SEDE AMMINISTRATIVA: UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI MATEMATICA

SCUOLA DI DOTTORATO IN SCIENZE MATEMATICHE
INDIRIZZO: MATEMATICA COMPUTAZIONALE
CICLO: XXIV

# Wishart processes: theory and applications in mathematical finance

*Direttore della Scuola*

Chiar.mo Prof. Paolo Dai Pra

*Coordinatore d'Indirizzo*

Chiar.mo Prof. Paolo Dai Pra

*Supervisori*

Chiar.mo Prof. Martino Grasselli

Chiar.mo Prof. Wolfgang Johann Runggaldier

*Dottorando*
Alessandro Gnoatto

# Wishart processes: theory and applications in mathematical finance

Alessandro Gnoatto

VIA IOLANDA, 15
36056, TEZZE SUL BRENTA - ITALY
*Current address*: Via Trieste, 63
35121, Padova - Italy
*E-mail address*: alessandro@alessandrognoatto.com
*URL*: http://www.alessandrognoatto.com

1991 *Mathematics Subject Classification.* Primary 91B26, 91B25, 60H30; Secondary 91B30, 91B70, 60H10

# Aknowledgments

First and foremost I would like to thank my supervisors, Prof. Martino Grasselli and Prof. Wolfgang Johann Runggaldier for assisting me during these intense years of research. Prof. Grasselli, in particular, introduced me to the study of mathematical finance and has constantly encouraged my interest in this area of mathematics. He was not only a supervisor but a true friend. I am indebted to Prof. Paolo Dai Pra and the board of the graduate school in mathematical sciences for promoting a perfect environment for my PhD studies.

This PhD thesis is also the result of many collaborations and stimulating experiences with people from the Academia and the Industry. First and foremost I would like to thank Prof. José Da Fonseca, who is co-author of the paper on Libor rates which constitutes the topic of Chapter 3. Prof. Da Fonseca was an invaluable source of suggestions for the numerical implementations and experiments on the Wishart process that can be found throughout the thesis.

During the course of my PhD studies I was granted the opportunity to study at the ETH in Zurich. There I had the privilege to work and interact with many top researchers in the field. I feel particularly grateful to Prof. Josef Teichmann, Prof. Martin Keller-Ressel and Dr. Christa Cuchiero for their very interesting lectures and for the many fruitful discussions.

The final part of this thesis is devoted to the study of foreign exchange markets. The opportunity was given by a collaboration with Prof. Grasselli and the head of the foreign exchange model validation team at UBS, Dr. Alvise De Col. Dr. De Col is co-author of the paper on foreign exchange options which constitutes the topic of Chapter 4 of the present thesis. It was a real pleasure for me to work with him as he also gave a substantial contribution in terms of hints coming from the world of practitioners.

I also feel very grateful to Michele Leoncelli and his group at Prometeia Spa.

Last but not least, I would like to thank the most important people in my life: my family and Elisa.

# Contents

# Prefazione

I mercati finanziari sono caratterizzati da un crescente livello di complessità. Gli ultimi decenni hanno visto la comparsa di nuove asset classes, come ad esempio le materie prime o la volatilità, e l'introduzione di nuovi prodotti finanziari sempre più complessi. Questi nuovi strumenti sono caratterizzati o da dipendenze non banali dai movimenti di mercato di un singolo sottostante, oppure dalla dipendenza simultanea da più titoli primari. Ciò evidenzia come il crescente grado di complessità nei mercati finanziari sia ascrivibile a due fenomeni naturali: l'introduzione di nuovi strumenti volti a soddisfare le esigenze di investitori via via più esperti, o il ruolo crescente giocato dal co-movimento di entità finanziarie. In particolare, lo studio dei movimenti congiunti di entità finanziarie, con particolare riferimento alle recenti crisi finanziarie, risulta di particolare importanza. Le recenti turbolenze finanziarie possono essere pensate anche come dei fallimenti nell'ambito del risk-management, dovuti prevalentemente ad un insieme di pratiche di mercato prive di reale fondamento, che si sono tradotte spesso nella sottostima del rischio implicito in determinate strutture complesse. Queste pratiche di mercato discutibili sembrano prevalentemente di due tipi: sottostima della possibilità di osservare valori estremi di variabili aleatorie da una parte, studio carente dei movimenti congiunti delle variabili stesse dall'altra. Dal punto di vista della gestione del rischio, i campi coinvolti sono la teoria dei valori estremi e le copule, si vedano McNeil et al. (2005), Embrechts et al. (1997).

Se da un lato la crisi sottolinea determinate pratiche di mercato infondate applicate dall'industria finanziaria, dall'altro anche i ricercatori nell'ambito della finanza matematica dovrebbero trarre alcuni insegnamenti dalla situazione attuale. Maggiore energia dovrebbe essere investita nella ricerca di modelli in grado di descrivere portafogli di grandi dimensioni o fattori di rischio multi-fattoriali. Questa tesi vuole rappresentare un passo in questa direzione.

In questo lavoro saranno introdotti nuovi e promettenti modelli, in un contesto multi-fattoriale o in presenza di più sottostanti. Nell'opinione dell'autore questi modelli costituiscono dei candidati interessanti per applicazioni a livello industriale. L'elevato grado di trattabilità analitica di questi modelli è dovuto alla particolare classe di processi stocastici che si è scelto di impiegare che è quella dei processi (esponenzialmente) affini, i quali costituiscono un insieme di processi Markoviani per i quali la funzione caratteristica può essere scritta in una forma molto esplicita.

Verranno impiegati processi affini a valori in due spazi, lo spazio "canonico" $\mathbb{R}^n_{\geq 0} \times \mathbb{R}^m$, considerato in Duffie et al. (2003), e il cono delle matrici $d \times d$ semidefinite positive, che è stato caratterizzato in Cuchiero et al. (2009). La maggior parte dei modelli sarà sviluppata sotto quest'ultima assunzione sullo spazio.

Il principale esempio che verrà considerato è il processo di Wishart, un processo puramente diffusivo a valori in $S_d^+$ che può essere visto come l'analogo matriciale del classico processo square root abitualmente utilizzato nelle applicazioni finanziarie. Il processo di Wishart è stato introdotto in Bru (1991) dove viene costruito a partire dal quadrato di moti Browniani matriciali e in seguito, nella sua forma più generale, come quadrato di processi Ornstein-Uhlenbeck matriciali.

Negli ultimi anni si è assistito all'introduzione di numerosi modelli basati su questo tipo di processo, prevalentemente in quanto matrici semidefinite positive emergono naturalmente in finanza, ad esempio sotto forma di matrici di varianze e covarianze, come nell'ambito dell'ottimizzazione di portafoglio. Altri esempi di processi appartenenti a questa classe sono dei processi composti di Poisson aventi intensità costante e distribuzione dei salti con supporto dato da $S_d^+$. Nella presente tesi vengono proposti diversi contributi alla letteratura sui processi affini e le loro applicazioni in finanza matematica.
Nel Capitolo 1 viene considerato il processo di Wishart e viene proposta una generalizzazione dei risultati di Bru in merito alla trasformata di Fourier/Laplace del processo. Test numerici mostrano l'elevato grado di accuratezza del nuovo approccio, che viene confrontato con gli approcci esistenti in letteratura.
La seconda parte della tesi è dedicata a modelli per il mercato dei tassi. Nel capitolo 2 viene presentato un modello per il tasso a breve mosso dal processo di Wishart, inizialmente proposto da Grasselli and Tebaldi (2008) e in seguito analizzato da Buraschi et al. (2008) e Chiarella et al. (2010) rispetto a questioni quali i fatti stilizzati nell'ambito del mercato dei tassi e la scelta del premio al rischio. Il contributo della tesi in questo ambito è duplice: viene introdotta una nuova formula per il prezzaggio dei titoli di puro sconto e viene inoltre presentato un insieme di condizioni sufficienti tali da garantire che il modello replichi determinate forme della struttura a termine dei tassi. La dimostrazione è ispirata alle tecniche introdotte in Keller-Ressel and Steiner (2008), ottenute sotto l'assunzione che il tasso a breve fosse descritto da un processo affine scalare. Il nuovo risultato non fornisce una tassonomia completa delle curve che possono essere replicate, tuttavia può essere utile in sede di calibrazione del modello in quanto è possibile specificare dei vincoli sullo stato iniziale del processo, tali che il modello replichi la forma della curva osservata sul mercato. Viene infine proposta un'analisi dell'impatto dei parametri del modello sulla struttura a termine dei tassi, fornendo un buon livello di intuizione.
Nel capitolo 3 viene esteso il modello di Keller-Ressel et al. (2009) per la valutazione di caps floors e swaptions in un contesto multi-fattoriale. Ciò non implica semplicemente la proposizione di un nuovo modello, bensì di un'intera classe di possibili modelli, dal momento che la costruzione generale dell'approccio può essere effettuata per qualsiasi processo a valori nel cono delle matrici strettamente definite positive. Ciò significa che possono essere analizzati modelli Libor puramente diffusivi, puramente a salti oppure diffusivi con salti. Per illustrare la metodologia, particolare attenzione è riservata al modello Libor di Wishart. La nuova metodologia sembra interessante in quanto scevra dai problemi che caratterizzano gli approcci tradizionali dei modelli di mercato Libor. A titolo di esempio si ricorda la ben nota inconsistenza tra i modelli di mercato Libor e Swap. Dal momento che la grandezza di riferimento è il rapporto di prezzi di titoli di puro sconto è possibile esprimere entrambi i problemi di valutazione attraverso questo rapporto, evitando così l'assunzione sin troppo semplicistica che il tasso swap sia mosso da un processo scalare, che non tiene conto di rilevanti fenomeni correlazione. Grazie alla buona trattabilità analitica dei dei processi affini in $S_d^{++}$, viene presentata una formula di valutazione semi-chiusa per i caplet. Per quanto concerne le swaptions, si mostra che è possibile applicare l'espansione di Edgeworth come in Collin-Dufresne and Goldstein (2002) , che permette di approssimare le probabilità di esercizio coinvolte nella formula di valutazione. Vengono presentati degli esperimenti numerici che consistono nello studio delle superfici di volatilità generate dal modello Libor di Wishart, dando così una prima dimostrazione delle potenzialità dell'approccio.
La terza e ultima parte della tesi è dedicata al mercato dei tassi di cambio. Si tratta di un mercato molto liquido e di grandi dimensioni, dove la descrizione delle correlazioni è di fondamentale importanza dal momento che il prodotto/la frazione di due tassi di cambio è ancora un tasso di cambio. Lo scopo della terza parte è fornire un'estensione del modello classico di Garman and Kohlhagen (1983), per la

valutazione di opzioni europee in presenza di volatilità stocastica. Vengono proposti due modelli: il capitolo 4 introduce un modello di Heston multi-fattoriale teoricamente coerente con relazioni triangolari tra tassi di cambio. In primis il focus è sui tassi di cambio rispetto ad un numerario universale (che potrebbe essere l'oro, a titolo di esempio), in seguito i tassi di cambio tra le valute sono costruiti come rapporti o prodotti a partire da questi tassi di cambio base. Così procedendo la struttura del modello è modulare e consente pertanto la valutazione di un intero book di derivati sui tassi, indipendentemente dal numero di valute coinvolte. Il modello di Heston multi-fattoriale è caratterizzato dalla presenza di uno skew e da una matrice di varianze e covarianze entrambi stocastici. La presenza di più misure risk-neutral, legate ai diversi paesi, richiede precise relazioni tra i parametri del modello nelle differenti economie. Tale aspetto viene completamente chiarito. Nuovamente, grazie alle proprietà dei processi affini, viene derivata una formula chiusa per la trasformata di Fourier/Laplace del tasso di cambio logaritmico. Inoltre, vengono presentate delle espansioni asintotiche che forniscono un'alternativa veloce per la calibrazione del modello limitatamente a maturità basse. Il modello è stato calibrato con successo contemporaneamente su tre superfici di volatilità di un triangolo di valute.
Il capitolo 5 presenta un'estensione del modello precedente, considerando il processo di Wishart al posto di un vettore di processi square root. Di interesse è il fatto che tutte le proprietà del precedente modello sono conservate. Quest'ultimo approccio potenzialmente si presta a descrivere fenomeni di correlazione più complessi e pertanto sembra essere promettente nell'ambito della valutazione di derivati esotici complessi.

# Preface

Modern financial markets are deeply characterised by an incresing level of complexity. The last few decades witnessed the appearence of new asset classes, such as commodities or volatility, and the emergence of new and sophisticated financial instruments. These products may involve complex dependencies on the market movements of a single reference underlying, or may depend on the price pattern of more securities simultaneously. Thus we realize that complexity in financial markets arises from two natural phenomena: the introduction of new instruments aiming at satisfying the qualified demand from more and more sophisticated investors, or the increasing role played by the co-movement of financial entities. In particular, co-movement is important, particularly with reference to the 2007-2009 financial crisis. This recent turmoil which was observed on the market may be seen also as a major risk management failure, due also to a set of market-bad-practices, resulting in an underestimation of the risk involved in complex structures. These market practices may be classified as misunderstanding of the behaviour of extreme values, and joint movements of random variables. From a risk management perspective, the fields which are involved in this sense are extreme value theory and copulas, see McNeil et al. (2005), Embrechts et al. (1997).

As the financial crisis stresses some critical bad practices in the financial industries, also researchers in the field of mathematical finance should learn some lessons from the present situation. More energies should be invested in the search for models which are able to describe larger portfolios or multifactor sources of risk. The present thesis is an attempt towards this direction.

We will introduce new and powerful models, either in the multifactor or in the multiasset setting. These models will be shown to exhibit a level of descriptive power and analytical tractability such that they constitute in our opinion strong candidates for real life applications. The analytical tractability will be ensured by the specific class of stochastic processes we decide to adopt. In particular, throughout this thesis, we will be always working with affine processes, which constitute a class of stochastically continuous Markov processes for which the characteristic function can be written in a very explicit form. We will consider affine processes taking values on two different state spaces: the standard polyhedral state space $\mathbb{R}^n_{\geq 0} \times \mathbb{R}^m$, already considered in Duffie et al. (2003), and the cone of positive semidefinite $d \times d$ matrices that we will denote by $S_d^+$, which has been characterised in Cuchiero et al. (2009). Most examples and models will be developed under this last general state space.

The most important stochastic process we will be working with is the Wishart process, which is a pure diffusion stochastic process on $S_d^+$ and may be seen as a matrix analogue of the standard square root process which is commonly applied in finance. The Wishart process has been introduced in Bru (1991) where it is constructed first as a square of a matrix Brownian motion and then, in its general form, as the square of a matrix Ornstein-Uhlenbeck process. The last few years have witnessed the introduction of many models driven by this stochastic process, mainly because in finance positive semidefinite matrices arise very naturally, an immediate example being given by the necessity to consider variance-covariance matrices e.g. for portfolio optimization purposes. Other examples include Poisson processes with constant intensity and jump distribution with support on $S_d^+$ and then also jump-diffusion processes.

By working with these processes, in the present thesis we propose several contributions to the literature on affine processes and their applications in mathematical finance. In Chapter 1 we consider the Wishart process per se, and provide a generalization of a result due to Bru (1991), thus providing an alternative explicit formula for the computation of the Fourier/Laplace transform of the Wishart process and its time integral. Numerical tests show the high degree of accuracy of this approach.

The second part of the thesis is devoted to models for the fixed income market. In Chapter 2 we consider the Wishart process as driving noise for a short rate model. This kind of short rate model has already been considered in the literature: it was first hinted at in Grasselli and Tebaldi (2008), and then thoroughly analyzed in Buraschi et al. (2008) and Chiarella et al. (2010) with respect to issues like stylized facts in the fixed income market and the impact of the specification of the risk premium. Our contribution is concerned with the presentation of a new closed form solution for the prices of bonds under the same conditions presented in Chapter 1. Moreover, for this short rate model we then derive a set of sufficient conditions ensuring that the yield curve replicates a set of basic shapes. The derivation of these sufficient conditions is inspired by the arguments due to Keller-Ressel and Steiner (2008), which were obtained under the assumption that the short rate model is driven by a scalar affine process. This result does not provide a full taxonomy of all possible shapes of the yield curve that can be attained, however it is of interest in a calibration perspective, since we can provide a constraint on the initial state of the process, such that the short rate model replicates some of the shapes of the curve that may be observed on the market. Finally we also propose a simple analysis of the impact of the rich family of model parameters on the yield curve, providing a good level of intuition.

