In figures (7.3), (7.4) and (7.5) however, we look at the pricing error when the strike

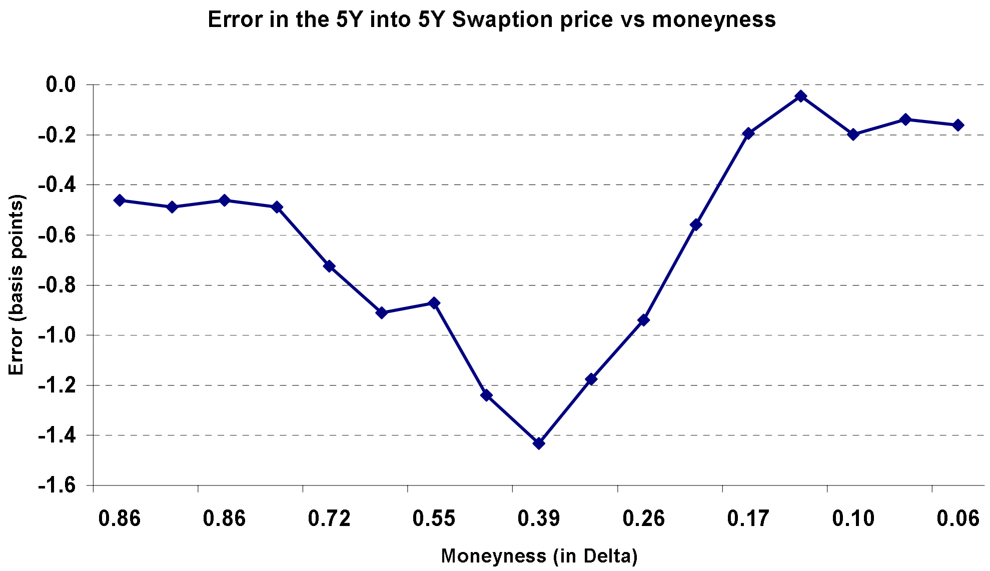

Figure 7.3: Absolute error on the 5Y into 5Y swaption price for various strikes. We compare the order zero price approximation versus the Libor market model prices estimated using Monte-Carlo simulation.

varies. Again, this plot is based on the prices obtained by calibrating the model to EURO swaption prices on November 6 2000. We have used all cap volatilities and the following swaptions: 2Y into 5Y, 5Y into 5Y, 5Y into 2Y, 10Y into 5Y, 7Y into 5Y, 10Y into 2Y, 10Y into 7Y, 2Y into 2Y, 1Y into 9Y.

7.3 Market Implied Covariance Factor Analysis

We use the data set from Nov. 6 2000 and we calibrate by fitting all caplets up to 20 years plus the following set of swaptions: 5Y into 5Y, 5Y into 2Y, 5Y into 10Y, 2Y into 2Y, 2Y into 5Y, 7Y into 5Y, 10Y into 5Y, 10Y into 2Y, 10Y into 10Y, 7Y into 3Y, 4Y into 6Y, 17Y into 3Y (again the motivation behind this choice of swaptions is liquidity). For simplicity, all frequencies are annual. The resulting covariance matrix is plotted in figure (7.6). In figure (7.7) and (7.8) we plot the eigenvectors of this matrix. The first vector has a level shape. The second one is close to a spread of rates. We notice, this purely market implied covariance factor structure closely matches the results obtained by historical time series analysis.

7.4 Relative caps-swaption prices

We use again the data set from Nov. 6 2000 and we plot in figure (7.9) the upper and lower bound given by maximizing (resp. minimizing) the volatility of a given swaption provided that the Libor

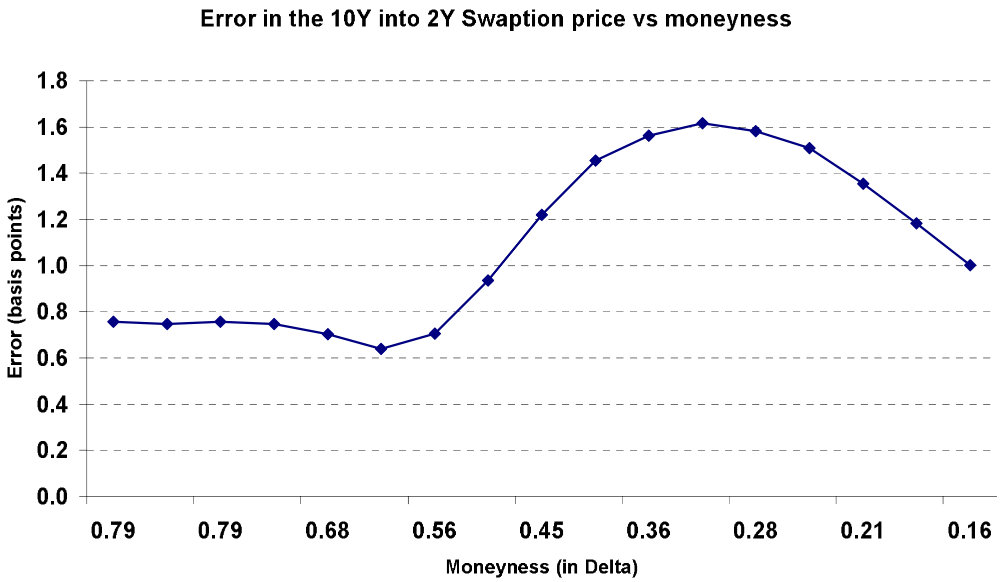

Figure 7.4: Absolute error on the 10Y into 2Y swaption price for various strikes. We compare the order zero price approximation versus the Libor market model prices estimated using Monte-Carlo simulation.

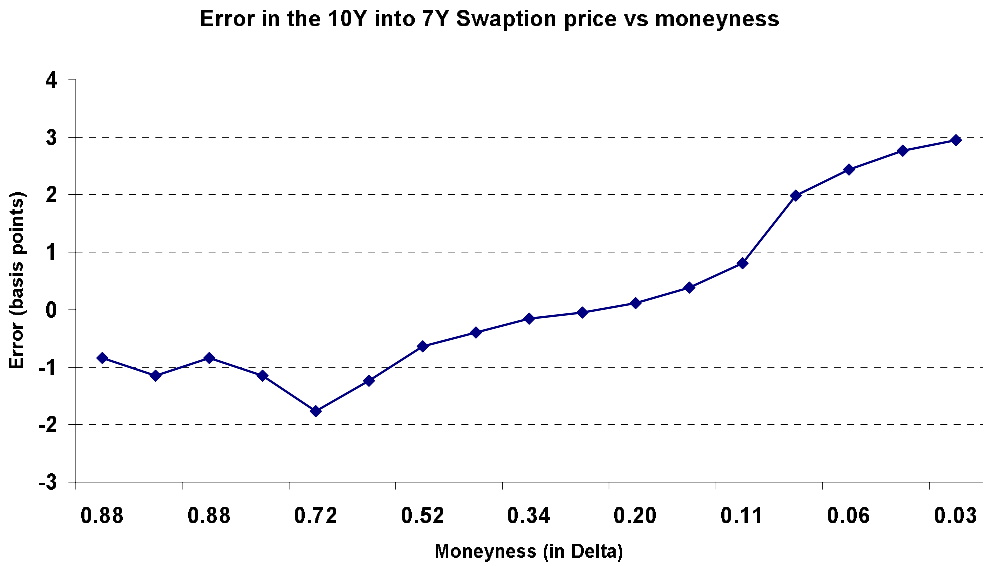

Figure 7.5: Absolute error on the 10Y into 7Y swaption price for various strikes. We compare the order zero price approximation versus the Libor market model prices estimated using Monte-Carlo simulation.

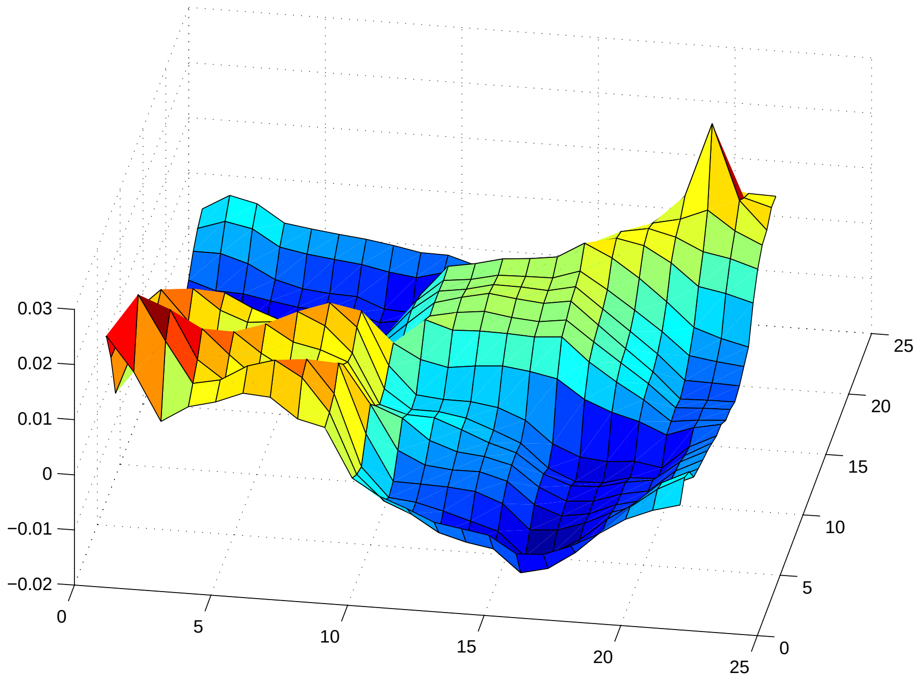

Figure 7.6: Smooth calibrated covariance matrix.

covariance matrix remains positive semidefinite and that it matches the calibration instruments. We calibrate on the same products as in the last section. The result is plotted in the first figure.