In Chapter 3 we extend the methodology introduced by Keller-Ressel et al. (2009) for the evaluation of caps floors and swaptions to the multifactor setting. In doing this we do not introduce a single model, but a whole family of models, since the general construction of the model can be carried out for any process on the interior of the cone $S_d^+$: it means that we can think about pure diffusion Libor models driven by a Wishart process, or we can treat compound Poisson processes or even jump-diffusions. To illustrate the methodology, we concentrate on the Wishart Libor model. This approach is very interesting because it provides a modelling framework which is free from some known problematic issues arising with standard market models. To give an example we can name the well known inconsistency between the Libor and the swap market model: since our reference quantity is a ratio of zero coupon bonds, we are able to express both pricing problems in terms of this ratio, thus avoiding the oversimplifying assumption that the swap rate is driven by a scalar process, which does not take into account all relevant correlation effects. Thanks to the good analytical tractability of affine processes on $S_d^{++}$, we derive a closed form valuation formula for caplets. As far as swaptions are concerned, we find that it is possible to employ the cumulant expansion introduced in Collin-Dufresne and Goldstein (2002) for the approximation of the exercise probabilities appearing in the general (model-free) valuation formula. We perform numerical experiments on caplet implied volatility surfaces generated by the Wishart Libor model, thus providing a first view on the potential of this methodology.

The third and final part of the thesis is concerned with the foreign exchange market. This is a very liquid and large market, where a precise description of correlations is a crucial point since suitable ratios/products of exchange rates are still exchange rates. Our aim here is to provide an extension of the standard Garman and Kohlhagen (1983) model, in order to evaluate European plain vanilla derivatives in the presence of stochastic volatility. We propose two models: in Chapter 4 we outline an extension of the multi-Heston model which is theoretically coherent with triangular relations among exchange rates. We look first at exchange rates with respect to a universal numeraire (which may be gold, just to provide an example), and then construct exchange rates among currencies as ratios of these basic exchange rates with respect to the universal numeraire. In this way the structure of the model is completely modular and is thus able to provide a consistent valuation approach for a whole book of foreign exchange derivatives, no matter the number of currencies involved. The very structure of the model is consistent with triangular relations among currencies, which we observe since suitable products/ratios of exchange rates are still exchange rates. Our FX multi-Heston model can account for a stochastic skew and for a stochastic covariance matrix. The co-existence of many risk neutral measures on the market (one for each country) requires a detailed discussion of the relation between the parameters of the model in the different economies. We provide this discussion and provide a simple but effective condition on the measure changes such that no arbitrage holds.

Again, thanks to the affine property, we derive a closed form formula for the Fourier/Laplace transform of the log-exchange rate. After that, we propose some asymptotic expansions which may provide an alternative and faster calibration of the model when we restrict our attention to short maturities. We perform a calibration to the three volatility surfaces of a standard currency triangle and conclude that we are in front of the first stochastic volatility model able to capture the shape of the three surfaces simultaneously.

Finally, in Chapter 5 we propose an ambitious generalization of the setup introduced in Chapter 4, by considering a Wishart process as driver for the stochastic volatility and replicate the procedure introduced in Chapter 4. It is interesting to note that we are able to recover all the nice properties of the previous approach, while we introduce a setting which has much more flexibility to describe complex correlation phenomena. This last model is promising in view of the evaluation of complex exotic derivatives written on multiple currencies.

# Notations and General Results

## Notations

| | |
|---|---|
| **Spaces of Matrices** | |
| $M_d$ | Set of $d \times d$ square matrices. |
| $S_d$ | Cone of symmetric $d \times d$ matrices. |
| $S_d^+$ | Cone of symmetric $d \times d$ positive semidefinite matrices. |
| $S_d^{++}$ | Cone of symmetric $d \times d$ positive definite matrices. |
| $S_d^-$ | Cone of symmetric $d \times d$ negative semidefinite matrices. |
| $S_d^{--}$ | Cone of symmetric $d \times d$ negative definite matrices. |
| $Tr$ | Trace operator. |
| $\succeq$ | Partial order relation on $S_d^+$. Given $U, V \in S_d^+$ we write $U \succeq V$ if $U - V \in S_d^+$. |
| **Matrix Functions** | |
| $\sqrt{X}$ | Square root of the matrix $X \in S_d^+$. We write $X = P^\top Diag(\lambda_1, ..., \lambda_d)P$, where $\lambda_i, i = 1, .., d$ are the eigenvalues of $X$ and $P$ is a unitary matrix. Then $\sqrt{X} = P^\top Diag(\sqrt{\lambda_1}, ..., \sqrt{\lambda_d})P$. |
| $e^X$ | Matrix exponential defined as: $e^X = \sum_{k=0}^{\infty} \frac{X^k}{k!}$. |
| $\sinh(X)$ | $\frac{e^X - e^{-X}}{2}$. |
| $\cosh(X)$ | $\frac{e^X + e^{-X}}{2}$. |
| **Chapter 1** | |
| $WIS_d(S_0, \alpha, M, Q)$ | Law of the Wishart process of dimension $d$, initial state $S_0$, Gindikin parameter/matrix $\alpha$, mean reversion matrix $M$ and volatility matrix $Q$. |
| $\psi(t)$ | Coefficient of the linear part of the cumulant generating function of the process $X$. |
| $\phi(t)$ | Constant part of the cumulant generating function of the process $X$. |
| $\psi'$ | A solution of the algebraic Riccati ODE. |

| | |
|---|---|
| **Chapter 2** | |
| $\psi(\tau, B)$ | Coefficient of the linear part of the cumulant generating function of the process $X$. |
| $\phi(\tau, B)$ | Constant part of the cumulant generating function of the process $X$. |
| $\tilde{\psi}(\tau, B), \tilde{\psi}(\tau)$ | Coefficient of the linear component of the bond log-price. |
| $\tilde{\phi}(\tau, B), \tilde{\psi}(\tau)$ | Constant component of the bond log-price. |
| $\sigma(X)$ | The set of the eigenvalues of the matrix X. |
| $\lambda(X)$ | An eigenvalue of the matrix X. |
| **Chapter 3** | |
| $Wis_d\left(n, \mathcal{Q}\right)$ | Wishart distribution. |
| $Wis_d\left(n, \mathcal{Q}, \mathcal{M}\right)$ | Non-central Wishart distribution. |
| $\beta_d^I(a, b)$ | Beta type I distribution. |
| $\beta_d^{II}(a, b)$ | Beta type II distribution. |
| ${}_mF_n$ | Hypergeometric function of matrix argument. |
| $\Gamma_d(a)$ | Multivariate Gamma function. |
| $\Psi(a; b; R)$ | Confluent hypergeometric function. |

## General Results

For the reader's convenience, we report here some results which may be found in Cuchiero et al. (2009), which constitute the theoretical framework we will be working with. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space, with the filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfying the usual assumptions. Let $S_d^+$ denote the cone of positive semidefinite $d \times d$ matrices, endowed with the scalar product $\langle x, y\rangle = Tr\,[xy]$. We consider a Markov process $X = (X_t)_{t\geq 0}$ with state space $S_d^+$, transition probability $p_t(X_0, A) = \mathbb{P}(X_t \in A)$ for $A \in S_d^+$, and transition semigroup $(P_t)_{t\geq 0}$ acting on bounded functions $f \in S_d^+$:

DEFINITION 0.1. *(Cuchiero et al. (2009) Definition 2.1) The Markov process $X$ is called affine if:*

1. *it is stochastically continuous, that is,* $\lim_{s\to t} p_s(X_0, \cdot) = p_t(X_0, \cdot)$ *weakly on* $S_d^+$ $\forall t$, $X_0 \in S_d^+$, *and*
2. *its Laplace transform has exponential-affine dependence on the initial state:*

$$P_t e^{-Tr[uX_0]} := \int_{S_d^+} e^{-Tr[u\xi]} p_t(X_0, d\xi) = e^{-\phi(t,u) - Tr[\psi(t,u)X_0]}, \tag{0.1}$$

$\forall t$ *and* $u, X_0 \in S_d^+$, *for some functions* $\phi : \mathbb{R}_{\geq 0} \times S_d^+ \to \mathbb{R}_{\geq 0}$ *and* $\psi : \mathbb{R}_{\geq 0} \times S_d^+ \to S_d^+$.

Note that in the definition above we assumed that the process is stochastically continuous, a feature that implies, according to Proposition 3.4 in Cuchiero et al. (2009), that the process is regular in the following sense:

DEFINITION 0.2. *(Cuchiero et al. (2009) Definition 2.2) The affine process $X$ is called regular if the derivatives*

$$F(u) = \frac{\partial \phi(t, u)}{\partial t}|_{t=0^+}, \qquad R(u) = \frac{\partial \psi(t, u)}{\partial t}|_{t=0^+} \tag{0.2}$$

*exist and are continuous at* $u = 0$.

DEFINITION 0.3. *(Cuchiero et al. (2009) Definition 2.3) Let $\chi : S_d^+ \to S_d^+$ be some bounded continuous truncation function with $\chi(\xi) = \xi$ in a neighborhood of 0. An admissible parameter set $(\alpha, b, \beta^{ij}, c, \gamma, m, \mu)$ associated with $\chi$ consists of:*

- *a linear diffusion coefficient*

$$\alpha \in S_d^+, \tag{0.3}$$

- *a constant drift term*

$$b \succeq (d-1)\alpha, \tag{0.4}$$

- *a constant killing rate term*

$$c \in \mathbb{R}^+, \tag{0.5}$$

- *a linear killing rate coefficient*

$$\gamma \in S_d^+, \tag{0.6}$$

- *a constant jump term: a Borel measure $m$ on $S_d^+ \setminus \{0\}$ satisfying*

$$\int_{S_d^+ \setminus \{0\}} (\| \xi \| \wedge 1)\, m(d\xi) < \infty, \tag{0.7}$$

- *a linear jump coefficient: a $d \times d$ matrix $\mu = (\mu_{ij})$ of finite signed measure on $S_d^+ \setminus \{0\}$, such that $\mu(E) \in S_d^+$ $\forall E \in \mathcal{B}(S_d^+)$ and the kernel*

$$M(x, d\xi) := \frac{Tr\,[x\mu(d\xi)]}{\| \xi \|^2 \wedge 1} \tag{0.8}$$

*satisfies*

$$\int_{S_d^+ \setminus \{0\}} Tr\,[\chi(\xi)u]\, M(x, d\xi) < \infty, \tag{0.9}$$

$\forall x, u \in S_d^+$ *s.t.* $Tr\,[xu] = 0$.

- *a linear drift coefficient: a family $\beta^{ij} = \beta^{ji} \in S_d^+$ s.t. the linear map $\beta : S_d \to S_d$ of the form*

$$\beta(x) = \sum_{i,j} \beta^{ij} x_{ij}, \tag{0.10}$$

*satisfies*

$$Tr\,[\beta(x)u] - \int_{S_d^+ \setminus \{0\}} Tr\,[\chi(\xi)u]\, M(x, d\xi) \geq 0 \tag{0.11}$$

$\forall x, u \in S_d^+$ *with* $Tr\,[xu] = 0$.

The following theorem closes our survey on affine processes. It is a generalization of the result by Duffie et al. (2003) to the state space $S_d^+$. Denote by $\mathcal{S}_d$ the space of rapidly decreasing real valued $C^\infty$-functions on $S_d^+$ (for their definition, see Cuchiero et al. (2009)) and let $\mathcal{D}(\mathcal{A})$ be the domain of the generator of the process.

THEOREM 0.1. *(Cuchiero et al. (2009) Theorem 2.4) Suppose $X$ is an affine process on $S_d^+$. Then $X$ is regular and has the Feller property. Let $\mathcal{A}$ be its infinitesimal generator on $C_0(S_d^+)$. Then $\mathcal{S}_d \subset \mathcal{D}(\mathcal{A})$ and there exists an admissible parameter set $\left(\alpha, b, \beta^{ij}, c, \gamma, m, \mu\right)$ associated to the truncation function $\xi$ such that, for $f \in \mathcal{S}_d$*

$$\begin{aligned}\mathcal{A}f(x) = \frac{1}{2}\sum_{i,j,k,l} A_{ijkl}(x)\frac{\partial^2 f(x)}{\partial x_{ij}\partial x_{kl}} + \sum_{i,j}\left(b_{ij} + \beta_{ij}(x)\right)\frac{\partial f(x)}{\partial x_{ij}} - \left(c + Tr\left[\gamma x\right]\right) f(x) \\ + \int_{S_d^+\setminus\{0\}} \left(f(x+\xi) - f(x)\right) m(d\xi) \\ + \int_{S_d^+\setminus\{0\}} \left(f(x+\xi) - f(x) - Tr\left[\chi(\xi)\nabla f(x)\right]\right) M(x, d\xi)\end{aligned} \tag{0.12}$$

*where $\beta(x)$ is defined by* (0.10)*, $M(x, d\xi)$ by* (0.8) *and*

$$A_{ijkl}(x) = x_{ik}\alpha_{jl} + x_{il}\alpha_{jk} + x_{jk}\alpha_{il} + x_{jl}\alpha_{ik} \tag{0.13}$$

*Moreover, $\phi(t,u)$ and $\psi(t,u)$ in Definition 0.1 solve the generalized Riccati differential equations, for $u \in S_d^+$,*

$$\frac{\partial \phi(t,u)}{\partial t} = F(\psi(t,u)), \qquad \phi(0,u) = 0, \tag{0.14}$$

$$\frac{\partial \psi(t,u)}{\partial t} = R(\psi(t,u)), \qquad \phi(0,u) = u, \tag{0.15}$$

*with*

$$F(u) = Tr\left[bu\right] + c - \int_{S_d^+\setminus\{0\}} \left(e^{-Tr[u\xi]} - 1\right) m(d\xi), \tag{0.16}$$

$$\begin{aligned}R(u) &= -2u\alpha u + \beta^T(u) + \gamma \\ &\quad - \int_{S_d^+\setminus\{0\}} \left(\frac{e^{-Tr[u\xi]} - 1 - Tr\left[\chi(\xi)u\right]}{\parallel \xi \parallel^2 \wedge 1}\right) \mu(d\xi),\end{aligned} \tag{0.17}$$

*where $\beta_{ij}^T(u) = Tr\left[\beta^{ij}u\right]$.*
*Conversely, let $\left(\alpha, b, \beta^{ij}, c, \gamma, m, \mu\right)$ be an admissible parameter set associated to the truncation function $\xi$. Then there exists a unique affine process on $S_d^+$ with infinitesimal generator given by* (0.12) *and such that the affine transform formula* (0.1) *holds for all $(t,u) \in \mathbb{R}_{\geq 0} \times S_d^+$, where $\phi(t,u)$ and $\psi(t,u)$ are given by* (0.14) *and* (0.15)*.*

# Part 1

# Wishart processes

CHAPTER 1

# The explicit Laplace transform of the Wishart process

## 1. Introduction

In this chapter we propose an analytical approach for the computation of the moment generating function for the Wishart process which has been introduced by Bru (1991), as an extension of square Bessel processes (Pitman and Yor (1982), Revuz and Yor (1994)) to the matrix case. Wishart processes belong to the class of affine processes and they generalise the notion of positive factor insofar as they are defined on the set of positive semidefinite real $d \times d$ matrices, denoted by $S_d^+$.

Given a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ satisfying the usual assumptions and a $d \times d$ matrix Brownian motion $B$ (i.e. a matrix whose entries are independent Brownian motions under $\mathbb{P}$), a Wishart process on $S_d^+$ is governed by the SDE

$$dS_t = \sqrt{S_t} dB_t Q + Q^\top dB_t^\top \sqrt{S_t} + \left(MS_t + S_t M^\top + \alpha Q^\top Q\right) dt, \quad S_0 \in S_d^+, \quad t \geq 0 \tag{1.1}$$

where $Q \in GL_d$ (the set of invertible real $d \times d$ matrices), $M \in M_d$ (the set of real $d \times d$ matrices) with all eigenvalues on the negative real line in order to ensure stationarity, and where the (Gindikin) real parameter $\alpha > d - 1$ grants the positive semi definiteness of the process, in analogy with the Feller condition for the scalar case (Bru (1991)). In the dynamics above $\sqrt{S_t}$ denotes the square root in matrix sense. We denote by $WIS_d(S_0, \alpha, M, Q)$ the law of the Wishart process $(S_t)_{t \geq 0}$. The starting point of the analysis was given by considering the square of a matrix Brownian motion $S_t = B_t^\top B_t$, while the generalization to the particular dynamics (1.1) was introduced by looking at squares of matrix Ornstein-Uhlenbeck processes (see Bru (1991)).

Bru proved many interesting properties of this process, like non-collision of the eigenvalues (when $\alpha \geq d + 1$) and the additivity property shared with square Bessel processes. Moreover, she computed the Laplace transform of the Wishart process and its integral (the *Matrix Cameron-Martin formula* using her terminology), which plays a central role in the applications:

$$\mathbb{E}_{S_0}^{\mathbb{P}} \left[ \exp \left\{ -Tr \left[ wS_t + \int_0^t vS_s ds \right] \right\} \right], \tag{1.2}$$

where $Tr$ denotes the trace operator and $w, v$ are symmetric matrices for which the expression (1.2) makes sense. Bru found an explicit formula for (1.2) (formula (4.7) in Bru (1991)) under the assumption that the symmetric diffusion matrix $Q$ and the mean reversion matrix $M$ commute.

Positive (semi)definite matrices arise in finance in a natural way and the nice analytical properties of affine processes on $S_d^+$ opened the door to new interesting models which are able to overcome the shortcomings of previous affine models. In fact, the non linearity of $S_d^+$ is the key ingredient that allows for non trivial correlations among positive factors, a feature which is precluded in classic (linear) state space domains like $\mathbb{R}^n_{\geq 0} \times \mathbb{R}^m$ (see Duffie et al. (2003)). Not surprisingly, the last years have witnessed the birth of a whole branch of literature on applications of affine processes on $S_d^+$. The first proposals were formulated in Gourieroux et al. (2005), Gourieroux and Sufana (2003), Gourieroux and Sufana (2005), Gourieroux (2006) both in discrete and continuous time. Applications to multifactor volatility and stochastic correlation can be found in Da Fonseca et al. (2008), Da Fonseca et al. (2007b), Da Fonseca et al. (2009), Da Fonseca et al. (2007a), Da Fonseca and Grasselli (2011), Buraschi et al. (2010), and Buraschi et al. (2008) both in option pricing and portfolio management. These contributions consider the case of continuous path Wishart processes. As far as jump processes on $S_d^+$ are concerned we recall the proposals by Barndorff-Nielsen and Stelzer (2007), Muhle-Karbe et al. (2010) and Pigorsch and Stelzer (2009). Leippold and Trojani (2010) and Cuchiero et al. (2009) consider jump-diffusions models in this class, while Grasselli and Tebaldi (2008) investigate processes lying in the more general symmetric cones state space domain, including the interior of the cone $S_d^+$.

The main contribution of this chapter consists in relaxing the commutativity assumption made in Bru (1991) and proving that it is possibile to characterize explicitly the joint distribution of the Wishart process and its time integral for general (even not symmetric) mean-reversion and diffusion matrices satisfying the assumptions above. The proof of our general Cameron Martin formula is in line with that of theorem 2" in Bru and we will provide a step-by-step derivation.

The chapter is organized as follows: in section 2 we prove our main result, which extends the original approach by Bru. In section 3 we briefly review some other existing methods which have been employed in the past literature for the computation of the Laplace transform: the variation of constant, the linearization and the Runge-Kutta method. The first two methods provide analytical solutions, so they should be considered as *competitors* of our new methodology. We show that the variation of constants method is unfeasible for real-life computations, hence the truly analytic competitor is the linearization procedure.
In Section 4 we provide two applications of our result to the setting of a matrix extension of the Heston model, proposed in Da Fonseca et al. (2008) and to a stochastic correlation model, due to Da Fonseca et al. (2007b). Finally, in the Appendix we extend our formula to the case where the Gindikin term $\alpha Q^\top Q$ is replaced by a general symmetric matrix $b$ satisfying $b-(d-1)Q^\top Q \in S_d^+$ according to Cuchiero et al. (2009).

## 2. The Matrix Cameron-Martin Formula

**2.1. Statement of the result.** In this section we proceed to prove the main result of this chapter. We report a formula completely in line with the Matrix Cameron-Martin formula given by Bru (1991).

THEOREM 1.1. *Let $S \in WIS_d(S_0,\alpha,M,Q)$ be the Wishart process solving (1.1), assume*

$$M^\top \left(Q^\top Q\right)^{-1} = \left(Q^\top Q\right)^{-1} M, \tag{2.1}$$

*let $\alpha \geq d+1$ and define the set of convergence of the Laplace transform*

$$\mathcal{D}_t = \left\{ w, v \in S_d : \mathbb{E}^{\mathbb{P}}_{S_0}\left[\exp\left\{-Tr\left[wS_t + \int_0^t vS_s ds\right]\right\}\right] < +\infty \right\}.$$

*Then the joint moment generating function of the process and its integral is given by:*

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}_{S_0}\left[\exp\left\{-Tr\left[wS_t+\int_0^t vS_s ds\right]\right\}\right]\\
&=\det\left(e^{-Mt}\left(\cosh(\sqrt{\bar{v}}t)+\sinh(\sqrt{\bar{v}}t)k\right)\right)^{\frac{\alpha}{2}}\\
&\times\exp\left\{Tr\left[\left(\frac{Q^{-1}\sqrt{\bar{v}}kQ^{\top^{-1}}}{2}-\frac{\left(Q^\top Q\right)^{-1}M}{2}\right)S_0\right]\right\},
\end{aligned}$$

*where the matrices $k,\bar{v},\bar{w}$ are given by:*

$$\begin{aligned}
k&=-\left(\sqrt{\bar{v}}\cosh(\sqrt{\bar{v}}t)+\bar{w}\sinh(\sqrt{\bar{v}}t)\right)^{-1}\left(\sqrt{\bar{v}}\sinh(\sqrt{\bar{v}}t)+\bar{w}\cosh(\sqrt{\bar{v}}t)\right),\\
\bar{v}&=Q\left(2v+M^\top Q^{-1}Q^{\top^{-1}}M\right)Q^\top, \qquad (2.2)\\
\bar{w}&=Q\left(2w-\left(Q^\top Q\right)^{-1}M\right)Q^\top.
\end{aligned}$$

*Moreover,*

$$\mathcal{D}_t=\left\{w,v\in S_d:\sqrt{\bar{v}}\cosh(\sqrt{\bar{v}}t)+\bar{w}\sinh(\sqrt{\bar{v}}t),\cosh(\sqrt{\bar{v}}t)+\sinh(\sqrt{\bar{v}}t)k\in GL_d\right\}. \qquad (2.3)$$

REMARK 1.1. *In the previous formulation we recognize the exponential affine shape with respect to the state variable $S$:*

$$\mathbb{E}^{\mathbb{P}}_{S_0}\left[\exp\left\{-Tr\left[wS_t+\int_0^t vS_s ds\right]\right\}\right]=\exp\left\{-\phi(t)-Tr\left[\psi(t)S_0\right]\right\},$$

*where the functions $\psi$ and $\phi$ are given by:*

$$\psi(t)=\frac{\left(Q^\top Q\right)^{-1}M}{2}-\frac{Q^{-1}\sqrt{\bar{v}}kQ^{\top^{-1}}}{2}, \qquad (2.4)$$

$$\phi(t)=-\frac{\alpha}{2}\log\left(\det\left(e^{-Mt}\left(\cosh(\sqrt{\bar{v}}t)+\sinh(\sqrt{\bar{v}}t)k\right)\right)\right). \qquad (2.5)$$

REMARK 1.2. *The derivation of Theorem 1.1 involves a change of probability measure that will be illustrated in the sequel. This change of measure introduces a lack of symmetry which does not allow to derive a fully general formula. However, under the assumption* (2.1) *we are able to span a large class of processes. To be more precise, in the two dimensional case, let:*

$$\left(Q^\top Q\right)^{-1}=A=\begin{pmatrix}a&b\\b&c\end{pmatrix},\quad M=\begin{pmatrix}x&y\\z&t\end{pmatrix}, \qquad (2.6)$$

*then condition* (2.1) *can be expressed as:*

$$bx+cz=ay+tb, \qquad (2.7)$$

*meaning that we can span a large class of parameters, thus going far beyond the commutativity assumption $QM=MQ$ for $Q\in S_d^+, M\in S_d^-$ as in Bru (1991).*

**2.2. Proof of Theorem 1.1.** We will prove the theorem in several steps. We first consider a simple Wishart process with $M=0$ and $Q=I_d$, defined under a measure $\tilde{\mathbb{P}}$ equivalent to $\mathbb{P}$. The second step will be given by the introduction of the volatility matrix $Q$, using an invariance result. Finally, we will prove the extension for the full process by relying on a measure change from $\tilde{\mathbb{P}}$ to $\mathbb{P}$. Under this last measure, the Wishart process will be defined by the dynamics (1.1).

As a starting point we fix a probability measure $\tilde{\mathbb{P}}$ such that $\tilde{\mathbb{P}} \approx \mathbb{P}$. Under the measure $\tilde{\mathbb{P}}$ we consider a matrix Brownian motion $\hat{B} = (\hat{B}_t)_{t\geq 0}$, which allows us to define the process $\Sigma_t \in WIS_d(S_0, \alpha, 0, I_d)$, i.e. a process that solves the following matrix SDE:

$$d\Sigma_t = \sqrt{\Sigma_t} d\hat{B}_t + d\hat{B}_t^\top \sqrt{\Sigma_t} + \alpha I_d dt. \tag{2.8}$$

For this process, relying on Pitman and Yor (1982) and Bru (1991), we are able to calculate the Cameron-Martin formula. For the sake of completeness we report the result in Bru (1991), which constitutes an extension of the methodology introduced in Pitman and Yor (1982).

PROPOSITION 1.1. *(Bru (1991) Proposition 5 p.742) If $\mathbf{\Phi} : \mathbb{R}_+ \to S_d^+$ is continuous, constant on $[t, \infty[$ and such that its right derivative (in the distribution sense) $\mathbf{\Phi}'_d : \mathbb{R}_+ \to S_d^-$ is continuous, with $\mathbf{\Phi}_d(0) = I_d$, and $\mathbf{\Phi}'_d(t) = 0$, then for every Wishart process $X_t \in WIS_d(S_0, \alpha, 0, I_d)$ we have:*

$$\mathbb{E}\left[\exp\left\{-\frac{1}{2}Tr\left[\int_0^t \mathbf{\Phi}''_d(s)\mathbf{\Phi}_d^{-1}(s)X_s ds\right]\right\}\right] = (\det \mathbf{\Phi}_d(t))^{\alpha/2} \exp\left\{\frac{1}{2}Tr\left[X_0 \mathbf{\Phi}_d^+(0)\right]\right\},$$

*where*

$$\mathbf{\Phi}_d^+(0) := \lim_{t\searrow 0} \mathbf{\Phi}'_d(t).$$

As a direct application we obtain the following:

PROPOSITION 1.2. *Let $\Sigma \in WIS_d(S_0, \alpha, 0, I_d)$, then*

$$\begin{aligned} \mathbb{E}\left[\exp\left\{-\frac{1}{2}Tr\left[w\Sigma_t + \int_0^t v\Sigma_s ds\right]\right\}\right] &= \det\left(\cosh\left(\sqrt{v}t\right) + \sinh\left(\sqrt{v}t\right)k\right)^{\frac{\alpha}{2}} \\ &\times \exp\left\{\frac{1}{2}Tr\left[\Sigma_0\sqrt{v}k\right]\right\}, \end{aligned} \tag{2.9}$$

*where $k$ is given by:*

$$k = -\left(\sqrt{v}\cosh(\sqrt{v}t) + w\sinh(\sqrt{v}t)\right)^{-1}\left(\sqrt{v}\sinh(\sqrt{v}t) + w\cosh(\sqrt{v}t)\right).$$

PROOF. By Proposition 1.1 we have to solve the ODE:

$$\begin{aligned} &\mathbf{\Phi}''_d(s) = v\mathbf{\Phi}_d(s) \quad s \in (0,t), \\ &\mathbf{\Phi}'^-_d(t) = -w\mathbf{\Phi}_d(t), \\ &\mathbf{\Phi}_d(0) = I_d. \end{aligned}$$

The general solution to this ODE is given by $\mathbf{\Phi}_d(s) = \cosh\left(\sqrt{v}s\right)k_1 + \sinh\left(\sqrt{v}s\right)k$. The condition $\mathbf{\Phi}_d(0) = I_d$ implies $k_1 = I_d$. In order to determine $k$ we look at the boundary condition on $\mathbf{\Phi}'^-_d(t)$ and hence write

$$\sqrt{v}\sinh\left(\sqrt{v}t\right) + \sqrt{v}\cosh\left(\sqrt{v}t\right)k = -w\left(\cosh\left(\sqrt{v}t\right) + \sinh\left(\sqrt{v}t\right)k\right).$$

The value for $k$ easily follows. Equipped with the value for $k$, we can proceed to compute the first derivative of $\mathbf{\Phi}_d$:

$$\mathbf{\Phi}'_d(s) = \sqrt{v}\sinh\left(\sqrt{v}s\right) + \sqrt{v}\cosh\left(\sqrt{v}s\right)k,$$

such that

$$\lim_{s\searrow 0}\mathbf{\Phi}'_d(s) = \sqrt{v}k.$$

By noting that $\mathbf{\Phi}_d(\infty) = \mathbf{\Phi}_d(t)$ since $\mathbf{\Phi}_d$ is constant on $[t, \infty)$, we obtain the claim. □

**Invariance under transformations**. We proceed by introducing a volatility matrix in the Wishart dynamics (2.8) via a transformation. The process is again considered with respect to the equivalent measure $\tilde{\mathbb{P}}$. We define the transformation $S_t = Q^\top \Sigma_t Q$, which is governed by the SDE:

$$dS_t = \sqrt{S_t} d\tilde{B}_t Q + Q^\top d\tilde{B}_t^\top \sqrt{S_t} + \alpha Q^\top Q dt, \tag{2.10}$$

where the process $\tilde{B} = (\tilde{B}_t)_{t\geq 0}$ defined by $d\tilde{B}_t = \sqrt{S_t}^{-1} Q^\top \sqrt{\Sigma} d\hat{B}_t$ is a Brownian motion under $\tilde{\mathbb{P}}$. We proceed to show this in detail.