Quite surprisingly considering the simplicity of the model (stationarity of the sliding dynamics in $L(t, \theta)$, the converging dynamics are of course not stationary), figure (7.9) shows that all swaptions seem to fit reasonably well inside the bounds imposed by the model except for the 10Y underlying, this in line with the findings of Longstaff et al. (2000). We can also calibrate by fitting all caplets up to 20 years and a smaller (more liquid) set of swaptions: 2Y into 5Y, 5Y into 5Y, 5Y into 2Y, 10Y into 5Y, 7Y into 5Y, 10Y into 2Y, 10Y into 7Y, 2Y into 2Y. This choice of swaptions was motivated by liquidity (where all swaptions on underlying and maturity in 2Y, 5Y, 7Y, 10Y are meant to be liquid). We plot the result in figure (7.10). Table (7.1) details the market caplet volatilities, while table (7.2) shows the swaption volatilities and the corresponding $\hat{\omega}_i$ weights (Sep. 03 2000, data courtesy of BNP Paribas, London).

Quite surprisingly considering the simplicity of the model (stationarity of the *sliding* Libor dynamics $L(t, \theta)$), figure (7.9) shows that all swaptions seem to fit reasonably well in the bounds imposed by the model except for the 10Y underlying. This is in line with the findings of Longstaff et al. (2000). In table (7.3) and (7.4), we show the market volatility movement vector with largest impact on the covariance matrix (first vector in the singular value decomposition of the sensitivity matrix in 5.12), computed in the A.H.O. case, using the same dataset above and a minimum trace objective.

We can also study the evolution of the price bounds on a particular swaption (the 5Y into 3Y) as more and more instruments are added in the calibration set (in that sense, we plot a graph that is ”transversal” to the one in the first figure). The initial calibration set is composed of all caps and the 2Y into 5Y swaption, it then evolves as in figure (7.11), where the flat line between the upper and lower bounds represents the actual market price of the 5Y into 3Y swaption.

Caplet Vols (%, 1Y to 10Y)	14.3	15.6	15.4	15.1	14.8	14.5	14.2	14.0	13.9	13.3
Caplet Vols (%, 11Y to 20Y)	13.0	12.7	12.4	12.2	12.0	11.9	11.8	11.8	11.7	12.0

Table 7.1: Caplet volatilities.

Swaption	**Vol (%)**	$\hat{\omega}_i$						
2Y into 5Y	12.4	0.22	0.20	0.20	0.19	0.18		
5Y into 5Y	11.7	0.22	0.21	0.20	0.19	0.18		
5Y into 2Y	14.0	0.51	0.49					
10Y into 5Y	10.0	0.22	0.21	0.20	0.19	0.18		
7Y into 5Y	11.0	0.23	0.21	0.20	0.19	0.18		
10Y into 2Y	12.2	0.51	0.49					
10Y into 7Y	9.6	0.17	0.16	0.15	0.14	0.13	0.13	0.12
2Y into 2Y	14.8	0.52	0.48					

Table 7.2: Swaption volatilities and weights.

Caplet (1Y to 10Y)	0.00	-0.01	0.01	0.21	-0.25	-0.08	0.04	-0.17	-0.09	0.18
Caplet (11Y to 20Y)	0.15	-0.01	0.19	-0.18	0.29	0.04	0.09	-0.52	0.09	0.00

Table 7.3: Maximum sensitivity vector: caplet components.

Swaption	(2Y,5Y)	(5Y,5Y)	(5Y,2Y)	(10Y,5Y)	(7Y,5Y)	(10Y,2Y)	(10Y,7Y)	(2Y,2Y)
Sensitivity	-0.05	0.16	0.28	0.17	-0.18	-0.38	0.15	-0.08

Table 7.4: Maximum sensitivity vector: swaption components.

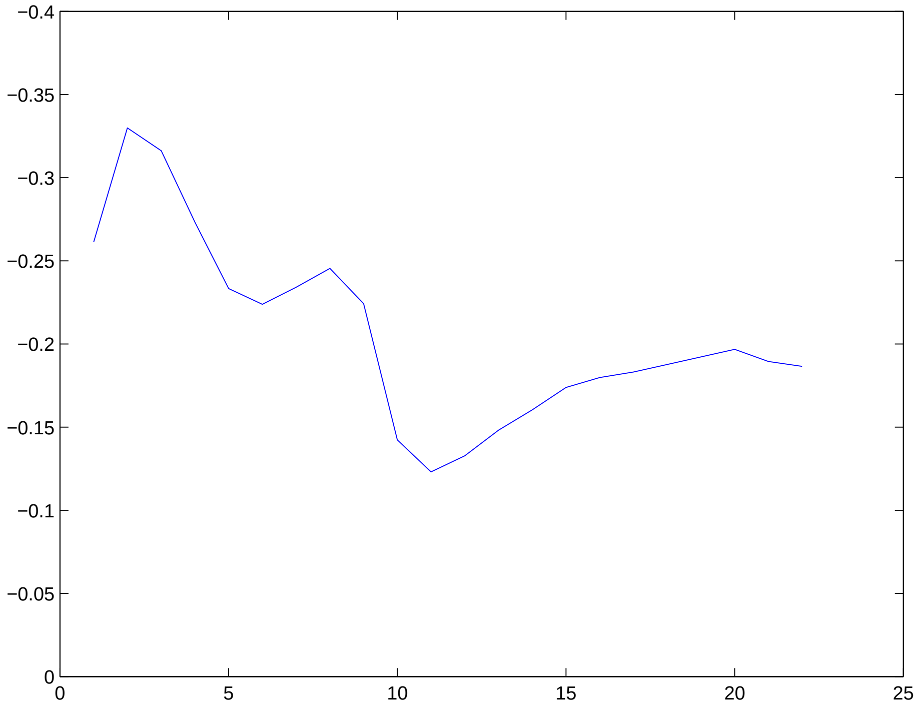

Figure 7.7: First eigenvector ”level”.

In the last two figures we study the sensitivity of the bounds obtained with the entire calibration set by looking at the respective dual solutions.

We have plotted the sensitivity versus a particular swaption in the calibration set. The singular behavior of the 7Y into 5Y swaption in the calibration set is probably due to its particular pricing by the market that day, making it particularly attractive component of the optimal hedging portfolio in terms of price vs. diversification.

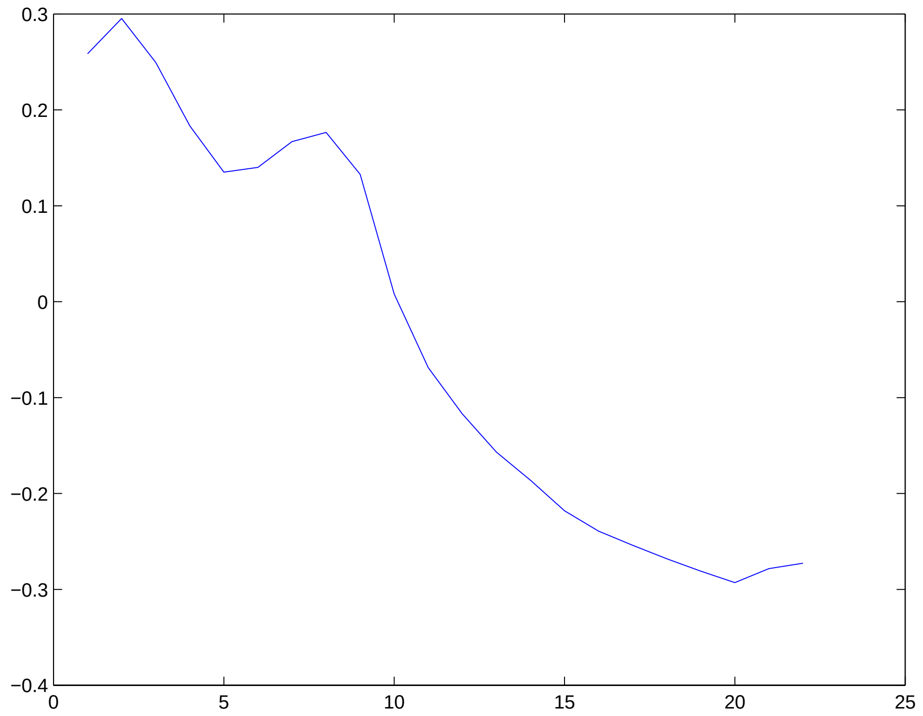

Figure 7.8: Second eigenvector ”spread”

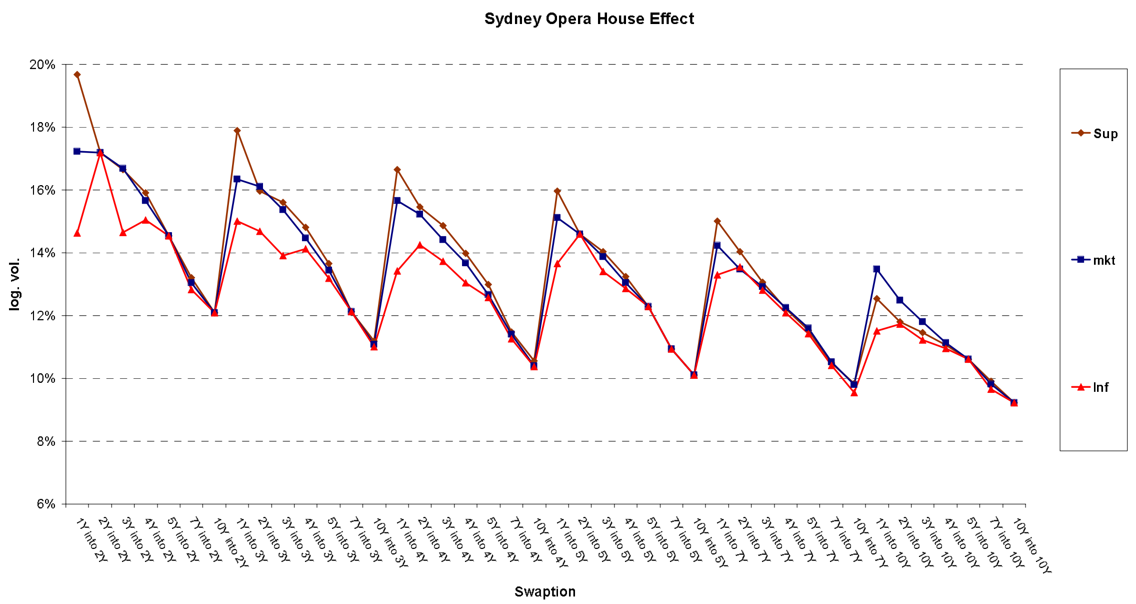

Figure 7.9: Calibration result and price bounds on a ”Sidney opera house” set of swaptions.