LEMMA 1.1. *The process* $\tilde{B}_t := \sqrt{S_t}^{-1} Q^\top \sqrt{\Sigma} \hat{B}_t$ *is a matrix Brownian motion.*

PROOF. Recall that it is sufficient to prove that, for any $\alpha, \beta \in \mathbb{R}^d$, the following holds true:

$$Cov(\tilde{B}_t \alpha, \tilde{B}_t \beta) = \mathbb{E}\left[ (d\tilde{B}_t \alpha)(d\tilde{B}_t \beta)^\top \right] = \alpha^\top \beta I_d dt \tag{2.11}$$

In our case, recalling that we defined $S_t = Q^\top \Sigma_t Q$

$$\begin{aligned}
&\mathbb{E}\left[ (\sqrt{S_t}^{-1} Q^\top \sqrt{\Sigma_t} d\hat{B}_t \alpha)(\sqrt{S_t}^{-1} Q^\top \sqrt{\Sigma_t} d\hat{B}_t \beta)^\top \right] \\
&\mathbb{E}\left[ \mathbb{E}\left[ (\sqrt{S_t}^{-1} Q^\top \sqrt{\Sigma_t} d\hat{B}_t \alpha)(\sqrt{S_t}^{-1} Q^\top \sqrt{\Sigma_t} d\hat{B}_t \beta)^\top \Big| \Sigma_t \right] \right] \\
&= \alpha^\top \beta \mathbb{E}\left[ \sqrt{S_t}^{-1} Q^\top \sqrt{\Sigma_t} \sqrt{\Sigma_t} Q \sqrt{S_t}^{-1} \right] dt \\
&= \alpha^\top \beta \mathbb{E}\left[ \sqrt{S_t}^{-1} Q^\top \Sigma_t Q \sqrt{S_t}^{-1} \right] dt \\
&= \alpha^\top \beta \mathbb{E}\left[ \sqrt{S_t}^{-1} S_t \sqrt{S_t}^{-1} \right] dt \\
&= \alpha^\top \beta \mathbb{E}\left[ \sqrt{S_t}^{-1} \sqrt{S_t} \sqrt{S_t} \sqrt{S_t}^{-1} \right] dt \\
&= \alpha^\top \beta I_d dt && (2.12)
\end{aligned}$$

□

From Bru (1991), we know the Laplace transform of this process (which was computed in the simplest case by relying on the associated backward Kolmogorov equation). Upon the introduction of the volatility matrix $Q$ we have:

$$\begin{aligned}
\mathbb{E}_{S_0}^{\tilde{\mathbb{P}}}\left[ e^{-Tr[uS_t]} \right] &= \mathbb{E}_{(Q^\top)^{-1} S_0 Q^{-1}}^{\tilde{\mathbb{P}}}\left[ e^{-Tr\left[uQ^\top \Sigma Q\right]} \right] \\
&= \mathbb{E}_{\Sigma_0}^{\tilde{\mathbb{P}}}\left[ e^{-Tr\left[\left(QuQ^\top\right)\Sigma\right]} \right] \\
&= \left( \det\left( I_d + 2tQuQ^\top \right) \right)^{-\frac{\alpha}{2}} \times \\
&\exp\left\{ -Tr\left[ S_0 Q^{-1} \left( I_d + 2tQuQ^\top \right)^{-1} Qu \right] \right\}.
\end{aligned}$$

Using the Taylor expansion $(I+A)^{-1} = I - A + A^2 - A^3 + ...$, we have

$$\begin{aligned}
Q^{-1}\left( I + 2tQuQ^\top \right)^{-1} Q &= Q^{-1}\left( I_d - 2tQuQ^\top + 4t^2 \left(QuQ^\top\right)\left(QuQ^\top\right) - \ldots \right) Q \\
&= I_d - 2tuQ^\top Q + 4t^2 \left(uQ^\top Q\right)\left(uQ^\top Q\right) - \ldots,
\end{aligned}$$

then, using Sylvester's law of inertia,

$$\det(I_d + AB) = \det(I_d + BA),$$

we obtain

$$\det\left(I_d + 2tQuQ^\top\right) = \det\left(I_d - 2tuQ^\top Q\right).$$

**Inclusion of the drift - Girsanov transformation.** The final step consists in introducing a measure change from $\tilde{\mathbb{P}}$, where the process has no mean reversion, to the measure $\mathbb{P}$ that will allow us to consider the general process governed by the dynamics in equation (1.1). We now define a matrix Brownian motion under the probability measure $\mathbb{P}$ as follows:

$$B_t = \tilde{B}_t - \int_0^t \sqrt{S_s} M^\top Q^{-1} ds = \tilde{B}_t - \int_0^t H_s ds.$$

The Girsanov transformation is given by the following stochastic exponential (see e.g. Donati-Martin et al. (2004)):

$$\begin{aligned} \left.\frac{\partial \mathbb{P}}{\partial \tilde{\mathbb{P}}}\right|_{\mathcal{F}_t} &= \exp\left\{ \int_0^t Tr\left[ H^\top d\tilde{B}_s \right] - \frac{1}{2}\int_0^t Tr\left[ HH^\top \right] ds \right\} \\ &= \exp\left\{ \int_0^t Tr\left[ Q^{-1^\top} M \sqrt{S_s} d\tilde{B}_s \right] - \frac{1}{2}\int_0^t Tr\left[ S_s M^\top Q^{-1} Q^{-1^\top} M \right] ds \right\}. \end{aligned}$$

We concentrate on the stochastic integral term, which may be rewritten as:

$$\int_0^t Tr\left[ \left(Q^\top Q\right)^{-1} M \sqrt{S_s} d\tilde{B}_s Q \right]$$

which, under the parametric restriction (2.1) can be expressed as:

$$\frac{1}{2}\int_0^t Tr\left[ \left(Q^\top Q\right)^{-1} M \left( \sqrt{S_s} d\tilde{B}_s Q + Q^\top d\tilde{B}_s^\top \sqrt{S_s} \right) \right]$$

and then we can write:

$$\frac{1}{2}\int_0^t Tr\left[ \left(Q^\top Q\right)^{-1} M \left( dS_s - \alpha Q^\top Q ds \right) \right].$$

In summary, the stochastic exponential may be written as:

$$\left.\frac{\partial \mathbb{P}}{\partial \tilde{\mathbb{P}}}\right|_{\mathcal{F}_t} = \exp\left\{ \frac{\left(Q^\top Q\right)^{-1} M}{2} \left( S_t - S_0 - \alpha Q^\top Q t \right) - \frac{1}{2}\int_0^t Tr\left[ S_s M^\top Q^{-1} Q^{-1^\top} M \right] ds \right\}.$$

Mayerhofer (2012) shows that under the assumption $\alpha \geq d+1$ (which is a sufficient condition ensuring that the process does not hit the boundary of the cone $S_d^+$) the stochastic exponential is a true martingale.

**Derivation of the Matrix Cameron-Martin formula**. We finally consider the process under $\mathbb{P}$:

$$dS_t = \sqrt{S_t} dB_t Q + Q^\top dB_t^\top \sqrt{S_t} + \left( M S_t + S_t M^\top + \alpha Q^\top Q \right) dt.$$

For the reader's convenience, we would like to summarise the procedure we are adopting. Recall that we know the Cameron Martin formula for the Wishart process when $M = 0$ and $Q = I_d$. We are proceeding in two separate steps in order to treat the general version of the process by considering the invariance under transformation property and a change of probability measure. Recall that under $\tilde{\mathbb{P}}$, we have:

$$dS_t = \sqrt{S_t} d\tilde{B}_t Q + Q^\top d\tilde{B}_t^\top \sqrt{S_t} + \alpha Q^\top Q dt,$$

then $\Sigma_t = Q^{-1^\top} S_t Q^{-1}$ solves:

$$d\Sigma_t = \sqrt{\Sigma_t} d\hat{B}_t + d\hat{B}_t^\top \sqrt{\Sigma_t} + \alpha I_d dt.$$

We are now ready to apply the change of measure along the following steps:

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}_{S_0}\left[\exp\left\{-\frac{1}{2}Tr\left[wS_t+\int_0^t vS_s ds\right]\right\}\right]\\
&=\mathbb{E}^{\tilde{\mathbb{P}}}_{S_0}\left[\exp\left\{-\frac{1}{2}Tr\left[wS_t+\int_0^t vS_s ds\right]-\frac{\alpha}{2}tTr\left[M\right]\right.\right.\\
&-Tr\left[\frac{\left(Q^\top Q\right)^{-1}M}{2}S_0\right]+Tr\left[\frac{\left(Q^\top Q\right)^{-1}M}{2}S_t\right]\\
&\left.\left.-\frac{1}{2}\int_0^t Tr\left[S_sM^\top Q^{-1}Q^{-1^\top}M\right]ds\right\}\right]\\
&=\exp\left\{-\frac{\alpha}{2}tTr\left[M\right]-Tr\left[\frac{\left(Q^\top Q\right)^{-1}M}{2}S_0\right]\right\}\\
&\times\mathbb{E}^{\tilde{\mathbb{P}}}_{S_0}\left[\exp\left\{-\frac{1}{2}Tr\left[\left(w-\left(Q^\top Q\right)^{-1}M\right)S_t\right.\right.\right.\\
&\left.\left.\left.+\int_0^t\left(v+M^\top Q^{-1}Q^{-1^\top}M\right)S_s ds\right]\right\}\right].
\end{aligned}$$

But $S_t = Q^\top \Sigma_t Q$, then:

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}}_{S_0}\left[\exp\left\{-\frac{1}{2}Tr\left[wS_t+\int_0^t vS_s ds\right]\right\}\right]\\
&=\exp\left\{-\frac{\alpha}{2}tTr\left[M\right]-Tr\left[\frac{\left(Q^\top Q\right)^{-1}M}{2}S_0\right]\right\}\\
&\times\mathbb{E}^{\tilde{\mathbb{P}}}_{Q^{\top^{-1}}S_0Q^{-1}}\left[\exp\left\{-\frac{1}{2}Tr\left[Q\left(w-\left(Q^\top Q\right)^{-1}M\right)Q^\top\Sigma_t\right.\right.\right.\\
&\left.\left.\left.+\int_0^t Q\left(v+M^\top Q^{-1}Q^{-1^\top}M\right)Q^\top\Sigma_s ds\right]\right\}\right].
\end{aligned}$$

The expectation may be computed via a direct application of formula (2.9) and after some standard algebra we get the result of Theorem 1.1, with the obvious substitutions $v \to 2v$ and $w \to 2w$.

## 3. Alternative existing methods

**3.1. Variation of Constants Method.** Since the process is affine, it is possible to reduce the PDE associated to the computation of (1.2) to a non linear (matrix Riccati) ODE.

PROPOSITION 1.3. *Let $S_t \in WIS_d(S_0,\alpha,M,Q)$ be the Wishart process defined by (1.1), then*

$$\mathbb{E}^{\mathbb{P}}_{S_0}\left[\exp\left\{-Tr\left[wS_t+\int_0^t vS_s ds\right]\right\}\right]=\exp\left\{-\phi(t)-Tr\left[\psi(t)S_0\right]\right\},$$

*where the functions $\psi$ and $\phi$ satisfy the following system of ODE's.*

$$\frac{d\psi}{dt}=\psi M+M^\top\psi-2\psi Q^\top Q\psi+v \quad \psi(0)=w, \tag{3.1}$$

$$\frac{d\phi}{dt}=Tr\left[\alpha Q^\top Q\psi(t)\right] \quad \phi(0)=0. \tag{3.2}$$

PROOF. See Cuchiero et al. (2009). □

The idea underlying the variation of constants method is that in order to compute the solution of the system of matrix ODE's (3.1), (3.2), it is sufficient to find a particular solution to the equation for $\psi$, since the solution for $\phi$ will be obtained via direct integration. We will therefore proceed in two steps: first, we will solve the equation for $v = 0$ and then provide the most general form. To this aim, we first introduce the following lemma.

LEMMA 1.2. *Let $\psi' \in S_d$ be a symmetric solution to the algebraic Riccati equation:*

$$\psi' M + M^\top \psi' - 2\psi' Q^\top Q \psi' + v = 0, \tag{3.3}$$

*then the function $Z(t) = \psi(t) - \psi'$ solves the following matrix ODE:*

$$\frac{dZ}{dt} = Z(t)M' + M'^\top Z(t) - 2Z(t)Q^\top Q Z(t), \tag{3.4}$$

*with $Z(0) = w'$, $w' = w - \psi'$, $M' = M - 2Q^\top Q\psi'$.*

PROOF. We replace $\psi$ by $Z(t) + \psi'$ in equation (3.1) and obtain:

$$\begin{aligned}\frac{dZ}{dt} &= (Z(t) + \psi')\, M + M^\top (Z(t) + \psi') - 2\,(Z(t) + \psi')\, Q^\top Q\,(Z(t) + \psi') + v \\ &= Z(t)\left(\overbrace{M - 2Q^\top Q\psi'}^{M'}\right) + \left(\overbrace{M^\top - 2\psi' Q^\top Q}^{M'^\top}\right) Z(t) - 2Z(t)Q^\top Q Z(t) \\ &+ \underbrace{\psi' M + M^\top \psi' - 2\psi' Q^\top Q \psi' + v}_{=0}.\end{aligned}$$

From $Z(t) = \psi(t) - \psi'$, we obtain $Z(0) = w - \psi'$. □

The next step is the computation of the solution for $Z(t)$, which is given by the following:

LEMMA 1.3. *The solution to the equation* (3.4) *is given by:*

$$Z(t) = e^{M'^\top t}\left({w'}^{-1} + \int_0^t e^{M's} Q^\top Q e^{M'^\top s} ds\right)^{-1} e^{M't}.$$

PROOF. Consider the function $f(t)$ defined by $Z(t) = e^{M'^\top t} f(t) e^{M't}$. By differentiating this expression we get:

$$\frac{dZ}{dt} = M'^\top Z(t) + Z(t)M' + e^{M'^\top t}\frac{df(t)}{dt}e^{M't}.$$

A comparison with (3.4) gets

$$e^{M'^\top t}\frac{df(t)}{dt}e^{M't} = -2e^{M'^\top t} f(t) e^{M't} Q^\top Q e^{M'^\top t} f(t) e^{M't},$$

which implies

$$\begin{aligned}\frac{df(t)}{dt} &= -2f(t)e^{M't}Q^\top Q e^{M'^\top t} f(t) \quad f(0) = w' \\ -f(t)^{-1}\frac{df(t)}{dt}f(t)^{-1} &= 2e^{M't}Q^\top Q e^{M'^\top t}.\end{aligned}$$

From Faraut and Korànyi (1994) we know that this is equivalent to

$$\frac{df(t)^{-1}}{dt} = 2e^{M't}Q^\top Q e^{M'^\top t} \quad f(0) = {w'}^{-1}.$$

Direct integration of this ODE and substitution of the solution in the identity defining Z yields the desiderd result. □

If we combine the two lemmas above, we obtain the solution for the functions $\psi$ and $\phi$. This is stated in the following proposition.

PROPOSITION 1.4. *The solutions for $\psi(t), \phi(t)$ in Proposition 1.3 are given by:*

$$\begin{aligned}\psi(t) = \psi' + e^{\left(M^\top - 2\psi' Q^\top Q\right)t}\Bigg[&(w-\psi')^{-1}\\ &+ 2\int_0^t e^{\left(M-2Q^\top Q\psi'\right)s} Q^\top Q e^{\left(M^\top - 2\psi' Q^\top Q\right)s} ds\Bigg]^{-1} e^{\left(M-2Q^\top Q\psi'\right)t},\\ \phi(t) = Tr\left[\alpha Q^\top Q \int_0^t \psi(s)ds\right],&\end{aligned}$$

*where $\psi'$ is a symmetric solution to the algebraic Riccati equation (3.3).*

REMARK 1.3. *In this subsection we followed closely Gourieroux and Sufana (2005), who solved the Riccati ODE (3.1) by using the variation of constants method (see also Gourieroux et al. (2005), Gourieroux and Sufana (2003), Gourieroux and Sufana (2005)). This is also equivalent to the procedure followed by Ahdida and Alfonsi (2010) and Mayerhofer (2010) who found the Laplace transform of the Wishart process alone (i.e. corresponding to $v = 0$ in (1.2)). The variation of constants method represents the first solution provided in literature for the solution of the matrix ODE's (3.1) and (3.2), and despite its theoretical simplicity, it turns out to be very time consuming, as we will show later in the numerical exercise.*

**3.2. Linearization of the Matrix Riccati ODE.** The second approach we consider is the one proposed by Grasselli and Tebaldi (2008), who used the Radon lemma in order to linearize the matrix Riccati ODE (3.1) (see also Levin (1959), Jong and Zhou (1999) and Anderson and Moore (1971)). As usual we are interested in the computation of the moment generating function of the process and of the integrated process, hence we look at the system of equations (3.1) and (3.2).

PROPOSITION 1.5. *The functions $\psi(t), \phi(t)$ in Proposition 1.3 are given by*

$$\begin{aligned}\psi(t) &= \left(w\psi_{12}(t) + \psi_{22}(t)\right)^{-1}\left(w\psi_{11}(t) + \psi_{21}(t)\right),\\ \phi(t) &= \frac{\alpha}{2} Tr\left[log\left(w\psi_{12}(t) + \psi_{22}(t)\right) + M^\top t\right],\end{aligned}$$

*where*

$$\begin{pmatrix} \psi_{11}(t) & \psi_{12}(t) \\ \psi_{21}(t) & \psi_{22}(t) \end{pmatrix} = exp\left\{t\begin{pmatrix} M & 2Q^\top Q \\ v & -M^\top \end{pmatrix}\right\}.$$

PROOF. We begin by writing

$$\psi(t) = F(t)^{-1}G(t), \tag{3.5}$$

for $F(t) \in GL(d)$ and $G(t) \in M_d$, the set of $d \times d$ matrices. Then we can write the time derivative as

$$\frac{d}{dt}\left[F(t)\psi(t)\right] - \frac{d}{dt}\left[F(t)\right]\psi(t) = F(t)\frac{d}{dt}\left[\psi(t)\right]$$

and this implies that

$$\begin{aligned}\frac{d}{dt}\left[F(t)\psi(t)\right] - \frac{d}{dt}\left[F(t)\right]\psi(t) =& F(t)\psi(t)M + F(t)M^\top\psi(t)\\ &- 2F(t)\psi(t)Q^\top Q\psi(t) + F(t)v.\end{aligned} \tag{3.6}$$

From (3.5) we immediately obtain $G(t) = F(t)\psi(t)$. Using this fact we can obtain a system of ODE's for $F$ and $G$ by recognizing the remaining terms in $\psi(t)$ in Equation (3.6).

$$\frac{d}{dt}G(t) = G(t)M + F(t)v, \quad G(0) = w, \tag{3.7}$$

$$\frac{d}{dt}F(t) = -F(t)M^\top + 2G(t)Q^\top Q, \quad F(0) = I_d. \tag{3.8}$$

which is solved by

$$(G(t), F(t)) = (w, I_d)\begin{pmatrix} \psi_{11}(t) & \psi_{12}(t) \\ \psi_{21}(t) & \psi_{22}(t) \end{pmatrix},$$

where

$$\begin{pmatrix} \psi_{11}(t) & \psi_{12}(t) \\ \psi_{21}(t) & \psi_{22}(t) \end{pmatrix} = exp\left\{ t\begin{pmatrix} M & 2Q^\top Q \\ v & -M^\top \end{pmatrix}\right\}.$$

Consequently, the solution for $\psi(t)$ is given by:

$$\psi(t) = \left(w\psi_{12}(t) + \psi_{22}(t)\right)^{-1}\left(w\psi_{11}(t) + \psi_{21}(t)\right), \tag{3.9}$$

since we have $\psi(0) = w$. As usual, a direct integration of (3.2) allows us to derive the solution for $\phi(t)$:

$$\phi(t) = \int_0^t Tr\left[\alpha Q^\top Q\psi(s)\right] ds. \tag{3.10}$$

It turns out that this integral is quite cumbersome, hence we adopt the strategy in Da Fonseca et al. (2008), which allows us to avoid this numerical integration. We notice that equation (3.8) may be written as

$$\frac{1}{2}\left(\frac{d}{dt}F(t) + F(t)M^\top\right)\left(Q^\top Q\right)^{-1} = G(t),$$

then, from (3.5), we have:

$$\psi(t) = \frac{1}{2}\left(F(t)^{-1}\frac{d}{dt}F(t) + M^\top\right)\left(Q^\top Q\right)^{-1}.$$

We substitute this last expression in (3.2) and use the properties of the trace operator, so that we can write:

$$\frac{d\phi(t)}{dt} = \frac{\alpha}{2}Tr\left[F(t)^{-1}\frac{d}{dt}F(t) + M^\top\right],$$

which is solved by

$$\phi(t) = \frac{\alpha}{2}Tr\left[log\left(F(t)\right) + M^\top t\right]. \tag{3.11}$$

Equations (3.9) and (3.11) represent the solution to the system of ODE we were looking for, hence we are able to compute the joint Laplace transform in Proposition 1.3. □

**3.3. Runge-Kutta Method.** The Runge-Kutta method is a classical approach for the numerical solution of ODE's. For a detailed treatment, see e.g. Quarteroni et al. (2000). If we want to solve numerically the system of equations (3.1) and (3.2), the most commonly used Runge-Kutta scheme is the fourth order one:

$$
\begin{aligned}
\psi(t_{n+1}) &= \psi(t_n) + \frac{1}{6}h\left(k_1 + 2k_2 + 2k_3 + k_4\right),\\
t_{n+1} &= t_n + h,\\
k_1 &= g(t_n, \psi(t_n)),\\
k_2 &= g(t_n + \frac{1}{2}h, \psi(t_n) + \frac{1}{2}hk_1),\\
k_3 &= g(t_n + \frac{1}{2}h, \psi(t_n) + \frac{1}{2}hk_2),\\
k_4 &= g(t_n + h, \psi(t_n) + hk_3),
\end{aligned}
$$

where the function $g$ is given by:

$$
g(t_n, \psi(t_n)) = g(\psi(t_n)) = \psi(t_n)M + M^\top \psi(t_n) - 2\psi(t_n)Q^\top Q\psi(t_n) + v.
$$

**3.4. Comparison of the methods.** A formal numerical analysis of the various methods is beyond the scope of this paper. Anyhow, we would like to stress some important points, which we believe are sufficient to highlight the importance of our new methodology. First of all we compare the results of the four different methods. We consider different time horizons $t \in [0, 0.3]$ and use the following values for the parameters:

$$
S_0 = \begin{pmatrix} 0.0120 & 0.0010 \\ 0.0010 & 0.0030 \end{pmatrix}; \quad Q = \begin{pmatrix} 0.141421356237310 & -0.070710678118655 \\ 0 & 0.070710678118655 \end{pmatrix};
$$

$$
M = \begin{pmatrix} -0.02 & -0.02 \\ -0.01 & -0.02 \end{pmatrix}; \quad \alpha = 3;
$$

$$
v = \begin{pmatrix} 0.1000 & 0.0400 \\ 0.0400 & 0.1000 \end{pmatrix}; \quad w = \begin{pmatrix} 0.1100 & 0.0300 \\ 0.0300 & 0.1100 \end{pmatrix}.
$$

The value for $Q$ was obtained along the following steps: given a matrix $A \in S_d^+$ such that $AM = M^\top A$, we compute its inverse and let $Q$ be obtained from a Cholesky factorization of this inverted matrix.
Table (1) shows the value of the moment generating function for different values of the time horizon $t$. The four methods lead to values which are very close to each other, and this constitutes a first test proving that the new methodology produces correct results.

The next important point that we should consider is the execution speed. In order to obtain a good degree of precision for the variation of constants method, we were forced to employ a fine integration grid. This results in a poor performance of this method in terms of speed. In Figure (3) we compare the time spent by the three analytical methods for the calculation of the moment generating function. As $t$ gets larger, the execution time for the variation of constants method grows exponentially, whereas the time required by the linearization and the new methodology is the same.

The Runge-Kutta method is a numerical solution to the problem, so the real competitors of our methodology are the variation of constants and the linearization method. As we saw above, the variation of constant method is quite cumbersome. This is because in order to implement the variation of constants methods we have to solve numerically an algebraic Riccati equation, then we have to perform a first numerical integration in order to determine $\psi$ and another numerical integration to compute $\phi$. The above procedure is obviously time consuming, hence we believe that this method is not suitable for applications, in particular in a calibration setting.

Finally, we would like to compare the linearization of the Riccati ODE to the new methodology. In terms of precision and execution speed the two methodologies seem to provide the same performance, up to the fourteenth digit. This shows that, under the parametric restriction of Theorem 1.1 our methodology represents a valid alternative.

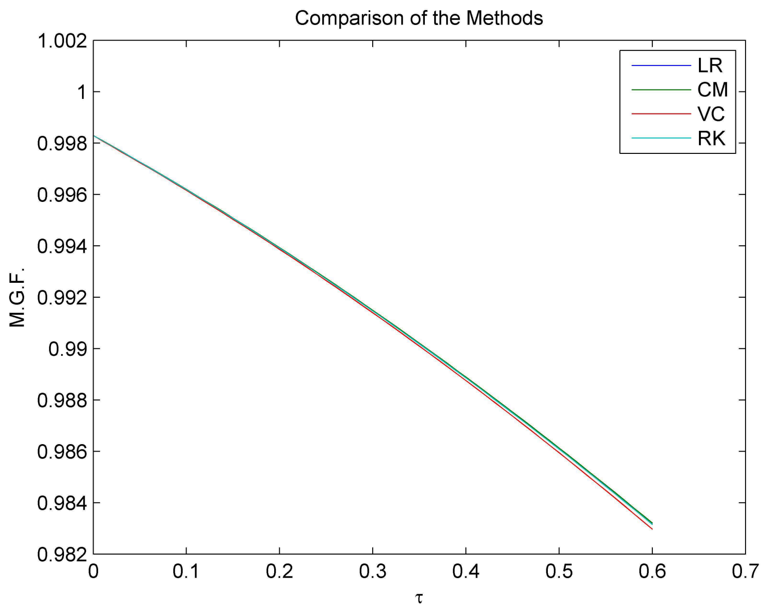



FIGURE 1. In this image we plot the value of the joint moment generating function of the Wishart process and its time integral for different time horizons $\tau$. All four methods are considered. It should be noted that the variation of constants method requires a very fine integration grid in order to produce precise values that can be compared with the results of the other methods.

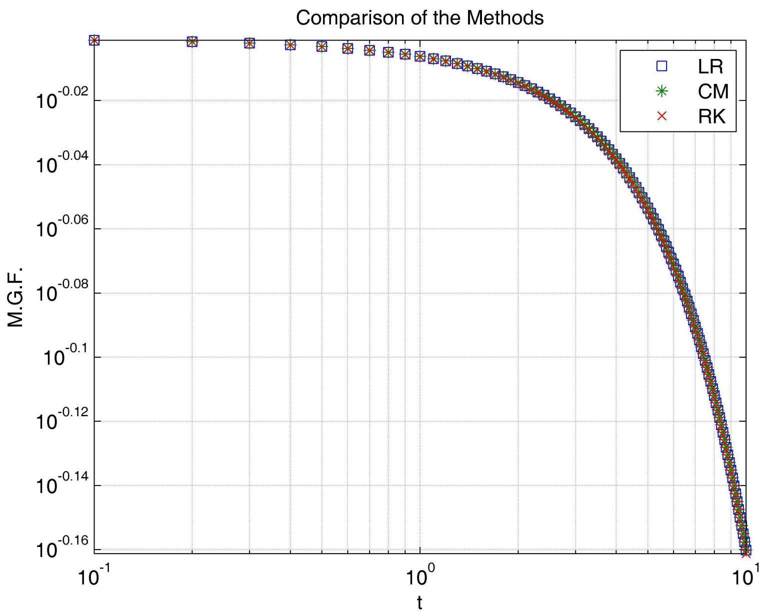



FIGURE 2. In this case we exclude the variation of constants method and treat a larger time horizon.

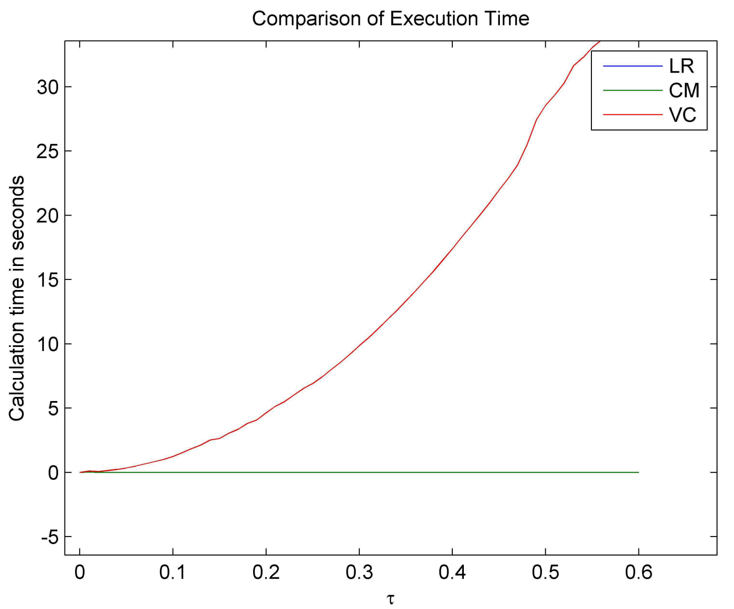



FIGURE 3. In this image we plot the time spent by the three analytical methods to compute the joint moment generating function of the Wishart process and its time integral for different time horizons.

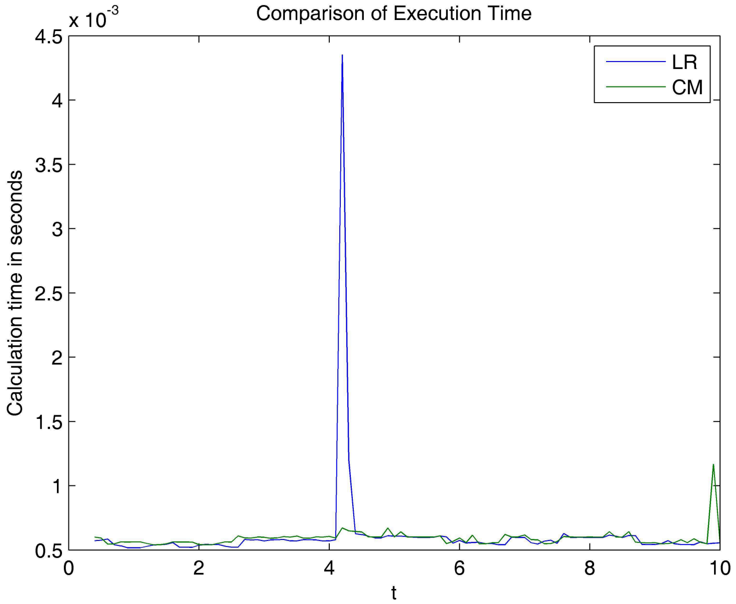



FIGURE 4. We restrict the previous comparison of execution times to the linearization method and the Cameron-Martin formula.

| Time Horizon | Lin. | C.-M. | Var. Const. | R.-K. |
|---|---|---|---|---|
| 0 | 0.998291461216988 | 0.998291461216988 | 0.998291461216988 | 0.998291461216988 |
| 0.1 | 0.997303305375919 | 0.997303305375919 | 0.997306285702955 | 0.997271605593416 |
| 0.2 | 0.996253721242885 | 0.996253721242885 | 0.996258979717961 | 0.996190498718109 |
| 0.3 | 0.995143124879428 | 0.995143124879428 | 0.995142369361912 | 0.995048563313279 |
| 0.4 | 0.993971944944528 | 0.993971944944528 | 0.993956917727425 | 0.993846234580745 |
| 0.5 | 0.992740622447456 | 0.992740622447456 | 0.992703104707601 | 0.992583959952442 |
| 0.6 | 0.991449610496379 | 0.991449610496380 | 0.991381426685627 | 0.991262198836806 |
| 0.7 | 0.990099374042951 | 0.990099374042951 | 0.989992396217013 | 0.989881422361235 |
| 0.8 | 0.988690389623073 | 0.988690389623073 | 0.988536541704748 | 0.988442113110838 |
| 0.9 | 0.987223145094070 | 0.987223145094070 | 0.987014407067708 | 0.986944764863693 |
| 1.0 | 0.985698139368470 | 0.985698139368470 | 0.985426551402640 | 0.985389882322825 |
| 1.1 | 0.984115882144609 | 0.984115882144608 | 0.983773548640023 | 0.983777980845113 |
| 1.2 | 0.982476893634278 | 0.982476893634278 | 0.982055987194167 | 0.982109586167352 |
| 1.3 | 0.980781704287638 | 0.980781704287638 | 0.980274469607849 | 0.980385234129674 |
| 1.4 | 0.979030854515581 | 0.979030854515582 | 0.978429612191836 | 0.978605470396549 |
| 1.5 | 0.977224894409802 | 0.977224894409802 | 0.976522044659620 | 0.976770850175581 |
| 1.6 | 0.975364383460752 | 0.975364383460752 | 0.974552409757698 | 0.974881937934301 |
| 1.7 | 0.973449890273708 | 0.973449890273707 | 0.972521362891742 | 0.972939307115188 |
| 1.8 | 0.971481992283166 | 0.971481992283166 | 0.970429571748981 | 0.970943539849112 |
| 1.9 | 0.969461275465768 | 0.969461275465768 | 0.968277715917132 | 0.968895226667421 |
| 2.0 | 0.967388334051965 | 0.967388334051964 | 0.966066486500228 | 0.966794966212865 |
| 2.1 | 0.965263770236630 | 0.965263770236631 | 0.963796585731653 | 0.964643364949586 |
| 2.2 | 0.963088193888842 | 0.963088193888842 | 0.961468726584740 | 0.962441036872358 |
| 2.3 | 0.960862222260992 | 0.960862222260992 | 0.959083632381238 | 0.960188603215284 |
| 2.4 | 0.958586479697485 | 0.958586479697484 | 0.956642036397998 | 0.957886692160160 |
| 2.5 | 0.956261597343174 | 0.956261597343174 | 0.954144681472186 | 0.955535938544691 |
| 2.6 | 0.953888212851758 | 0.953888212851759 | 0.951592319605355 | 0.953136983570760 |
| 2.7 | 0.951466970094322 | 0.951466970094322 | 0.948985711566685 | 0.950690474512938 |
| 2.8 | 0.948998518868209 | 0.948998518868209 | 0.946325626495722 | 0.948197064427429 |
| 2.9 | 0.946483514606424 | 0.946483514606425 | 0.943612841504911 | 0.945657411861629 |
| 3.0 | 0.943922618087738 | 0.943922618087738 | 0.940848141282233 | 0.943072180564490 |