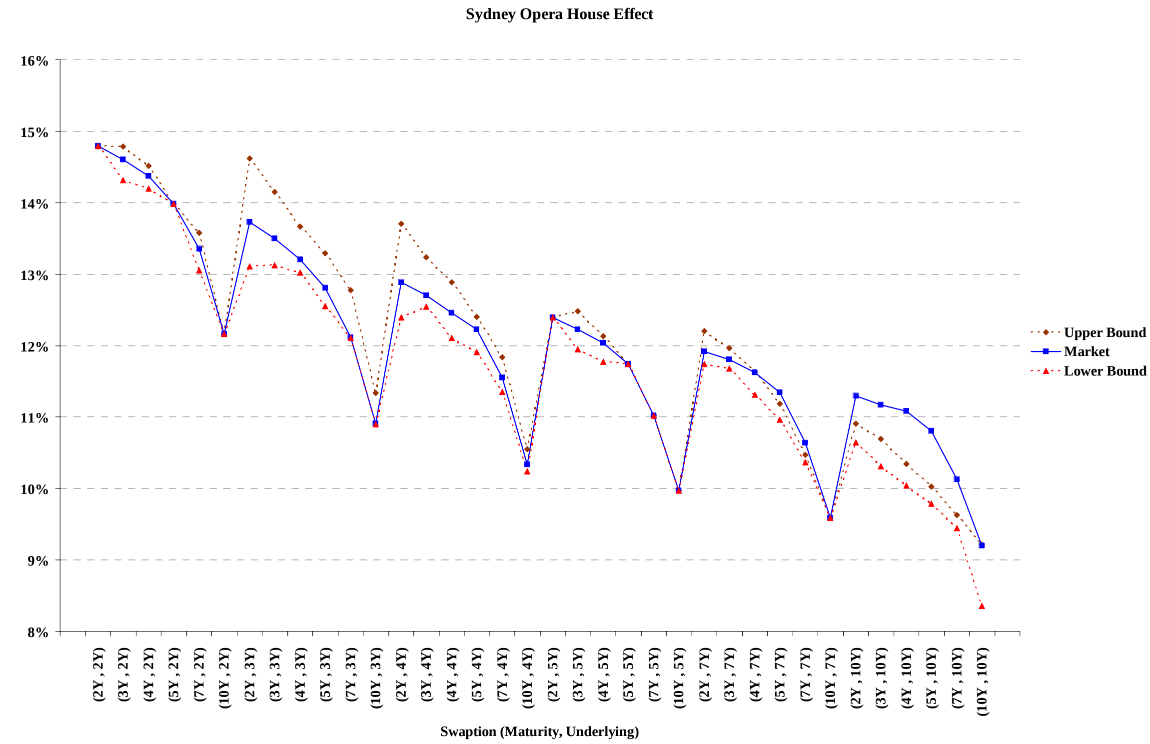

Figure 7.10: Calibration result and price bounds on a "Sydney opera house" set of swaptions.

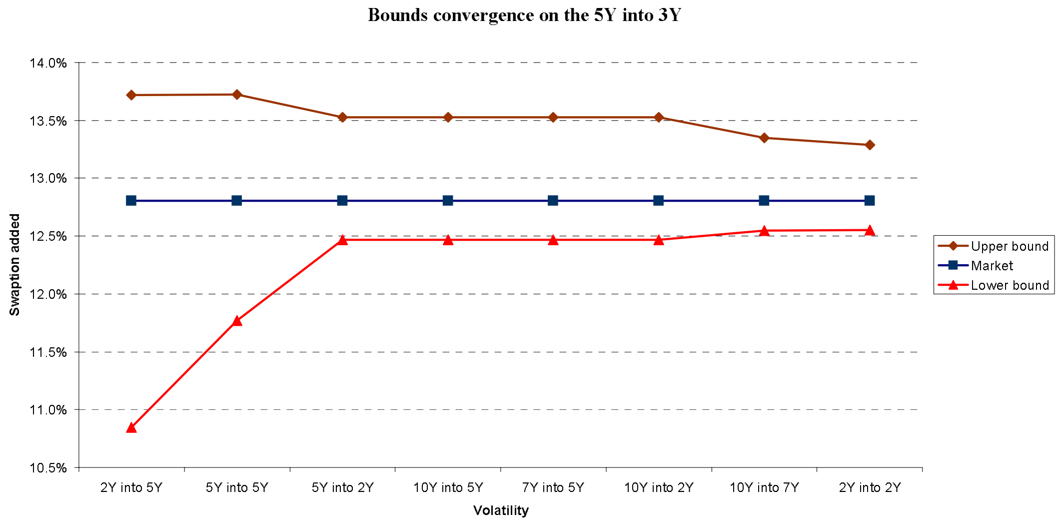

Figure 7.11: Convergence of the price bounds as more and more swaptions are added.

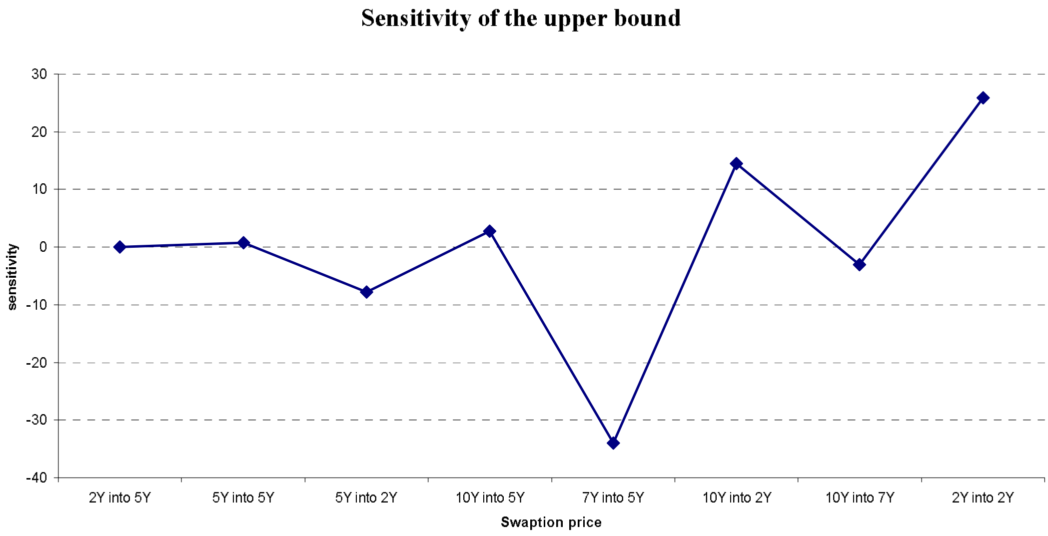

Figure 7.12: Dual solution to the upper bound problem

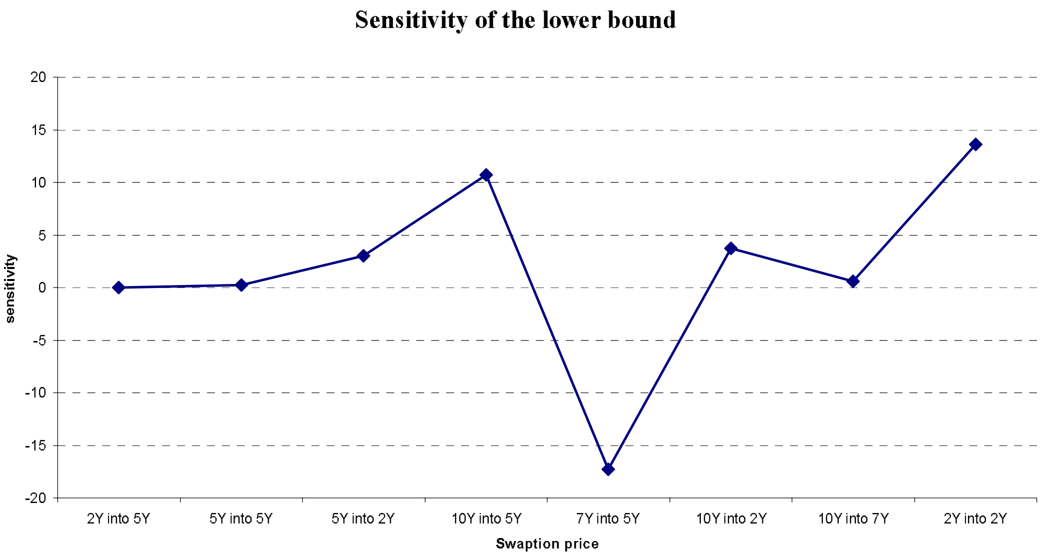

Figure 7.13: Dual solution to the lower bound problem

Chapter 8

Conclusion

The methods described in this work are organized around one central objective: the design of a true ”black-box” calibration and risk-management tool for classic multifactor interest rate models. In particular, the performance guarantee given by the numerical methods used here makes it possible to design a calibration procedure that does not require numerical baby-sitting. Furthermore, the possibility of stabilizing the calibration result should induce significant savings in hedging transaction costs by suppressing the possibility of purely numerical P&L hikes.

In practice however, two important obstacles remain in the design of a ”Swiss army knife” interest rate model: smile modelling and rank reduction. The first problem has already been mentioned in the first part, it is at this point not possible to globally calibrate the model to the smile and to the covariance structure, instead, one has to apply a two-step procedure to first calibrate the correct smile structure and then recover the covariance information. This makes it impossible to jointly optimize the calibration result on the smile and the covariance structure (smoothness, etc...). The second problem is rank reduction: numerical methods for American-style securities pricing are only efficient for models with a small number of factors. This makes rank reduction a backward compatibility problem: the semidefinite programming methods detailed above cannot guarantee a minimum rank for the solution. In fact, we know that the minimum rank problem is NP-hard. However, recent advances in quantization methods (see Bally & Pages (2000)) or American Monte-Carlo (see Longstaff & Schwartz (1998) for example) make it reasonable to expect significant progress in this area.

Part IV

Appendix

8.1 Second order term in the basket price approximation

Here, we show how we can apply the same expansion technique as in Chapter 2 to compute the second-order term in the basket option price expansion. We notice that the computations become a bit heavy and at this point it would probably be easier to compute this term and all the following ones using a formal calculus software such as Mathematica.

We look here for the second-order term using a procedure that exactly mimics the one used for $C^{(1)}$. We start by writing the P.D.E. solved by the second order term in the development.

Lemma 38 *The second order term $C^{(2)}(s,x,y)$ can be computed by solving:*

$$\begin{cases} 0 = L_0^0 C^{(2)} + 2L_1^0 C^{(1)} + L_2 C^0 \\ \\ 0 = C^{(2)} \textit{ for } s = T \end{cases} \tag{8.1}$$

With the terms $L_1^0 C^{(1)}$ and $L_2 C^0$ given by:

$$\begin{aligned} L_1^0 C^{(1)} \quad = \quad & \left(2\sum_{j=1}^n y_j \xi_s^j\right) \frac{x^2}{2} \frac{\partial^2 C^{(1)}}{\partial x^2} + \sum_{j=1}^n \left(\sum_{k=1}^n \left\langle \xi_s^j - \sigma_s^\omega, \xi_s^k \right\rangle\right) x y_j \frac{\partial^2 C^{(1)}}{\partial x \partial y_j} \\ & - \sum_{j=1}^n \left(2\sum_{k=1}^n y_k \left\langle \xi_s^j, \xi_s^k \right\rangle\right) \frac{y_j^2}{2} \frac{\partial^2 C^{(1)}}{\partial y_j^2} \\ & + \sum_{j=1}^n \left(\sum_{k=1}^n y_k \left\langle \xi_s^j - \sigma_s^\omega, \xi_s^k \right\rangle\right) y_j \frac{\partial C^{(1)}}{\partial y_j} \end{aligned}$$

and

$$L_2 C^0 = \left(2\left\|\sum_{k=1}^n y_k \xi_s^k\right\|^2\right) \frac{x^2}{2} \frac{\partial^2 C^0}{\partial x^2}$$

where $C^0 = BS(s,x,V_s)$ given by the Black & Scholes (1973) formula as in (3.3) and $C^{(1)}(s,x,y)$ has been computed in (3.5) above.

Proof. As in Fournié et al. (1997) and Lebuchoux & Musiela (1999) we can differentiate twice with respect to ε the P.D.E. obtained in the previous lemma to get:

$$\begin{cases} 0 = L_0^\varepsilon C^{(2)} + 2L_1^\varepsilon C^{(1)} + L_2 C^0 \\ \\ 0 = C^{(2)} \text{ for } s = T \end{cases} \tag{8.2}$$

where L_0^ε and L_1^ε where computed in the previous lemma and with:

$$\begin{aligned} L_2 C^{0,\varepsilon} \quad = \quad & \left(2\left\|\sum_{k=1}^n y_k \xi_s^k\right\|^2\right) \frac{x^2}{2} \frac{\partial^2 C^{0,\varepsilon}}{\partial x^2} - \sum_{j=1}^n \left(2\left\|\sum_{k=1}^n y_k \xi_s^k\right\|^2\right) x y_j \frac{\partial^2 C^{0,\varepsilon}}{\partial x \partial y_j} \\ & + \sum_{j=1}^n \left(2\left\|\sum_{k=1}^n y_k \xi_s^k\right\|^2\right) \frac{y_j^2}{2} \frac{\partial^2 C^{0,\varepsilon}}{\partial y_j^2} - \sum_{j=1}^n \left(2\left\|\sum_{k=1}^n y_k \xi_s^k\right\|^2\right) y_j \frac{\partial C^{0,\varepsilon}}{\partial y_j} \end{aligned}$$

and again, as in Lebuchoux & Musiela (1999) and Karatzas & Shreve (1991), we take the limit as $\varepsilon \to \infty$ and compute $C^{2)}$ as the solution to the P.D.E.

$$\begin{cases} 0 = L_0^0 C^{(2)} + 2L_1^0 C^{(1)} + L_2 C^0 \\ 0 = C^{(2)} \text{ for } s = T \end{cases} \tag{8.3}$$

with $L_0^0 C^{(2)}$, $L_1^0 C^{(1)}$ and $L_2 C^0$ defined as above. ∎

We can now compute a closed-form solution to the equation verified by $C^{(2)}$ using its Feynman-Kac representation and the formula for $C^{(1)}(s,x,y)$ computed in (3.5) above.

Proposition 39 *The derivative* $C^{(2)}\left(t, F_t^\omega, (\widehat{\omega}_{j,t})_{j=1,\dots,n}\right)$ *can be computed as:*

$$\begin{aligned} C^{(2)} &= \int_t^T E\left[\left(2\sum_{j=1}^n \widehat{\omega}_{j,s}\left\langle \xi_s^j, \sigma_s^\omega \right\rangle\right)\frac{(F_s^\omega)^2}{2}\frac{\partial^2 C^{(1)}}{\partial x^2}\right] ds \\ &+ \int_t^T E\left[\sum_{j=1}^n\left(\sum_{k=1}^n \widehat{\omega}_{k,s}\left\langle \xi_s^j - \sigma_s^\omega, \xi_s^k\right\rangle\right) F_s^\omega \widehat{\omega}_{j,s}\frac{\partial^2 C^{(1)}}{\partial x \partial y_j}\right] ds \\ &+ \int_t^T E\left[\sum_{j=1}^n\left(\sum_{k=1}^n \widehat{\omega}_{k,s}\left\langle \xi_s^j - \sigma_s^\omega, \xi_s^k\right\rangle\right) \widehat{\omega}_{j,s}\frac{\partial C^{(1)}}{\partial y_j}\right] ds \\ &+ \int_t^T E\left[\left(2\left\|\sum_{k=1}^n \widehat{\omega}_{k,s}\xi_s^k\right\|^2\right)\frac{(F_s^\omega)^2}{2}\frac{\partial^2 C^0}{\partial x^2}\right] ds \end{aligned} \tag{8.4}$$

Proof. Because $C^{(2)}$ solves the P.D.E. (3.4) in the above lemma, we can write the Feynman-Kac representation of the solution as:

$$C^{(2)} = \int_t^T E\left[2L_1^\varepsilon C^{(1)}\left(s, F_s^\omega, (\widehat{\omega}_{j,s})_{j=1,\dots,n}\right) + L_2 C^0\left(s, F_s^\omega, (\widehat{\omega}_{j,s})_{j=1,\dots,n}\right)\right] ds$$

which together with $\partial^2 C^{(1)}/\partial y_j^2 = 0$ gives the desired result. ∎

We first introduce a set of convenient notations and then begin with a sequence of purely technical lemmas that will help us compute the final expression of $C^{(2)}\left(t, F_t^\omega, (\widehat{\omega}_{j,t})_{j=1,\dots,n}\right)$.

Notation 40 *We note*

$$h(s,u,x,K) = \frac{\ln\frac{x}{K} + \int_s^u \left\langle \xi_v^j, \sigma_v^\omega\right\rangle dv + \frac{1}{2}V_{t,T}}{\sqrt{V_T}}$$

and

$$\varphi_s^j(u) = \frac{\left\langle \xi_u^j, \sigma_u^\omega\right\rangle}{\sqrt{V_T}} \exp\left(2\int_s^u \left\langle \xi_v^j, \sigma_v^\omega\right\rangle dv\right)$$

to simplify the computations that follow.

We now compute a generic formula for a term that will appear frequently in the expression of $C^{(2)}$.

Lemma 41 *Suppose* (z_1, z_2, z_3) *is a centered Gaussian vector with covariance matrix* V. *We have:*

$$E\left[\exp(z_1)\, n\left(\frac{z_2+a}{b}\right) z_3\right]$$
$$=\exp\left(\frac{1}{2}V_{1,1}\right)\left(\frac{V_{3,1}}{\sqrt{(V_{2,2}/b^2+1)}}+\frac{bV_{2,3}}{V_{2,2}}\frac{(V_{2,1}+a)/b}{(V_{2,2}/b^2+1)^{\frac{3}{2}}}\right) n\left(\frac{(V_{2,1}+a)/b}{\sqrt{(V_{2,2}/b^2+1)}}\right)$$

where we have noted as above $n(x) = 1/\sqrt{2\pi}\exp\left(-\frac{1}{2}x^2\right)$.

Proof. First we can use the Cameron-Martin formula to get:

$$\begin{aligned} & E\left[\exp(z_1)\, n\left(\frac{z_2+a}{b}\right) z_3\right] \\ = \; & \exp\left(\frac{1}{2}V_{1,1}\right) E\left[n\left(\frac{z_2+V_{2,1}+a}{b}\right)(z_3+V_{3,1})\right] \end{aligned}$$

and because if g is Gaussian with mean m and variance v^2 we have:

$$E[n(g)] = \frac{1}{\sqrt{(v^2+1)}} n\left(\frac{m}{\sqrt{(v^2+1)}}\right)$$

and

$$E[gn(g)] = \frac{m}{(v^2+1)^{\frac{3}{2}}} n\left(\frac{m}{\sqrt{(v^2+1)}}\right)$$

hence

$$\begin{aligned} & E\left[n\left(\frac{z_2+V_{2,1}+a}{b}\right) V_{3,1}\right] \\ = \; & \frac{V_{3,1}}{\sqrt{(V_{2,2}/b^2+1)}} n\left(\frac{(V_{2,1}+a)/b}{\sqrt{(V_{2,2}/b^2+1)}}\right) \end{aligned}$$

and

$$\begin{aligned} & E\left[n\left(\frac{z_2+V_{2,1}+a}{b}\right) z_3\right] \\ = \; & E\left[n\left(\frac{z_2+V_{2,1}+a}{b}\right) \frac{z_2}{b}\frac{bV_{2,3}}{V_{2,2}}\right] \\ = \; & \frac{bV_{2,3}}{V_{2,2}}\frac{(V_{2,1}+a)/b}{(V_{2,2}/b^2+1)^{\frac{3}{2}}} n\left(\frac{(V_{2,1}+a)/b}{\sqrt{(V_{2,2}/b^2+1)}}\right) \end{aligned}$$

which is the desired result. ■

For simplicity, we introduce the following notation based on the previous lemma.

Notation 42 *We note*

$$\Phi_u(\sigma_{1,v},\sigma_{2,v},\sigma_{3,v})=\left(\frac{V_{3,1}}{\sqrt{(V_{2,2}/b^2+1)}}+\frac{bV_{2,3}}{V_{2,2}}\frac{(V_{2,1}+a)/b}{(V_{2,2}/b^2+1)^{\frac{3}{2}}}\right)n\left(\frac{(V_{2,1}+a)/b}{\sqrt{(V_{2,2}/b^2+1)}}\right) \tag{8.5}$$

where $\sigma_{1,u}\in\mathbb{R}^d$ *for* $u\in[t,T]$ *and* $V_{i,j}=\int_t^u\langle\sigma_{1,v},\sigma_{2,v}\rangle\,dv$.