TABLE 1. This table visualizes in more detail the numerical values for the joint moment generating function plotted in Figure (1).

| Time horizon | Lin. | C.M. | R.-K. |
|---|---|---|---|
| 0 | 0.998291461216988 | 0.998291461216988 | 0.998291461216988 |
| 0.1 | 0.997303305375919 | 0.997303305375919 | 0.997271605593416 |
| 0.2 | 0.996253721242885 | 0.996253721242885 | 0.996190498718109 |
| 0.3 | 0.995143124879428 | 0.995143124879428 | 0.995048563313279 |
| 0.4 | 0.993971944944528 | 0.993971944944528 | 0.993846234580745 |
| 0.5 | 0.992740622447456 | 0.992740622447456 | 0.992583959952442 |
| 1.0 | 0.985698139368470 | 0.985698139368470 | 0.985389882322825 |
| 2.0 | 0.967388334051965 | 0.967388334051964 | 0.966794966212865 |
| 3.0 | 0.943922618087738 | 0.943922618087738 | 0.943072180564490 |
| 4.0 | 0.915938197508059 | 0.915938197508059 | 0.914862207389661 |
| 5.0 | 0.884120166104796 | 0.884120166104796 | 0.882852196560219 |
| 10.0 | 0.691634000576684 | 0.691634000576684 | 0.689897813632122 |
| 100.0 | 0.000001636282753 | 0.000001636282753 | 0.000001629036716 |

TABLE 2. In this table we do not include the results for the variation of constants method. This allows us to look at a longer time horizon and appreciate the precision of the new methodology also in this case.

## 4. Applications

**4.1. A stochastic volatility model.** In this subsection we consider the model proposed in Da Fonseca et al. (2008) and show that it is possible to derive the explicit Laplace transform of the log-forward-price using our new methodology. As a starting point, we report the dynamics defining the model:

$$\frac{dF_t}{F_t} = Tr\left[\sqrt{S_t}\left(dW_t R^\top + dB_t\sqrt{I_d - RR^\top}\right)\right],$$
$$dS_t = \left(\alpha Q^\top Q + MS_t + S_t M^\top\right)dt + \sqrt{S_t}dW_t Q + Q^\top dW_t^\top\sqrt{S_t},$$

where $F_t$ denotes the forward-price of the underlying asset, and the Wishart process acts as a multifactor source of stochastic volatility. $W$ and $B$ are independent matrix Brownian motions and the matrix $R$ parametrizes all possible correlation structures preserving the affinity. This model is a generalization of the (multi-)Heston model, see Heston (1993) and Christoffersen et al. (2009), and it offers a very rich structure for the modelization of stochastic volatilities as the factors governing the instantaneous variance are non-trivially correlated. It is easy to see that the log-forward-price $Y$ is given as

$$dY = -\frac{1}{2}Tr\left[S_t\right]dt + Tr\left[\sqrt{S_t}\left(dW_t R^\top + dB_t\sqrt{I_d - RR^\top}\right)\right].$$

We are interested in the Laplace transform of the log-price, i.e.:

$$\varphi(\tau) = \mathbb{E}\left[e^{-\omega Y_T}\left|\mathcal{F}_t\right.\right], \quad \tau := T - t.$$

This expectation satisfies a backward Kolmogorov equation, see Da Fonseca et al. (2008) for a detailed derivation. Since the process $S = (S_t)_{0\le t\le T}$ is affine, we make a guess of a solution of the form

$$\varphi(\tau) = \exp\left\{-\omega\ln F_t - \phi(\tau) - Tr\left[\psi(\tau)S_t\right]\right\}.$$

By substituting it into the PDE, we obtain the system of ODE's:

$$\frac{d\psi}{d\tau} = \psi\left(M - \omega Q^\top R^\top\right) + \left(M^\top - \omega RQ\right)\psi - 2\psi Q^\top Q\psi - \frac{\omega^2+\omega}{2}I_d, \tag{4.1}$$

$$\psi(0) = 0, \tag{4.2}$$

$$\frac{d\phi}{d\tau} = Tr\left[\alpha Q^\top Q\psi(\tau)\right], \tag{4.3}$$

$$\phi(0) = 0. \tag{4.4}$$

If we look at the first ODE, we recognize the same structure as in (3.1): instead of $M$ and $v$ we have respectively $M - \omega Q^\top R^\top$ and $-\frac{\omega^2+\omega}{2}I_d$. This means that we can rewrite the solution for $\psi$, using Remark 1.1, as:

$$\begin{aligned}
\psi(\tau) &= \frac{\left(Q^\top Q\right)^{-1}\left(M - \omega Q^\top R^\top\right)}{2} \\
&\quad - \frac{Q^{-1}\sqrt{\bar{v}}kQ^{\top^{-1}}}{2}, \\
\phi(\tau) &= -\frac{\alpha}{2}\log\left(\det\left(e^{\left(M-\omega Q^\top R^\top\right)\tau}\left(\cosh(\sqrt{\bar{v}}\tau) + \sinh(\sqrt{\bar{v}}\tau)k\right)\right)\right), \\
\bar{v} &= Q\left(2\left(-\frac{\omega^2+\omega}{2}I_d\right) + \left(M^\top - \omega RQ\right)Q^{-1}Q^{\top^{-1}}\left(M - \omega Q^\top R^\top\right)\right)Q^\top, \\
\bar{w} &= Q\left(-\left(Q^\top Q\right)^{-1}\left(M - \omega Q^\top R^\top\right)\right)Q^\top, \\
k &= -\left(\sqrt{\bar{v}}\cosh(\sqrt{\bar{v}}\tau) + \bar{w}\sinh(\sqrt{\bar{v}}\tau)\right)^{-1}\left(\sqrt{\bar{v}}\sinh(\sqrt{\bar{v}}\tau) + \bar{w}\cosh(\sqrt{\bar{v}}\tau)\right).
\end{aligned}$$

Once the Laplace transform of the log-forward price is computed, one can easily perform the Fast Fourier transform in order to price options.

**4.2. A stochastic correlation model.** In this subsection we consider the model introduced in Da Fonseca et al. (2007b). In this framework we consider a vector of forward prices together with a stochastic variance-covariance matrix:

$$\begin{aligned}
dF_t &= Diag(F_t)\sqrt{S_t}\left(dW_t\rho + \sqrt{1-\rho^\top\rho}\,dB_t\right), \\
dS_t &= \left(\alpha Q^\top Q + MS_t + S_tM^\top\right)dt + \sqrt{S_t}dW_tQ + Q^\top dW_t^\top\sqrt{S_t},
\end{aligned}$$

where now the vector Brownian motion $Z = W_t\rho + \sqrt{1-\rho^\top\rho}B_t$ is correlated with the matrix Brownian motion $W$ through the correlation vector $\rho$. Using exactly the same arguments as before, we are able to compute the joint conditional Laplace transform of the vector of the log-forward prices $Y_T = \log(F_T)$:

$$\varphi(\tau) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\omega^\top Y_T}|\mathcal{F}_t\right], \quad \tau := T - t.$$

As in the previous model, we have that the expectation corresponds to a PDE, see Da Fonseca et al. (2007b) for more details. Again the affine property allows us to write the associated system of matrix

Riccati ODE's, which is given as:

$$\begin{aligned}\frac{d\psi}{d\tau} &= \psi\left(M - Q^\top\rho\omega^\top\right) + \left(M^\top - \omega\rho^\top Q\right)\psi - 2\psi Q^\top Q\psi \\ &\quad - \frac{1}{2}\left(\sum_{i=1}^{d}\omega_i e_{ii} + \omega^\top\omega\right) I_d, \end{aligned} \tag{4.5}$$

$$\psi(0) = 0, \tag{4.6}$$

$$\frac{d\phi}{d\tau} = Tr\left[\alpha Q^\top Q\psi(\tau)\right], \tag{4.7}$$

$$\phi(0) = 0. \tag{4.8}$$

We recognize again the same structure as in Equations (3.2) and (3.1) where instead of $M$ and $v$, we now have $M - Q^\top\rho\omega^\top$ and $-\frac{1}{2}\left(\sum_{i=1}^{d}\omega_i e_{ii} + \omega^\top\omega\right) I_d$ respectively. Consequently, using Remark 1.1, we can compute the solution as:

$$\begin{aligned}\psi(\tau) &= \frac{\left(Q^\top Q\right)^{-1}\left(M - Q^\top\rho\omega^\top\right)}{2} \\ &\quad - \frac{Q^{-1}\sqrt{\bar{v}}kQ^{\top^{-1}}}{2}, \\ \phi(\tau) &= -\frac{\alpha}{2}\log\left(\det\left(e^{\left(M-Q^\top\rho\omega^\top\right)\tau}\left(\cosh(\sqrt{\bar{v}}\tau) + \sinh(\sqrt{\bar{v}}\tau)k\right)\right)\right), \\ \bar{v} &= Q\left(2\left(-\frac{1}{2}\left(\sum_{i=1}^{d}\omega_i e_{ii} + \omega^\top\omega\right) I_d\right) + \left(M^\top - \omega\rho Q\right) Q^{-1}Q^{\top^{-1}}\left(M - Q^\top\rho\omega^\top\right)\right) Q^\top, \\ \bar{w} &= Q\left(-\left(Q^\top Q\right)^{-1}\left(M - Q^\top\rho\omega^\top\right)\right) Q^\top, \\ k &= -\left(\sqrt{\bar{v}}\cosh(\sqrt{\bar{v}}\tau) + \bar{w}\sinh(\sqrt{\bar{v}}\tau)\right)^{-1}\left(\sqrt{\bar{v}}\sinh(\sqrt{\bar{v}}\tau) + \bar{w}\cosh(\sqrt{\bar{v}}\tau)\right).\end{aligned}$$

## 5. Conclusions

In this chapter we derived a new explicit formula for the joint Laplace transform of the Wishart process and its time integral based on the original approach of Bru (1991). Our methodology leads to a truly explicit formula that does not involve any additional integration (like the highly time consuming variation of constants method) or blocks of matrix exponentials (like the linearization method) at the price of a simple condition on the parameters. We showed some examples of applications in the context of multivariate stochastic volatility. There is ample room of future research in different domains, including portfolio management, interest rates and option pricing. In fact, we believe that our result can be useful for speeding up some numerical procedure and mostly for the computation of sensitivities in option pricing.

## 6. A generalization of the Wishart process

Affine processes on positive semidefinite matrices have been classified in full generality by Cuchiero et al. (2009). In this reference, a complete set of sufficient and necessary conditions providing a full characterization of this family of processes is derived. These conditions are the $S_d^+$ analogue of the concept of admissibility for the state space $\mathbb{R}^n_{\geq 0} \times \mathbb{R}^m$, which has been studied by Duffie et al. (2003) under the assumption of regularity of the process, and by Keller-Ressel (2008), who proved that regularity is a consequence of the stochastic continuity requirement in the definition of affine process. Cuchiero et al. (2009) showed that an admissible drift can be considerably different from the one considered in (1.1).

In particular, the Gindikin condition $\alpha > d-1$ can be generalized, thus leading to a generalized Wishart process with dynamics[1] with respect to $\mathbb{P}$:

$$dS_t = \sqrt{S_t}dB_tQ + Q^\top dB_t^\top \sqrt{S_t} + \left(MS_t + S_tM^\top + b\right)dt, \quad S_0 = s_0 \in S_d^+, \tag{6.1}$$

where the symmetric matrix $b$ satisfies

$$b - (d-1)Q^\top Q \in S_d^+. \tag{6.2}$$

We will denote by $WIS_d(s_0, b, M, Q)$ the law of the Wishart process $(S_t)_{t\geq 0}$ that satisfies (6.1).

In this Appendix we will find the explicit Cameron Martin formula for the more general specification (6.1). We will first characterize the distribution function of the process $WIS_d(s_0, b, 0, I_d)$ at a fixed time through the Laplace transform and then we will proceed with the analogue of Theorem 1.1.

**6.1. Laplace transform of $WIS_d(\Sigma_0, \tilde{b}, 0, I_d)$.** We fix a probability measure $\tilde{\mathbb{P}}$ such that $\tilde{\mathbb{P}} \approx \mathbb{P}$. Under the measure $\tilde{\mathbb{P}}$ we consider a matrix Brownian motion $\hat{B} = (\hat{B}_t)_{t\geq 0}$ that will allow us to define the process $\Sigma_t$ having law $WIS_d(\Sigma_0, \tilde{b}, 0, I_d), \tilde{b} \in S_d^+$, that is a process which solves the following SDE:

$$d\Sigma_t = \sqrt{\Sigma_t}d\hat{B}_t + d\hat{B}_t^\top \sqrt{\Sigma_t} + \tilde{b}dt, \quad \Sigma_0 \in S_d^+,$$

where the drift term $\tilde{b}$ satisfies the following condition:

$$\tilde{b} - (d-1)\,I_d \in S_d^+.$$

We may characterize the distribution of this process by means of its Laplace Transform.

THEOREM 1.2. *Let $\Sigma$ be a generalized Wishart process in $WIS_d(\Sigma_0, \tilde{b}, 0, I_d)$, then the distribution of $\Sigma_t$, for fixed t, under $\tilde{\mathbb{P}}$, is given by its Laplace transform:*

$$\mathbb{E}_{\Sigma_0}^{\tilde{\mathbb{P}}}\left[e^{-Tr[u\Sigma_t]}\right] = \det\left(e^{\tilde{b}\log(I_d+2tu)^{-\frac{1}{2}}}\right)e^{-Tr\left[(I_d+2tu)^{-1}u\Sigma_0\right]},$$

*for all $u \in S_d$ such that $(I_d + 2tu)$ is nonsingular.*

PROOF. We know that this process belongs to the class of affine processes on $S_d^+$, which means that we may write the Laplace Transform in the following way:

$$\varphi(t, \Sigma_0) = \mathbb{E}_{\Sigma_0}^{\tilde{\mathbb{P}}}\left[e^{-Tr[u\Sigma_t]}\right] = \exp\left\{-\phi(t) - Tr\left[\psi(t)\Sigma_0\right]\right\}.$$

We proceed along the lines of Bru (1991) and look for the solution for the Laplace transform by means of the associated backward Kolmogorov equation.

$$\frac{d}{dt}\varphi(t, \Sigma_0) = Tr\left[bD\varphi(t, \Sigma_0) + 2D\varphi(t, \Sigma_0)^2\right], \quad \varphi(0, \Sigma_0) = e^{-Tr[u\Sigma_0]},$$

where $D$ denotes the matrix differential operator whose $ij$-th element is given by

$$\frac{\partial}{\partial \Sigma_{ij}}.$$

[1] Cuchiero et al. (2009) consider a matrix jump diffusion dynamics with a more general structure for the drift parameter involving a linear operator that cannot be written a priori in the matrix form $MS_t + S_tM^\top$. In this appendix we restrict ourselves to the continuous path version with the usual matrix drift. We also emphasize that in view of applications the specification (1.1) is highly preferable since it is more parsimonius in terms of parameters: this is a delicate and crucial issue when calibrating any model.