We then compute the various derivatives of $C^{(1)}$, beginning with $\partial C^{(1)}/\,\partial x$ and $\partial^2C^{(1)}/\partial x^2$.

Lemma 43 *If* $C^{(1)}$ *is given by (3.5) above then*

$$\begin{aligned} &E\left[\left(\sum_{i=1}^n\widehat{\omega}_{i,s}\left\langle\xi_s^j,\sigma_s^\omega\right\rangle\right)(F_s^\omega)^2\frac{\partial^2C^{(1)}}{\partial x^2}\right] \qquad (8.6)\\ &= \sum_{i=1}^n\left\langle\xi_s^i,\sigma_s^\omega\right\rangle\int_s^T\left(\sum_{j=1}^n\varphi_s^j(u)\widehat{\omega}_{j,t}\widehat{\omega}_{i,t}\,(F_t^\omega)^2\,\phi(u)\left(H_u^1+H_u^2\right)\right)du\end{aligned}$$

where we have noted:

$$\begin{aligned} H_u^1 &= \frac{\exp\left(\int_t^u\left\|\eta_v^{i,j}\right\|^2dv\right)\exp\left(\frac{V_{t,u}\bar{\psi}(u,K)^2}{V_{t,u}+V_T}\right)\left(\frac{V_{t,u}/V_T}{\sqrt{V_{t,u}/V_T+V_T}}+\frac{\bar{\psi}(u,K)^2}{V_{t,u}/V_T+1}\right)\frac{(V_{t,u}/V_T+1)^{\frac{1}{4}}}{\sqrt{V_{t,u}/V_T}}}{F_t^\omega V_T\sqrt{2\pi V_{t,u}/V_T}}\\ H_u^2 &= \frac{\Phi_u(\eta_v^{i,j},\frac{\sigma_u^\omega}{\sqrt{V_T}},\frac{\sigma_u^\omega}{\sqrt{V_T}})+\psi(u,K)\Phi_u(\eta_v^{i,j},\frac{\sigma_u^\omega}{\sqrt{V_T}},0)}{F_t^\omega\sqrt{V_T}}\end{aligned}$$

and

$$\begin{aligned} \eta_v^{i,j} &= \left(\xi_v^j-\sigma_v^\omega+1_{\{v\le s\}}\left(\xi_v^i+2\sigma_v^\omega\right)\right)\\ \psi(u,K) &= \frac{\ln\frac{F_t^\omega}{K}+\int_t^u\left\langle\xi_v^j,\sigma_v^\omega\right\rangle dv+\frac{1}{2}V_T}{\sqrt{V_T}}\\ \phi(u) &= \exp\left(\int_t^u\left(\left\langle\xi_v^j,\sigma_v^\omega\right\rangle-\frac{1}{2}\left\|\xi_v^j\right\|^2+\frac{1}{2}V_{t,u}\right)dv\right)\\ &\quad\exp\left(\int_t^s\left(\left\langle\xi_v^i,\sigma_v^\omega\right\rangle-\frac{1}{2}\left\|\xi_v^i\right\|^2-V_{t,s}\right)dv\right)\end{aligned}$$

and

$$\begin{aligned} \bar{\psi}(u,K) &= \psi(u,K)+\int_t^u\left\|\left(\xi_v^j-\sigma_v^\omega+1_{\{v\le s\}}\left(\xi_v^j+2\sigma_v^\omega\right)\right)\right\|^2dv\\ &\quad+\frac{\int_t^u\left(\left\langle\xi_v^i,\sigma_v^\omega\right\rangle+1_{\{v\le s\}}\left\langle\xi_v^i+2\sigma_v^\omega,\sigma_v^\omega\right\rangle\right)dv-V_{t,u}}{\sqrt{V_T}}\end{aligned}$$

to simplify the expression above.

Proof. With $C^{(1)}$ written as:

$$C^{(1)}(s,x,y)=x\int_s^T\sum_{j=1}^n y_j\varphi_s^j(u)n\left(h_{s,u}(x,K)\right)du$$

we get:

$$\frac{\partial C^{(1)}}{\partial x} = \int_s^T \sum_{j=1}^n y_j \varphi_s^j(u) n\left(h_{s,u}(x,K)\right) du + \int_s^T \sum_{j=1}^n y_j \varphi_s^j(u) \frac{n\left(h_{s,u}(x,K)\right) h_{s,u}(x,K)}{\sqrt{V_T}} du$$

hence

$$\frac{\partial^2 C^{(1)}}{\partial x^2} = \int_s^T \sum_{j=1}^n y_j \varphi_s^j(u) \frac{n\left(h_{s,u}(x,K)\right) h_{s,u}(x,K)}{x\sqrt{V_T}} du + \int_s^T \sum_{j=1}^n y_j \varphi_s^j(u) \frac{n\left(h_{s,u}(x,K)\right)}{xV_T} \left(h_{s,u}^2(x,K)+1\right) du$$

we can then compute

$$E\left[\widehat{\omega}_{i,s}\left(F_s^\omega\right)^2 \frac{\partial^2 C^{(1)}(s)}{\partial x^2}\right]$$
$$= \int_s^T \sum_{j=1}^n E\left[\widehat{\omega}_{j,u}\widehat{\omega}_{i,s}\left(F_s^\omega\right)^2 \varphi_s^j(u) \frac{n\left(h_{s,u}(F_u^\omega,K)\right)}{F_u^\omega\sqrt{V_T}} \left(\frac{h_{s,u}^2(F_u^\omega,K)}{\sqrt{V_T}}\right)\right] du$$
$$+ \int_s^T \sum_{j=1}^n E\left[\widehat{\omega}_{j,u}\widehat{\omega}_{i,s}\left(F_s^\omega\right)^2 \varphi_s^j(u) \frac{n\left(h_{s,u}(F_u^\omega,K)\right)}{F_u^\omega\sqrt{V_T}} \left(1+\frac{1}{\sqrt{V_T}}\right)\right] du$$

which can be written:

$$\int_s^T \sum_{j=1}^n \varphi_s^j(u)\widehat{\omega}_{j,t}\widehat{\omega}_{i,t}\left(F_t^\omega\right)^2 \exp\left(\int_t^u \left(\left\langle \xi_v^j, \sigma_v^\omega\right\rangle - \frac{1}{2}\left\|\xi_v^i\right\|^2 + \frac{1}{2}V_{t,u}\right) dv\right)$$
$$\exp\left(\int_t^s \left(\left\langle \xi_v^i, \sigma_v^\omega\right\rangle - \frac{1}{2}\left\|\xi_v^i\right\|^2 - V_{t,s}\right) dv\right)$$
$$E[\exp\left(\int_t^u \left(\xi_v^j - \sigma_v^\omega + 1_{\{v\le s\}}\left(\xi_v^i + 2\sigma_v^\omega\right)\right) dW_v\right) \frac{n\left(h_{s,u}(F_u^\omega,K)\right)}{F_t^\omega\sqrt{V_T}}$$
$$\left(\frac{h_{s,u}^2(F_u^\omega,K)}{\sqrt{V_T}} + \left(1+\frac{1}{\sqrt{V_T}}\right)\right)]du$$

An because if we remember that

$$h(s,u,x,K) = \frac{\ln\frac{x}{K} + \int_s^u \left\langle \xi_v^j, \sigma_v^\omega\right\rangle dv + \frac{1}{2}V_T}{\sqrt{V_T}}$$

and if we note:

$$\begin{aligned} \eta_v^{i,j} &= \left(\xi_v^j - \sigma_v^\omega + 1_{\{v \leq s\}} \left(\xi_v^i + 2\sigma_v^\omega\right)\right) \\ Z_{1,u} &= \int_t^u \eta_v^{i,j} dW_v \text{ and } Z_2 = \frac{\int_t^s \sigma_u^\omega dW_u}{\sqrt{V_T}} \\ \psi(u) &= \frac{\ln \frac{F_t^\omega}{K} + \int_t^u \left\langle \xi_v^j, \sigma_v^\omega \right\rangle dv + \frac{1}{2} V_T}{\sqrt{V_T}} \\ \phi(u) &= \exp\left(\int_t^u \left(\left\langle \xi_v^j, \sigma_v^\omega \right\rangle - \frac{1}{2} \left\|\xi_v^j\right\|^2 + \frac{1}{2} V_{t,u}\right) dv\right) \\ &\quad \exp\left(\int_t^s \left(\left\langle \xi_v^i, \sigma_v^\omega \right\rangle - \frac{1}{2} \left\|\xi_v^i\right\|^2 - V_{t,s}\right) dv\right) \end{aligned}$$

we can get:

$$\begin{aligned} & E\left[\exp\left(Z_{1,u}\right) \frac{n\left(Z_{2,u} + \psi(u)\right)\left(Z_{2,u} + \psi(u)\right)^2}{F_t^\omega \sqrt{V_T}}\right] du \\ &= E\left[\exp\left(\int_t^u \eta_v^{i,j} dW_v\right) \frac{n\left(Z_{2,u} + \psi(u)\right)\left(Z_{2,u} + \psi(u)\right)^2}{F_t^\omega \sqrt{V_T}}\right] du \\ &= \exp\left(\int_t^u \left\|\eta_v^{i,j}\right\|^2 dv\right) E\left[\frac{n\left(Z_{2,u} + \bar{\psi}(u)\right)\left(Z_{2,u} + \bar{\psi}(u)\right)^2}{F_t^\omega \sqrt{V_T}}\right] du \end{aligned}$$

where

$$\begin{aligned} \bar{\psi}(u) &= \psi(u) + \int_t^u \left\|\left(\xi_v^j - \sigma_v^\omega + 1_{\{v \leq s\}} \left(\xi_v^i + 2\sigma_v^\omega\right)\right)\right\|^2 dv \\ &\quad + \frac{\int_t^u \left(\left\langle \xi_v^j, \sigma_v^\omega \right\rangle + 1_{\{v \leq s\}} \left\langle \xi_v^i + 2\sigma_v^\omega, \sigma_v^\omega \right\rangle\right) dv - V_{t,u}}{\sqrt{V_T}} \end{aligned}$$

hence we have

$$\begin{aligned} & E\left[\exp\left(Z_{1,u}\right) \frac{n\left(Z_{2,u} + \psi(u)\right)\left(Z_{2,u} + \psi(u)\right)^2}{F_t^\omega V_T}\right] \\ = \; & \frac{\exp\left(\int_t^u \left\|\eta_v^{i,j}\right\|^2 dv\right) \exp\left(\frac{V_{t,u} \bar{\psi}(u)^2}{V_{t,u} + V_T}\right)}{F_t^\omega V_T \sqrt{2\pi V_{t,u}/V_T}} \\ & \left(\frac{V_{t,u}/V_T}{\sqrt{V_{t,u}/V_T + V_T}} + \frac{\bar{\psi}(u)^2}{V_{t,u}/V_T + 1}\right) \frac{\left(V_{t,u}/V_T + 1\right)^{\frac{1}{4}}}{\sqrt{V_{t,u}/V_T}} \end{aligned}$$

We can then obtain the last term by computing

$$E\left[\exp\left(Z_{1,u}\right) \frac{n\left(h_{s,u}(F_u^\omega, K)\right)\left(h_{s,u}(F_u^\omega, K)\right)}{F_t^\omega \sqrt{V_T}}\right] = \frac{\Phi_u(\eta_v^{i,j}, \frac{\sigma_u^\omega}{\sqrt{V_T}}, \frac{\sigma_u^\omega}{\sqrt{V_T}}) + \psi(u) \Phi_u(\eta_v^{i,j}, \frac{\sigma_u^\omega}{\sqrt{V_T}}, 0)}{F_t^\omega \sqrt{V_T}}$$

which produces the desired result. ■

We then compute the term in $\partial C^{(1)} / \partial y_i$ in the expression of $C^{(2)}$.

Lemma 44 *If $C^{(1)}$ is given by (3.5) above then*

$$\begin{aligned} E\left[\widehat{\omega}_{j,s}\widehat{\omega}_{k,s}\frac{\partial C^{(1)}}{\partial y_j}\right] &= \int_s^T \widehat{\omega}_{j,t}\widehat{\omega}_{k,t}F_t^{\omega}\varphi_s^j(u)\exp\left(\int_t^s \left\langle \xi_v^j+\xi_v^k,\sigma_v^{\omega}\right\rangle dv\right) \\ &\exp\left(-\frac{1}{2}V_{t,u}\right)\exp\left(-\frac{1}{2}\int_t^s\left(\left\|\xi_v^j\right\|^2+\left\|\xi_v^k\right\|^2\right)dv\right) \\ &\Phi_u(\sigma_v^{\omega}+1_{\{v\leq s\}}\left(\xi_v^k+\xi_v^j\right),\frac{\sigma_u^{\omega}}{\sqrt{V_T}},0)du \end{aligned}$$

Proof. From the expression of $C^{(1)}$ we get:

$$\frac{\partial C^{(1)}}{\partial y_j} = x\int_s^T \varphi_s^j(u)n\left(h_{s,u}(x,K)\right)du$$

hence

$$\begin{aligned} E\left[\widehat{\omega}_{j,s}\widehat{\omega}_{k,s}\frac{\partial C^{(1)}}{\partial y_j}\right] &= \int_s^T E\left[\widehat{\omega}_{j,s}\widehat{\omega}_{k,s}F_u^{\omega}\varphi_s^j(u)n\left(h_{s,u}(x,K)\right)\right]du \\ &= \int_s^T \widehat{\omega}_{j,t}\widehat{\omega}_{k,t}F_t^{\omega}\varphi_s^j(u)\exp\left(\int_t^s \left\langle \xi_v^j+\xi_v^k,\sigma_v^{\omega}\right\rangle dv\right) \\ &\exp\left(-\frac{1}{2}V_{t,u}\right)\exp\left(-\frac{1}{2}\int_t^s\left(\left\|\xi_v^j\right\|^2+\left\|\xi_v^k\right\|^2\right)dv\right) \\ &E\left[\exp\left(\int_t^u \sigma_v^{\omega}+1_{\{v\leq s\}}\left(\xi_v^k+\xi_v^j\right)dW_v\right)n\left(h_{s,u}(x,K)\right)\right]du \end{aligned}$$

with

$$\begin{aligned} &E\left[\exp\left(\int_t^u \sigma_v^{\omega}+1_{\{v\leq s\}}\left(\xi_v^k+\xi_v^j\right)dW_v\right)n\left(Z_{2,u}+\psi(u)\right)\right] \\ &= \Phi_u(\sigma_v^{\omega}+1_{\{v\leq s\}}\left(\xi_v^k+\xi_v^j\right),\frac{\sigma_u^{\omega}}{\sqrt{V_T}},0) \end{aligned}$$

hence the desired result. ■

We can now compute the cross term in $\partial^2 C^{(1)}/\,\partial x\partial y_i$ in the expression of $C^{(2)}$.

Lemma 45 *If $C^{(1)}$ is given by (3.5) above then*

$$\begin{aligned} E\left[\widehat{\omega}_{k,s}F_s^{\omega}\widehat{\omega}_{j,s}\frac{\partial^2 C^{(1)}}{\partial x\partial y_j}\right] &= \int_s^T \widehat{\omega}_{j,t}\widehat{\omega}_{k,t}F_t^{\omega}\varphi_s^j(u)\exp\left(\int_t^s \left\langle \xi_v^j+\xi_v^k,\sigma_v^{\omega}\right\rangle dv\right) \\ &\exp\left(-\frac{1}{2}V_{t,u}\right)\exp\left(-\frac{1}{2}\int_t^s\left(\left\|\xi_v^j\right\|^2+\left\|\xi_v^k\right\|^2\right)dv\right) \\ &\Phi_u(1_{\{v\leq s\}}\left(\sigma_v^{\omega}+\xi_v^k+\xi_v^j\right),\frac{\sigma_u^{\omega}}{\sqrt{V_T}},0)du \end{aligned}$$

Proof. With $C^{(1)}$ written as:

$$C^{(1)}(s,x,y) = x\int_s^T \sum_{j=1}^n y_j\varphi_s^j(u)n\left(h_{s,u}(x,K)\right)du$$

we get:

$$\frac{\partial^2 C^{(1)}}{\partial x \partial y_j} = \int_s^T \varphi_s^j(u) n\left(h_{s,u}(x,K)\right) du$$

hence

$$\begin{aligned}
E\left[\widehat{\omega}_{j,s}\widehat{\omega}_{k,s}F_s^\omega \frac{\partial C^{(1)}}{\partial y_j}\right] &= \int_s^T E\left[\widehat{\omega}_{j,s}\widehat{\omega}_{k,s}F_s^\omega \varphi_s^j(u) n\left(h_{s,u}(x,K)\right)\right] du \\
&= \int_s^T \widehat{\omega}_{j,t}\widehat{\omega}_{k,t}F_t^\omega \varphi_s^j(u) \exp\left(\int_t^s \left\langle \xi_v^j + \xi_v^k, \sigma_v^\omega \right\rangle dv\right) \\
&\quad \exp\left(-\frac{1}{2}V_{t,u}\right) \exp\left(-\frac{1}{2}\int_t^s \left(\left\|\xi_v^j\right\|^2 + \left\|\xi_v^k\right\|^2\right) dv\right) \\
&\quad E\left[\exp\left(\int_t^s \sigma_v^\omega + \left(\xi_v^k + \xi_v^j\right) dW_v\right) n\left(h_{s,u}(x,K)\right)\right] du
\end{aligned}$$

with

$$\begin{aligned}
&E\left[\exp\left(\int_t^s \left(\sigma_v^\omega + \xi_v^k + \xi_v^j\right) dW_v\right) n\left(Z_{2,u} + \psi(u)\right)\right] \\
&= \Phi_u(1_{\{v \le s\}}\left(\sigma_v^\omega + \xi_v^k + \xi_v^j\right), \frac{\sigma_u^\omega}{\sqrt{V_T}}, 0)
\end{aligned}$$

hence the desired result. ■

And we now get the last term in the development of $C^{(2)}$:

Lemma 46 *If $C^{(1)}$ is given by (3.5) above then*

$$\begin{aligned}
& E\left[\widehat{\omega}_{i,s}\widehat{\omega}_{j,s}\left(F_s^\omega\right)^2 \frac{\partial^2 C^0}{\partial x^2}\right] \\
= \; & \int_s^T \widehat{\omega}_{i,t}\widehat{\omega}_{j,t}F_t^\omega \exp\left(\int_t^s \left\langle \xi_v^i + \xi_v^j, \sigma_v^\omega \right\rangle dv\right) \\
& \exp\left(-\frac{1}{2}V_{t,u}\right) \exp\left(-\frac{1}{2}\int_t^s \left(\left\|\xi_v^j\right\|^2 + \left\|\xi_v^i\right\|^2\right) dv\right) \\
& \frac{1}{\sqrt{V_{s,T}}}\Phi_s(\sigma_v^\omega + \left(\xi_v^i + \xi_v^j\right), \frac{\sigma_v^\omega}{\sqrt{V_T}}, 0)
\end{aligned}$$

Proof. As before we know that

$$\frac{\partial^2 C^0}{\partial x^2} = \frac{n(h(x, V_{s,T}))}{x\sqrt{V_{s,T}}}$$

which means that we need to compute:

$$\begin{aligned}
& E\left[\widehat{\omega}_{i,s}\widehat{\omega}_{j,s}F_s^\omega \frac{n(h(F_s^\omega, V_{s,T}))}{\sqrt{V_{s,T}}}\right] \\
= \; & \widehat{\omega}_{i,t}\widehat{\omega}_{j,t}F_t^\omega \exp\left(\int_t^s \left\langle \xi_v^i + \xi_v^j, \sigma_v^\omega \right\rangle dv\right) \\
& \exp\left(-\frac{1}{2}V_{t,u}\right) \exp\left(-\frac{1}{2}\int_t^s \left(\left\|\xi_v^j\right\|^2 + \left\|\xi_v^i\right\|^2\right) dv\right) \\
& E\left[\exp\left(\int_t^s \sigma_v^\omega + \left(\xi_v^i + \xi_v^j\right) dW_v\right) \frac{n(h(F_s^\omega, V_{s,T}))}{\sqrt{V_{s,T}}}\right]
\end{aligned}$$

with

$$
\begin{aligned}
& E\left[\exp\left(\int_t^s \sigma_v^\omega + \left(\xi_v^i + \xi_v^j\right) dW_v\right) \frac{n(Z_{2,s} + \psi(s))}{\sqrt{V_{s,T}}}\right] \\
= \quad & \frac{1}{\sqrt{V_{s,T}}} \Phi_s(\sigma_v^\omega + \left(\xi_v^i + \xi_v^j\right), \frac{\sigma_v^\omega}{\sqrt{V_T}}, 0)
\end{aligned}
$$

hence the desired result. ■

Finally we can assemble these results to compute the second-order term in the price development.