Upon substitution of the exponentially affine guess we obtain the following system of ODE's, which is a simplified version of (3.1), (3.2):

$$\frac{d\psi}{dt} = -2\psi(t)^2, \quad \psi(0) = u,$$
$$\frac{d\phi}{dt} = Tr\left[\tilde{b}\psi(t)\right], \quad \psi(0) = 0.$$

The solution for $\psi(t)$ is:

$$\psi(t) = (I_d + 2tu)^{-1} u,$$

while the solution for $\phi(t)$ is obtained via direct integration:

$$\phi(t) = Tr\left[\tilde{b}\int_0^t (I_d + 2su)^{-1} uds\right]$$
$$= Tr\left[\tilde{b}\frac{1}{2}\log(I_d + 2tu)\right].$$

Noting that

$$e^{-\phi(t)} = \det\left(e^{\tilde{b}\log(I_d+2tu)^{-\frac{1}{2}}}\right),$$

we obtain the result. □

**6.2. Cameron Martin formula for the process $WIS_d(s_0, b, M, Q)$.** We now consider, under the measure $\mathbb{P}$, the process governed by the SDE

$$dS_t = \sqrt{S_t}dB_tQ + Q^\top dB_t^\top\sqrt{S_t} + \left(MS_t + S_tM^\top + b\right)dt, \quad S_0 = s_0 \in S_d^+,$$

with $b \in S_d^+$ satisfying (6.2).
In this subsection we compute the joint moment generating function of the process and its time integral.

THEOREM 1.3. *The joint Laplace transform of the generalized Wishart process $S \in WIS_d(s_0, b, M, Q)$ and its time integral is given by:*

$$\mathbb{E}_{s_0}^{\mathbb{P}}\left[\exp\left\{-Tr\left[wS_t + \int_0^t vS_sds\right]\right\}\right] = \exp\left\{-\phi(t) - Tr\left[\psi(t)s_0\right]\right\},$$

*where the functions $\phi$ and $\psi$ are given by:*

$$\psi(t) = \frac{\left(Q^\top Q\right)^{-1}M}{2} - \frac{Q^{-1}\sqrt{\bar{v}}kQ^{\top^{-1}}}{2},$$
$$\phi(t) = Tr\left[b\frac{\left(Q^\top Q\right)^{-1}M}{2}\right]t$$
$$+\frac{1}{2}Tr\left[\left(Q^\top\right)^{-1}b\left(Q\right)^{-1}\log\left(\sqrt{\bar{v}}^{-1}\left(\sqrt{\bar{v}}\cosh(\sqrt{\bar{v}}t) + \bar{w}\sinh(\sqrt{\bar{v}}t)\right)\right)\right],$$

*with $k$ given by:*

$$k = -\left(\sqrt{\bar{v}}\cosh(\sqrt{\bar{v}}t) + \bar{w}\sinh(\sqrt{\bar{v}}t)\right)^{-1}\left(\sqrt{\bar{v}}\sinh(\sqrt{\bar{v}}t) + \bar{w}\cosh(\sqrt{\bar{v}}t)\right)$$

*and $\bar{v}, \bar{w}$ are defined as follows:*

$$\bar{v} = Q\left(2v + M^\top Q^{-1}Q^{\top^{-1}}M\right)Q^\top,$$
$$\bar{w} = Q\left(2w - \left(Q^\top Q\right)^{-1}M\right)Q^\top.$$

*The Laplace transform is defined for all $w, v \in S_d$ such that the matrix $\left(\sqrt{\bar{v}}\cosh(\sqrt{\bar{v}}t) + \bar{w}\sinh(\sqrt{\bar{v}}t)\right)$ is non singular.*

PROOF. The proof will be based on the previous discussion in Theorem 1.1 on the standard process with a scalar Gindikin parameter. First of all we have that the invariance under transformation may still be used.
We consider the process $\Sigma_t \in WIS_d(\Sigma_0, \left(Q^\top\right)^{-1} bQ^{-1}, 0, I_d)$, i.e. a process solving the following matrix SDE:

$$d\Sigma_t = \sqrt{\Sigma_t}d\hat{B}_t + d\hat{B}_t^\top\sqrt{\Sigma_t} + \left(Q^\top\right)^{-1} bQ^{-1}dt.$$

We define the following quantities: $S_t = Q^\top \Sigma_t Q$, $\Sigma_t = \left(Q^\top\right)^{-1} S_t Q^{-1}$. Under this transformation the process under $\tilde{\mathbb{P}}$ is governed by the SDE:

$$dS_t = \sqrt{S_t}d\tilde{B}_t Q + Q^\top d\tilde{B}_t^\top \sqrt{S_t} + bdt,$$

where the process $\tilde{B} = (\tilde{B}_t)_{t\geq 0}$ defined by $\sqrt{S_t}^{-1} Q^\top \sqrt{\Sigma_t} d\hat{B}_t$ is a matrix Brownian motion under $\tilde{\mathbb{P}}$. From Bru (1991) we have:

$$\begin{aligned}\mathbb{E}_{s_0}^{\tilde{\mathbb{P}}}\left[e^{-Tr[uS_t]}\right] &= \mathbb{E}^{\tilde{\mathbb{P}}}_{(Q^\top)^{-1}s_0Q^{-1}}\left[e^{-Tr\left[uQ^\top\Sigma_t Q\right]}\right] \\ &= \mathbb{E}_{\Sigma_0}^{\tilde{\mathbb{P}}}\left[e^{-Tr\left[\left(QuQ^\top\right)\Sigma_t\right]}\right].\end{aligned}$$

Again, by relying on a Taylor expansion and Sylvester's rule of inertia, we obtain the following closed-form formula for the Laplace transform:

$$\varphi(t, s_0) = \det\left(e^{\left(Q^\top\right)^{-1} bQ^{-1}\log\left(I_d + 2tuQ^\top Q\right)^{-\frac{1}{2}}}\right) e^{-Tr\left[\left(I_d+2tuQ^\top Q\right)^{-1}us_0\right]}.$$

Now we consider the process under $\mathbb{P}$:

$$dS_t = \sqrt{S_t}dB_t Q + Q^\top dB_t^\top \sqrt{S_t} + \left(MS_t + S_t M^\top + b\right)dt,$$

where $dB_t = d\tilde{B}_t - \sqrt{S_t}M^\top Q^{-1}dt$, with an associated Girsanov kernel satisfying (2.1) as before. The system of Riccati ODE's satisfied by the joint Laplace transform of the process $S_t$ under $\mathbb{P}$ is given by:

$$\begin{aligned}\frac{d\psi}{dt} &= \psi M + M^\top \psi - 2\psi Q^\top Q\psi + v, \quad \psi(0) = w, \\ \frac{d\phi}{dt} &= Tr\left[b\psi(t)\right], \quad \phi(0) = 0.\end{aligned}$$

We realize that the ODE for $\psi$ is the same as in (3.1) , whose solution is known from Remark 1.1. We can recover the solution also for $\phi$ upon a direct integration. We show the calculation in detail:

$$\begin{aligned}\frac{d\phi}{dt} &= Tr\left[b\psi(t)\right] \\ &= Tr\left[b\left(\frac{\left(Q^\top Q\right)^{-1}M}{2} - \frac{Q^{-1}\sqrt{\bar{v}}kQ^{\top^{-1}}}{2}\right)\right].\end{aligned}$$

Integrating the ODE yields

$$\phi(t) = Tr\left[b\frac{\left(Q^\top Q\right)^{-1}M}{2}\right]t - \frac{1}{2}Tr\left[\left(Q^\top\right)^{-1} bQ^{-1}\sqrt{\bar{v}}\int_0^t kds\right].$$

We concentrate on the integral appearing in the second term:

$$\int_0^t kds = \int_0^t -\left(\sqrt{\bar{v}}\cosh(\sqrt{\bar{v}}s) + \bar{w}\sinh(\sqrt{\bar{v}}s)\right)^{-1}\left(\sqrt{\bar{v}}\sinh(\sqrt{\bar{v}}s) + \bar{w}\cosh(\sqrt{\bar{v}}s)\right)ds.$$

Define $f(s) = \sqrt{\bar{v}}\cosh(\sqrt{\bar{v}}s) + \bar{w}\sinh(\sqrt{\bar{v}}s)$ and let us differentiate it:

$$\frac{df}{ds} = \left(\sqrt{\bar{v}}\sinh(\sqrt{\bar{v}}s) + \bar{w}\cosh(\sqrt{\bar{v}}s)\right)\sqrt{\bar{v}},$$

hence we can write

$$\begin{aligned}
&-\frac{1}{2}Tr\left[\left(Q^{\top}\right)^{-1}bQ^{-1}\sqrt{\bar{v}}\int_0^t kds\right] \\
&= \frac{1}{2}Tr\left[\left(Q^{\top}\right)^{-1}bQ^{-1}\left(\log\left(\sqrt{\bar{v}}\cosh(\sqrt{\bar{v}}t) + \bar{w}\sinh(\sqrt{\bar{v}}t)\right) - \log\left(\sqrt{\bar{v}}\right)\right)\right] \\
&= \frac{1}{2}Tr\left[\left(Q^{\top}\right)^{-1}bQ^{-1}\log\left(\sqrt{\bar{v}}^{-1}\left(\sqrt{\bar{v}}\cosh(\sqrt{\bar{v}}t) + \bar{w}\sinh(\sqrt{\bar{v}}t)\right)\right)\right]
\end{aligned}$$

and the proof is complete. □

# Part 2

# Fixed-income market

CHAPTER 2

# The Wishart short rate model

## 1. Introduction

In the present chapter we focus on models where the short rate is given as

$$r_t = a + Tr\left[BX_t\right], \tag{1.1}$$

where $a \in \mathbb{R}_{\geq 0}$, $B$ is a symmetric positive definite matrix and $X = (X_t)_{t\geq 0}$ is a stochastic process on the cone of positive semidefinite matrices. We will provide a fairly general pricing formula for zero coupon bonds for this family of models. Then we will restrict to the **Wishart short rate model**. This kind of model has been suggested in Grasselli and Tebaldi (2008), then Buraschi et al. (2008) investigated the properties of this model with respect to many issues concerning the yield curve and interest rate derivatives. An analysis of the impact of the specification of the risk-premium is provided in Chiarella et al. (2010).

The contribution of this chapter is given by a new and explicit closed-form formula, which is based on the theoretical framework we developed in Chapter 1. This allows the easy computation of Bond prices. Moreover, we provide a set of sufficient conditions ensuring that this short rate model produces certain shapes of the yield curve. Our analysis of yield curve shapes is inspired by the work of Keller-Ressel and Steiner (2008), where a set of restrictions on the shapes of the yield curve is derived, under the assumption that the driving process is affine in the sense of Duffie et al. (2003).

Affine processes have been applied in finance in many contexts, for an application to interest rates, see Duffie and Kan (1996). More recently Duffie et al. (2003) and Keller-Ressel (2008) provided a full characterization for the state space $\mathbb{R}^m_{\geq 0} \times \mathbb{R}^n$. In Cuchiero et al. (2009), an analogous characterization is provided for the state space $S_d^+$. An alternative characterization is proposed in Grasselli and Tebaldi (2008), where the concept of solvable affine term structure model is discussed, both for the state space $\mathbb{R}^m_{\geq 0} \times \mathbb{R}^n$ and $S_d^+$.

Affine processes on positive semidefinite matrices have witnessed an increasing importance in applications in finance. The first application is due to Gourieroux and Sufana (2003, 2005). Applications to multi-factor stochastic volatility, and stochastic correlations may be found in Da Fonseca et al. (2008, 2007b, 2009, 2007a), Gourieroux et al. (2005) and Buraschi et al. (2010), for the pure diffusion case. Leippold and Trojani (2010) introduce a class of jump diffusions where the intensity is an affine function of the state variable. Jump processes on $S_d^+$ are treated in Barndorff-Nielsen and Stelzer (2010); Barndorff-Nielsen and Stelzer (2007), Mayerhofer et al. (2009), Muhle-Karbe et al. (2010) and Pigorsch and Stelzer (2009).

The present chapter is organized as follows: first we introduce the general setup of affine processes on the cone state space $S_d^+$, thus providing the general setup for the pricing problem. Then we restrict to the Wishart process and provide our new closed form formula for zero coupon bonds. Finally, by assuming that the Wishart process lies in the interior of $S_d^+$, we are also able to provide a set of sufficient conditions on the initial state of the process, which ensure that the model replicates certain shapes of the yield curve.

**1.1. Bond Prices.** In this section we derive a fairly general pricing formula for zero coupon bonds. Before this, we would like to spend a couple of words to recall an important fact concerning the risk neutral measure that we will use for pricing purposes. From Björk (2004) we know that it is quite tempting to consider the short rate as a traded asset and treat zero coupon bonds as derivatives written on the short rate. Unfortunately, the short rate is not a traded asset, hence the bond market is arbitrage free but not complete. This means that in general there exist many risk neutral measures. This implies that the reference risk neutral measure $\mathbb{Q}$ will be inferred in general from market prices, and so will result from a calibration procedure. Let $B \in S_d^+$ then, according to Definition 0.1, we have:

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-Tr[BX_t]}\right] = e^{-\phi(t,B)-Tr[\psi(t,B)x]}. \tag{1.2}$$

More generally, for $t, s, > 0$:

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-Tr[BX_{t+s}]}|\mathcal{F}_t\right] = e^{-\phi(s,B)-Tr[\psi(s,B)x_t]}. \tag{1.3}$$

In what follows, by defining $\tau = T - t$, we will see that a similar formula holds for the price of a zero coupon bond which is computed, when the short rate is given as in (1.1), via the following expectation:

$$P_t(\tau) := \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T a+Tr[BX_u]du}|\mathcal{F}_t\right]. \tag{1.4}$$

This expectation satisfies the following Kolmogorov backward equation:

$$\frac{\partial P_t}{\partial \tau} = \mathcal{A}P - (a + Tr\left[BX\right])P, \quad P(0) = 1, \tag{1.5}$$

where the infinitesimal generator of the process $X$ is reported in Theorem 0.1. We introduce an exponentially affine guess given by the following:

$$P_t(\tau) = \exp\left\{-\tilde{\phi}(\tau, B) - Tr\left[\tilde{\psi}(\tau, B)X\right]\right\}, \tag{1.6}$$

so that:

$$\frac{\partial P_t}{\partial \tau} = \left(-\frac{\partial \tilde{\phi}}{\partial \tau} - Tr\left[\frac{\partial \tilde{\psi}}{\partial \tau}X\right]\right)P_t, \tag{1.7}$$

$$\mathcal{A}e^{-\tilde{\phi}(\tau,B)-Tr\left[\tilde{\psi}(\tau,B)X\right]} = e^{-\tilde{\phi}(\tau,B)}\mathcal{A}e^{-Tr\left[\tilde{\psi}(\tau,B)X\right]}, \tag{1.8}$$

and (see always Theorem 0.1):

$$\begin{aligned}
\mathcal{A}e^{-Tr\left[\tilde{\psi}(\tau,B)X\right]} &= \left(-F(\tilde{\psi}(\tau, B)) - Tr\left[R(\tilde{\psi}(\tau, B))X\right]\right)e^{-Tr\left[\tilde{\psi}(\tau,B)X\right]}\\
&= \Bigg\{-Tr\left[b\tilde{\psi}(\tau, B)\right] + \int_{S_d^+\setminus\{0\}}\left(e^{-Tr\left[\tilde{\psi}(\tau,B)\xi\right]} - 1\right)m(d\xi)\\
&+ Tr\Bigg[\Bigg(2\tilde{\psi}(\tau, B)\alpha\tilde{\psi}(\tau, B) - \beta^\top(\tilde{\psi}(\tau, B))\\
&+ \int_{S_d^+\setminus\{0\}}\frac{e^{-Tr\left[\tilde{\psi}(\tau,B)\xi\right]} - 1 + Tr\left[\chi(\xi)\tilde{\psi}(\tau, B)\right]}{\parallel \xi \parallel^2 \wedge 1}\mu(d\xi)\Bigg)X\Bigg]\Bigg\}\\
&\times e^{-Tr\left[\tilde{\psi}(\tau,B)X\right]}.
\end{aligned} \tag{1.9}$$