Proposition 47 *The derivative* $C^{(2)}\left(t, F_t^\omega, (\widehat{\omega}_{j,t})_{j=1,\dots,n}\right)$ *is given by:*

$$
C^{(2)} = \int_t^T (A_s + B_s + C_s + D_s)\, ds \tag{8.7}
$$

with

$$
A_s = \sum_{i=1}^n \left\langle \xi_s^i, \sigma_s^\omega \right\rangle \int_s^T \left(\sum_{j=1}^n \varphi_s^j(u) \widehat{\omega}_{j,t} \widehat{\omega}_{i,t} \left(F_t^\omega\right)^2 \phi(u) \left(H_u^1 + H_u^2\right)\right) du
$$

using the notations in (8.6). With

$$
\begin{aligned}
B_s \quad = \quad & \int_s^T \sum_{j=1}^n \sum_{k=1}^n \left\langle \xi_s^j - \sigma_s^\omega, \xi_s^k \right\rangle \widehat{\omega}_{j,t} \widehat{\omega}_{k,t} F_t^\omega \varphi_s^j(u) \exp\left(\int_t^s \left\langle \xi_v^j + \xi_v^k, \sigma_v^\omega \right\rangle dv\right) \\
& \exp\left(-\frac{1}{2} V_{t,u}\right) \exp\left(-\frac{1}{2} \int_t^s \left(\left\|\xi_v^j\right\|^2 + \left\|\xi_v^k\right\|^2\right) dv\right) \\
& \Phi_u(\sigma_v^\omega + 1_{\{v \le s\}} \left(\xi_v^k + \xi_v^j\right), \frac{\sigma_u^\omega}{\sqrt{V_T}}, 0) du
\end{aligned}
$$

and

$$
\begin{aligned}
C_s \quad = \quad & \int_s^T \sum_{j=1}^n \sum_{k=1}^n \left\langle \xi_s^j - \sigma_s^\omega, \xi_s^k \right\rangle \widehat{\omega}_{j,t} \widehat{\omega}_{k,t} F_t^\omega \varphi_s^j(u) \exp\left(\int_t^s \left\langle \xi_v^j + \xi_v^k, \sigma_v^\omega \right\rangle dv\right) \\
& \exp\left(-\frac{1}{2} V_{t,u}\right) \exp\left(-\frac{1}{2} \int_t^s \left(\left\|\xi_v^j\right\|^2 + \left\|\xi_v^k\right\|^2\right) dv\right) \\
& \Phi_u(1_{\{v \le s\}} \left(\sigma_v^\omega + \xi_v^k + \xi_v^j\right), \frac{\sigma_u^\omega}{\sqrt{V_T}}, 0) du
\end{aligned}
$$

and finally

$$
\begin{aligned}
D_s \quad = \quad & \int_s^T \sum_{k=1}^n \sum_{j=1}^n \left\langle \xi_s^k, \xi_s^j \right\rangle \widehat{\omega}_{i,t} \widehat{\omega}_{j,t} F_t^\omega \exp\left(\int_t^s \left\langle \xi_v^i + \xi_v^j, \sigma_v^\omega \right\rangle dv\right) \\
& \exp\left(-\frac{1}{2} V_{t,u}\right) \exp\left(-\frac{1}{2} \int_t^s \left(\left\|\xi_v^j\right\|^2 + \left\|\xi_v^i\right\|^2\right) dv\right) \\
& \frac{1}{\sqrt{V_{s,T}}} \Phi_s(\sigma_v^\omega + \left(\xi_v^i + \xi_v^j\right), \frac{\sigma_v^\omega}{\sqrt{V_T}}, 0) du
\end{aligned}
$$

Bibliography

Alizadeh, F. (1995), 'Interior point methods in semidefinite programming with applications to combinatorial optimization', *SIAM Journal on Optimization* **5**, 13–51.

Alizadeh, F., Haeberly, J. A. & Overton, M. L. (1998), 'Primal–dual interior–point methods for semidefinite programming: Convergence rates, stability and numerical results', *SIAM Journal on Optimization* **8**, 746–768.

Avellaneda, M., Levy, A. & Paras, A. (1995), 'Pricing and hedging derivative securities in markets with uncertain volatilities', *Applied Mathematical Finance* **2**, 73–88.

Avellaneda, M. & Paras, A. (1996), 'Managing the volatility risk of portfolios of derivative securities: the lagrangian uncertain volatility model', *Applied Mathematical Finance* **3**, 21–52.

Azencott, R. (1984), 'Densite des diffusions en temps petit: Developements asymptotiques.', *Lecture Notes in Mathematics* **1059**, 402–498.

Bally, V. & Pages, G. (2000), 'A quantization algorithm for solving multi-dimensional optimal stopping problems', *Preprint 628, Laboratoire de Probabilites, University de Paris VI.* .

Ben-Tal, A., El Ghaoui, L. & Lebret, H. (1998), Robust semidefinite programming, *in* S. Sastry, H. Wolkowitcz & L. Vandenberghe, eds, 'Handbook on Semidefinite Programming', Vol. 27 of *International Series in Operations Research and Management Science*.

Bertsimas, D. & Popescu, I. (2002), 'On the relation between option and stock prices: a convex optimization approach', *Operations Research* **50**(2).

Black, F. (1976), 'The pricing of commodity contracts.', *Journal of Financial Economics* **3**, 167–179.

Black, F. & Scholes, M. (1973), 'The pricing of options and corporate liabilities', *Journal of Political Economy* **81**, 637–659.

Boyd, S., El Ghaoui, L., Feron, E. & Balakrishnan, V. (1994), *Linear Matrix Inequalities in System and Control Theory*, SIAM.

Brace, A., Dun, T. & Barton, G. (1999), 'Towards a central interest rate model', *Working Paper. FMMA* .

Brace, A., Gatarek, D. & Musiela, M. (1997), 'The market model of interest rate dynamics', *Mathematical Finance* **7**(2), 127–155.

Brace, A. & Womersley, R. S. (2000), 'Exact fit to the swaption volatility matrix using semidefinite programming', *Working paper, ICBI Global Derivatives Conference* .

Breeden, D. T. & Litzenberger, R. H. (1978), 'Price of state-contingent claims implicit in option prices', *Journal of Business* **51**(4), 621–651.

Brixius, N., Potra, F. & Sheng, R. (1996), 'Solving semidefinite programs in mathematica', *Reports on Computational Mathematics. The University of Iowa* **97**.

Cont, R. (2001), 'Inverse problems in financial modeling: theoretical and numerical aspects of model calibration.', *Lecture Notes, Princeton University.* .

Conze, A. & Viswanathan (1991), 'Path dependent options: The case of lookback options.', *Journal of Finance* **46**(5), 1893–1907.

d'Aspremont, A. & El Ghaoui, L. (2002), 'Static arbitrage bounds on basket option prices', *EECS Technical report, UC Berkeley.* .

David, J. (1994), 'Algorithms for analysis and design of robust controllers.', *PhD thesis, Kat. Univ. Leuven, ESAT, 3001 Leuven, Belgium.* .

Davis, M. (1994), 'A general option pricing formula', *Working paper, Imperial College London* .

Douady, R. (1995), 'Optimisation du gamma d'une option sur panier ou sur spread en l'absence d'options croisées.', *Working paper* .

Duffie, D. (1996), *Dynamic asset pricing theory*, 2nd edn, Princeton University Press, Princeton, N.J.

Duffie, D. & Kan, R. (1996), 'A yield factor model of interest rates', *Mathematical Finance* **6**(4).

Dupire, B. (1994), 'Pricing with a smile', *Risk* **7**(1).

Edirisinghe, C., Naik, V. & Uppal, R. (1993), 'Optimal replication of options with transactions costs and trading restrictions', *Journal of Financial and Quantitative Analysis* **28**(1), 117–138.

El Ghaoui, L. (1997), 'Robust control', *Course given at the Ecole Polytechnique in 1997-98 (preliminary version, in french)* .

El Ghaoui, L. (2001), 'Non-parametric bounds on the price of a basket option', *Mimeo* .

El Karoui, N., Geman, H. & Rochet, J. C. (1995), 'Changes of numeraire, changes of probability measures and option pricing', *Journal of Applied Probability* **32**, 443–458.

El Karoui, N., Jeanblanc-Picqué, M. & Shreve, S. E. (1998), 'On the robustness of the black-scholes equation', *Mathematical Finance* **8**, 93–126.

El Karoui, N. & Lacoste, V. (1992), 'Multifactor analysis of the term structure of interest rates', *Proceedings, AFFI* .

El Karoui, N. & Quenez, M. (1991), 'Programmation dynamique et évaluation des actifs contingents en marchés incomplets.', *Comptes Rendus de l'Académie des Sciences de Paris, Série I* **313**, 851–854.

El Karoui, N. & Quenez, M. (1995), 'Dynamic programming and pricing of contingent claims in an incomplete market', *Siam Journal of Control and Optimization* **33**, 29–66.

Fazel, M., Hindi, H. & Boyd, S. (2000), 'A rank minimization heuristic with application to minimum order system approximation', *American Control Conference, September 2000* .

Fouque, J.-P., Papanicolaou, G. & Sircar, K. R. (2000), *Derivatives in Financial Markets with Stochastic Volatility*, Cambridge University Press.

Fournié, E., Lebuchoux, J. & Touzi, N. (1997), 'Small noise expansion and importance sampling', *Asymptotic Analysis* **14**, 361–376.

Glasserman, P. & Zhao, X. (2000), 'Arbitrage-free discretization of lognormal forward libor and swap rate models.', *Finance and Stochastics* (4), 35–68.

Gozzi, F. & Vargiolu, T. (2000), 'On the superreplication approach for european interest rates derivatives', *Proceedings of the Ascona '99 Seminar on Stochastic Analysis, Random Fields and Applications* .

Hamy, O. (1999), 'Statistics of the weights.', *F.I.R.S.T. working paper, Paribas Capital Markets, London* .

Harrison, J. & Kreps, D. (1979), 'Martingales and arbitrage in multiperiod securities markets', *J. Econ. Theory* **20**, 381–408.

Harrison, J. & Pliska, S. R. (1981), 'Martingales and stochastic integrals in the theory of continous trading', *Stochastic Processes and their Applications* **11**, 215–260.

Heath, D., Jarrow, R. & Morton, A. (1992), 'Bond pricing and the term structure of interest rates: A new methodology', *Econometrica* **61**(1), 77–105.

Helmberg, C., Rendl, F., Vanderbei, R. J. & Wolkowicz, H. (1996), 'An interior–point method for semidefinite programming', *SIAM Journal on Optimization* **6**, 342–361.

Hettich, R. & Kortanek, K. O. (1993), 'Semi-infinite programming: Theory, methods and applications', *SIAM Review* **35**(3), 380–429.

Hodges, S. & Neuberger, A. (1989), 'Optimal replication of contingent claims under transaction costs', *Review of Futures Markets* **8**, 222–239.

Huynh, C. B. (1994), 'Back to baskets', *Risk* **7**(5).

Jackwerth, J. & Rubinstein, M. (1996), 'Recovering probability distributions from option prices', *Journal of Finance* **51**(5), 1611–1631.

Jamshidian, F. (1997), 'Libor and swap market models and measures', *Finance and Stochastics* **1**(4), 293–330.

Janson, S. (1997), *Gaussian Hilbert spaces*, Cambridge tracts in mathematics ; 129., Cambridge University Press, Cambridge, U.K. ; New York. Svante Janson. Includes bibliographical references (p. 329-334) and indexes.

Jarrow, R. (1986), 'The relationship between arbitrage and first order stochastic dominance.', *Journal of Finance* **41**(4), 915–921.

Ji, J., Potra, F. & Sheng, R. (1997), On the local convergence of a predictor–corrector method for semidefinite programming Reports on Computational Mathematics, Technical report, Department of Mathematics, The University of Iowa.

Ju, N. (2002), 'Pricing asian and basket options via taylor expansion.', *Journal of Computational Finance* **5**(3), 79–103.

Karatzas, I. & Shreve, S. E. (1991), *Brownian motion and stochastic calculus*, Vol. 113 of *Graduate texts in mathematics*, Springer-Verlag, New York.

Karatzas, I. & Shreve, S. E. (1998), *Methods of mathematical finance*, Applications of mathematics ; 39., Springer, New York.

Kawata, T. (1972), *Fourier analysis in probability theory*, Academic Press, New York,. 24 cm.

Kojima, M., Shindoh, S. & Hara, S. (1997), 'Interior-point methods for the monotone semidefinite linear complementarity problem in symmetrix matrixes', *SIAM Journal on Optimization* **7**, 86–125.

Øksendal, B. K. (1998), *Stochastic differential equations : an introduction with applications*, 5th edn, Springer, Berlin ; New York. Bernt Øksendal. Universitext Includes bibliographical references (p. [311]-316) and index.

Lacoste, V. (1996), 'Wiener chaos: a new approach to option hedging.', *Mathematical Finance* **6**(2), 197–213.

Laurent, J. & Leisen, D. (2000), Building a consistent pricing model from observed option prices, *in* M. Avellaneda, ed., 'Quantitative Analysis in Financial Markets', World Scientific Publishing.

Lebuchoux, J. & Musiela, M. (1999), 'Market models and smile effects in caps and swaptions volatilities.', *Working paper, Paribas Capital Markets.* .

Lemaréchal, C. & Oustry, F. (1999), 'Semidefinite relaxations and lagrangian duality with application to combinatorial optimization.', *INRIA, Rapport de recherche* **3710**.

Lobo, M., Vandenberghe, L., Boyd, S. & Lebret, H. (1998), 'Applications of second-order cone programming', *Linear Algebra and its Applications* **284**, 193–228. Special Issue on Linear Algebra in Control, Signals and Image Processing.

Longstaff, F. A., Santa-Clara, P. & Schwartz, E. S. (2000), 'The relative valuation of caps and swaptions: Theory and empirical evidence.', *Working Paper, The Anderson School at UCLA.* .

Longstaff, F. A. & Schwartz, E. S. (1998), 'Valuing american options by simulation: A simple least-squares approach', *Finance Working Paper, UCLA.* .

Luenberger, D. G. (1968), *Optimization by vector space methods*, Series in decision and control, Wiley, New York,. [by] David G. Luenberger. illus. 23 cm.

Madan, D. & Milne, F. (1994), 'Contingent claims valued and hedged by pricing and investing in a basis', *Mathematical Finance* **4**(3), 223–245.

Malliavin, P. & Letac, G. (1995), *Integration and probability*, Graduate texts in mathematics ; 157., Springer-Verlag, New York. Paul Malliavin, in cooperation with Hélène Airault, Leslie Kay, Gérard Letac. Accompanied by: Exercises and solutions manual for Integration and probability / Gérard Letac. viii, 142 p. ; 24 cm. Includes index.

Merton, R. C. (1973), 'Theory of rational option pricing', *Bell Journal of Economics and Management Science* **4**, 141–183.

Mesbahi, M. & Papavassilopoulos, G. P. (1997), 'On the rank minimization problem over a positive semidefinite matrix inequality.', *IEEE Transactions on Automatic Control* **42**(2), 239–243.

Milevsky, M. & Posner, S. (1998), 'Asian options, the sum of lognormals and the reciprocal gamma distribution.', *Journal of Financial and Quantitative Analysis* **33**(3), 409–422.

Miltersen, K., Sandmann, K. & Sondermann, D. (1995), 'Closed form term structure derivatives in a heath-jarrow-morton-model with log-normal annually compounded interest rates', *Proceedings of the Seventh Annual European Futures Research Symposium, Chicago Board of Trade* pp. 145–164.

Miltersen, K., Sandmann, K. & Sondermann, D. (1997), 'Closed form solutions for term structure derivatives with log-normal interest rates', *Journal of Finance* **52**(1), 409–430.

Monteiro, R. (1997), 'Primal-dual path following algorithms for semidefinite programming', *SIAM Journal on Optimization* **7**, 663–678.

Musiela, M. & Rutkowski, M. (1997), *Martingale methods in financial modelling*, Vol. 36 of *Applications of mathematics*, Springer, Berlin.

Nesterov, I. E. & Nemirovskii, A. S. (1994), *Interior-point polynomial algorithms in convex programming*, Society for Industrial and Applied Mathematics, Philadelphia.

Nesterov, I. E. & Todd, M. (1997), 'Self-scaled barriers and interior-point methods in convex programming', *Mathematics of Operations Research* **22**, 1–42.

Nesterov, I. E. & Todd, M. (1998), 'Primal-dual interior-point methods for self-scaled cones.', *SIAM Journal on Optimization* **8**, 324–364.

Nualart, D. (1995), *The Malliavin calculus and related topics*, Springer-Verlag, New York.

Overton, M. L. & Womersley, R. S. (1992), 'On the sum of the largest eigenvalues of a symmetric matrix.', *SIAM Journal of Matrix Anal. Appl.* **13**, 41–45.

Poljak, S. & Wolkowicz, H. (1993), 'Convex relaxations of 0-1 quadratic programming', *Working paper, University of Waterloo* .

Potra, F. A. & Sheng, R. (1995), Homogeneous interior–point algorithms for semidefinite programming, Technical report, Reports on Computational Mathematics. Department of Mathematics, The University of Iowa.

Rebonato, R. (1998), *Interest-Rate Options Models*, Financial Engineering, Wiley.

Rebonato, R. (1999), 'On the simultaneous calibration of multi-factor log-normal interest rate models to black volatilities and to the correlation matrix.', *Working paper* .

Rockafellar, R. T. (1970), *Convex Analysis*, Princeton University Press., Princeton.

Romagnoli, S. & Vargiolu, T. (2000), 'Robustness of the black-scholes approach in the case of options on several assets', *Finance and Stochastics* **4**, 325–341.

Sandmann, K. & Sondermann, D. (1997), 'A note on the stability of lognormal interest rate models and the pricing of eurodollar futures', *Mathematical Finance* **7**, 119–128.

Schwartz, S. C. & Yeh, Y. S. (1981), 'On the distribution function and moments of power sums with log-normal components.', *The Bell System Technical Journal* **61**(7), 1441–1462.

Sidenius, J. (1998), 'Libor market models in practice', *Working paper, Skandinaviska Enskilda Banken,Copenhagen.* .

Singleton, K. J. & Umantsev, L. (2001), 'Pricing coupon-bond options and swaptions in affine term structure models', *Working paper, Stanford University Graduate School of Business.* .

Stroock, D. W. (1987), 'Homogeneous chaos revisited.', *Séminaire de Probabilités XXI, Lectures Notes in Math.* **1247**, 1–8.

Sturm, J. F. (1999), 'Using sedumi 1.0x, a matlab toolbox for optimization over symmetric cones.', *Working paper, Department of Quantitative Economics, Maastricht University, The Netherlands.* .

Tikhonov, A. N. (1963), 'Solution of incorrectly formulated problems and the regularization method', *Soviet Math. Dokl.* (4).

Todd, M. & Yildirim, E. A. (1999), 'Sensitivity analysis in linear programming and semidefinite programming using interior-points methods.', *Working paper, School of Operation Research and Industrial Engineering, Cornell University.* .

Toh, K. C., Todd, M. J. & Tutuncu, R. H. (1996), Sdpt3 – a matlab software package for semidefinite programming, Technical report, School of Operations Research and Industrial Engineering, Cornell University.

Ustunel, A. S. & Zakai, M. (1989), 'On independence and conditioning on wiener space.', *Annals of Probability* **17**(4), 1441–1453.

Vandenberghe, L. & Boyd, S. (1996), 'Semidefinite programming', *SIAM Review* **38**, 49–95.

Vasicek, O. (1977), 'An equilibrium characterization of the term structure', *Journal of Financial Economics* (5).