In summary, we obtain:

$$
\begin{aligned}
&-\frac{\partial \tilde{\phi}}{\partial \tau} - Tr\left[\frac{\partial \tilde{\psi}}{\partial \tau} X\right] = -Tr\left[b\tilde{\psi}(\tau, B)\right] + \int_{S_d^+ \setminus \{0\}} \left(e^{-Tr\left[\tilde{\psi}(\tau,B)\xi\right]} - 1\right) m(d\xi) \\
&+ Tr\left[\left(2\tilde{\psi}(\tau, B)\alpha\tilde{\psi}(\tau, B) - \beta^\top(\tilde{\psi}(\tau, B))\right.\right. \\
&\left.\left. + \int_{S_d^+ \setminus \{0\}} \frac{e^{-Tr\left[\tilde{\psi}(\tau,B)\xi\right]} - 1 + Tr\left[\chi(\xi)\tilde{\psi}(\tau, B)\right]}{\parallel \xi \parallel^2 \wedge 1} \mu(d\xi)\right) X\right] \\
&- (a + Tr\,[BX])\,. \qquad (1.10)
\end{aligned}
$$

Identify terms to obtain the system of (matrix) ODE's:

$$
\begin{aligned}
\frac{\partial \tilde{\phi}}{\partial \tau} &= \mathcal{F}\left(\tilde{\psi}(\tau, B)\right) = Tr\left[b\tilde{\psi}(\tau, B)\right] \\
&- \int_{S_d^+ \setminus \{0\}} \left(e^{-Tr\left[\tilde{\psi}(\tau,B)\xi\right]} - 1\right) m(d\xi) + a, \qquad (1.11)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \tilde{\psi}}{\partial \tau} &= \mathcal{R}\left(\tilde{\psi}(\tau, B)\right) = -2\tilde{\psi}(\tau, B)\alpha\tilde{\psi}(\tau, B) + \beta^\top(\tilde{\psi}(\tau, B)) \\
&- \int_{S_d^+ \setminus \{0\}} \frac{e^{-Tr\left[\tilde{\psi}(\tau,B)\xi\right]} - 1 + Tr\left[\chi(\xi)\tilde{\psi}(\tau, B)\right]}{\parallel \xi \parallel^2 \wedge 1} \mu(d\xi) + B. \qquad (1.12)
\end{aligned}
$$

We have thus proven the following:

PROPOSITION 2.1. *Let $X$ be a conservative affine process on $S_d^+$ under the risk neutral probability measure $\mathbb{Q}$. Let the short rate be given as:*

$$
r_t = a + Tr\,[BX_t]\,, \qquad (1.13)
$$

*then the price of a zero-coupon bond is given by:*

$$
\begin{aligned}
P_t(\tau) :&= \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T a + Tr[BX_u]du} | \mathcal{F}_t\right] \\
&= \exp\left\{-\tilde{\phi}(\tau, B) - Tr\left[\tilde{\psi}(\tau, B)X_t\right]\right\}, \qquad (1.14)
\end{aligned}
$$

*where $\tilde{\phi}$ and $\tilde{\psi}$ satisfy the following ODE's:*

$$
\begin{aligned}
&\frac{\partial \tilde{\phi}}{\partial \tau} = \mathcal{F}\left(\tilde{\psi}(\tau, B)\right) = Tr\left[b\tilde{\psi}(\tau, B)\right] - \int_{S_d^+ \setminus \{0\}} \left(e^{-Tr\left[\tilde{\psi}(\tau,B)\xi\right]} - 1\right) m(d\xi) + a, \\
&\tilde{\phi}(0, B) = 0, \qquad (1.15)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \tilde{\psi}}{\partial \tau} &= \mathcal{R}\left(\tilde{\psi}(\tau, B)\right) = -2\tilde{\psi}(\tau, B)\alpha\tilde{\psi}(\tau, B) + \beta^\top(\tilde{\psi}(\tau, B)) \\
&- \int_{S_d^+ \setminus \{0\}} \frac{e^{-Tr\left[\tilde{\psi}(\tau,B)\xi\right]} - 1 + Tr\left[\chi(\xi)\tilde{\psi}(\tau, B)\right]}{\parallel \xi \parallel^2 \wedge 1} \mu(d\xi) + B, \\
\tilde{\psi}(0, B) &= 0. \qquad (1.16)
\end{aligned}
$$

REMARK 2.1. *In summary, with respect to the case where we are interested in the Laplace transform of the process (Theorem 0.1), we notice that when we consider the integrated process we have the following:*

$$\mathcal{F}(u) := F(u) + a, \tag{1.17}$$

$$\mathcal{R}(u) := R(u) + B. \tag{1.18}$$

In the next sections, we will be working repeatedly with the functions $\mathcal{F}(u)$ and $\mathcal{R}(u)$ defined above. They will permit us to characterize in a very precise way the asymptotic behavior of the yield curve for large maturities and they will be a key ingredient to derive our sufficient conditions concerning the shapes of the curve.

## 2. The Wishart short rate model

**2.1. Some Properties of the Matrix Exponential.** We recall some background on the matrix exponential, which will be useful in the sequel. First, we provide the following:

DEFINITION 2.1. *Let $A$ be a matrix with entries in $\mathbb{C}$, then we define:*

$$e^{A\tau} := \sum_{k=0}^{\infty} \frac{A^k \tau^k}{k!}. \tag{2.1}$$

In the sequel we will look at the asymptotic behavior of the yield curve, so the following lemma will be useful:

LEMMA 2.1. *Let $A \in M_d$. Assume $\Re(\lambda(A)) < 0, \quad \forall \lambda \in \sigma(A)$, then:*

$$\lim_{\tau \to \infty} e^{A\tau} = 0 \in M_{d \times d}. \tag{2.2}$$

PROOF. We follow Damm (2009). Assume that $A = SJS^{-1}$ is the Jordan form. Then

$$e^{A\tau} = S \begin{bmatrix} e^{J_1 \tau} & & \\ & \ddots & \\ & & e^{J_m \tau} \end{bmatrix} S^{-1}. \tag{2.3}$$

It suffices to understand what is $e^{J\tau}$ for a Jordan block:

$$J = \lambda I + N, \quad N = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}. \tag{2.4}$$

Notice that:

$$N^2 = \begin{bmatrix} 0 & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & 0 & 0 \\ & & & & 0 \end{bmatrix}, ..., N^{\nu} = \begin{bmatrix} 0 & 0 & 0 & & 1 \\ & & \ddots & \ddots & \\ & & & 0 & 0 \\ & & & 0 & 0 \\ & & & & 0 \end{bmatrix}, N^{\nu+1} = 0, \tag{2.5}$$

thus we have:

$$\frac{\tau^k}{k!} J^k = \frac{\tau^k \lambda^k}{k!} I + \frac{\tau^k \lambda^{k-1}}{(k-1)!} N + \frac{\tau^k \lambda^{k-2}}{2(k-2)!} N^2 + ... + \frac{\tau^k \lambda^{k-\nu}}{\nu!(k-\nu)!} N^{\nu} \tag{2.6}$$

and

$$\sum_{k=0}^{\infty} \frac{\tau^k}{k!} J^k = e^{\tau\lambda} I + \tau e^{\tau\lambda} N + \tau^2 e^{\tau\lambda} N^2 + ...$$

$$= e^{\tau\lambda} \begin{bmatrix} 1 & \tau & \frac{\tau^2}{2} & \cdots & \frac{\tau^\nu}{\nu!} \\ & \ddots & & & \vdots \\ & & \ddots & & \vdots \\ & & & \ddots & \tau \\ & & & & 1 \end{bmatrix}. \tag{2.7}$$

Now if $\Re(\lambda) < 0$, then $\left| e^{\lambda\tau} \tau^\nu \right| < M e^{\gamma\tau}$, for some $M > 0$ and $0 > \gamma > \Re(\lambda)$, hence the same holds for $e^{A\tau}$, if $\sigma(A) \in \mathbb{C}_-$. Now let $\tau \to \infty$ and the result follows. □

This fact allows us to determine the asymptotic behavior of the following functions:

LEMMA 2.2. *Let $O \in S_d^+$, define:*

$$\sinh(O\tau) = \frac{e^{O\tau} - e^{-O\tau}}{2}, \quad \cosh(O\tau) = \frac{e^{O\tau} + e^{-O\tau}}{2} \tag{2.8}$$

*and*

$$\tanh(O\tau) = (\cosh(O\tau))^{-1} \sinh(O\tau), \quad \coth(O\tau) = (\sinh(O\tau))^{-1} \cosh(O\tau), \tag{2.9}$$

*then*

$$\lim_{\tau\to\infty} \tanh(O\tau) = \lim_{\tau\to\infty} \coth(O\tau) = I_d. \tag{2.10}$$

PROOF.

$$\lim_{\tau\to\infty} \tanh(O\tau) = \lim_{\tau\to\infty} \left(I_d + e^{-2O\tau}\right)^{-1} \left(I_d - e^{-2O\tau}\right) = I_d. \tag{2.11}$$

The second equality follows along the same lines. □

**2.2. Closed-Form Pricing Formulae in the General Diffusion Model.** In this section we consider a diffusion model for the short rate. The driving process we use was first considered in the seminal paper by Bru (1991), however, in the present dissertation, following the standard literature on Wishart process, we will be dealing with a slight generalization. Using the terminology of Bru, we assume that the law of $X_t$ is $WIS_d(x_0, \alpha, M, Q)$ under the risk neutral measure $Q$. $X_t$ is the solution of the following SDE:

$$dX_t = \left(b + MX_t + X_t M^\top\right) dt + \sqrt{X_t} dW_t Q + Q^\top dW_t^\top \sqrt{X_t}, \tag{2.12}$$

where $M, Q \in GL(d)$, $b = \alpha Q^\top Q$. We further assume[1] $\alpha \geq d + 1$ and $x_0 \in S_d^{++}$. These last assumptions, according to Theorem 2.2 in Mayerhofer et al. (2009), allow us to claim that there exists a strong solution to (2.12) on the interval $[0, \tau_0)$, where the stopping time $\tau_0$ is defined as:

$$\tau_0 = \inf\{t \geq 0 | \det X_t = 0\}. \tag{2.13}$$

[1]The assumption on $\alpha$ implies that the process lies in the interior of the cone $S_d^+$, that we denote by $S_d^{++}$. This more restrictive assumption is required in order to derive the conditions on the shapes of the yield curve. All bond pricing formulae that we outline in the sequel hold true also for $\alpha > d - 1$. For more details, see Bru (1991).

Moreover, we have $\tau_0 = +\infty$ a.s. Finally, we notice that, in full analogy with the scalar square root process, the term $-\alpha Q^\top Q$ is related to the long term matrix $X_\infty$ via the following Lyapunov equation:

$$-\alpha Q^\top Q = MX_\infty + X_\infty M^\top, \tag{2.14}$$

so that, for the rest of this chapter, we make the following standard assumption in order to grant the mean reverting feature of the process $X_t$.

ASSUMPTION 2.1. *We require $\Re\lambda < 0$, $\forall \lambda \in \sigma(M)$. This requirement implies the convergence of the improper integral $\alpha \int_0^\infty e^{Ms} Q^\top Q e^{M^\top s} ds$, which satisfies the equation. We further assume that $\forall \tau \in \mathbb{R}_{\geq 0} \cup \{\infty\}$ the matrix $M - 2Q^\top Q\tilde{\psi}(\tau)$ has negative real eigenvalues.*

For this driving process, the price of a zero coupon bond is computed as follows. We use the shorthand notation $\tilde{\phi}(\tau, B) = \tilde{\phi}(\tau)$ and $\tilde{\psi}(\tau, B) = \tilde{\psi}(\tau)$.

PROPOSITION 2.2. *Let the short rate be given as in* (1.1)*, for a process $X_t$ with law $WIS_d(x_0, \alpha, M, Q)$. Let $B \in S_d^{++}$ and set $\tau = T - t$. Then the price of a zero coupon bond is given by:*

$$\mathbb{E}^{\mathbb{Q}}_{X_t}\left[\exp\left\{-a\tau - Tr\left[\int_t^T BX_s ds\right]\right\}\right]$$
$$= \exp\left\{-\tilde{\phi}(\tau) - Tr\left[\tilde{\psi}(\tau)X_t\right]\right\}, \tag{2.15}$$

*where $\tilde{\psi}(\tau)$ and $\tilde{\phi}(\tau)$ solve the following system of ODE's:*

$$\frac{\partial\tilde{\phi}}{\partial\tau} = Tr\left[\alpha Q^\top Q\tilde{\psi}(\tau)\right] + a, \quad \tilde{\phi}(0) = 0, \tag{2.16}$$

$$\frac{\partial\tilde{\psi}}{\partial\tau} = \tilde{\psi}(\tau)M + M^\top\tilde{\psi}(\tau) - 2\tilde{\psi}(\tau)Q^\top Q\tilde{\psi}(\tau) + B, \quad \tilde{\psi}(0) = 0. \tag{2.17}$$

PROOF. Same arguments as in Proposition 2.1, given that (2.15) corresponds to the bond price in (1.14). □

REMARK 2.2. *In the present setting we have:*

$$\mathcal{R}\left(\tilde{\psi}(\tau)\right) = \tilde{\psi}(\tau)M + M^\top\tilde{\psi}(\tau) - 2\tilde{\psi}(\tau)Q^\top Q\tilde{\psi}(\tau) + B \tag{2.18}$$

$$\mathcal{F}\left(\tilde{\psi}(\tau)\right) = Tr\left[\alpha Q^\top Q\tilde{\psi}(\tau)\right] + a \tag{2.19}$$

*Moreover, a direct substitution of the terminal condition $\tilde{\psi}(0) = 0$ implies:*

$$\mathcal{R}\left(\tilde{\psi}(0)\right) = B. \tag{2.20}$$

The solution of the system of ODE's above may be computed by relying on the different approaches outlined in Chapter 1.

PROPOSITION 2.3. *The system of ODE's* (2.16)*,* (2.17) *admits the following solutions:*

- ***Matrix Cameron Martin approach (See Remark 1.1)***

$$\tilde{\phi}(\tau) = -\frac{\alpha}{2}\log\det\left(e^{-M\tau}\left(\cosh(\sqrt{\bar{v}}\tau) + \sinh(\sqrt{\bar{v}}\tau)k\right)\right) + a\tau, \tag{2.21}$$

$$\tilde{\psi}(\tau) = -\frac{Q^{-1}\sqrt{\bar{v}}kQ^{\top^{-1}}}{2} + \frac{\left(Q^\top Q\right)^{-1}M}{2}, \tag{2.22}$$

$$k = -\left(\sqrt{\bar{v}}\cosh(\sqrt{\bar{v}}\tau) + \bar{w}\sinh(\sqrt{\bar{v}}\tau)\right)^{-1}\left(\sqrt{\bar{v}}\sinh(\sqrt{\bar{v}}\tau) + \bar{w}\cosh(\sqrt{\bar{v}}\tau)\right),$$
$$\bar{v} = Q\left(2B + M^\top Q^{-1} Q^{\top^{-1}} M\right) Q^\top,$$
$$\bar{w} = Q\left(-\left(Q^\top Q\right)^{-1} M\right) Q^\top.$$

- ***Variation of constant approach (See the proof of Proposition 1.4)***

$$\tilde{\psi}(\tau) = \psi' + e^{\left(M^\top - 2\psi' Q^\top Q\right)\tau}\Bigg[(-\psi')^{-1}$$

$$+ 2\int_0^\tau e^{\left(M - 2Q^\top Q\psi'\right)s} Q^\top Q e^{\left(M^\top - 2\psi' Q^\top Q\right)s} ds\Bigg]^{-1} e^{\left(M - 2Q^\top Q\psi'\right)\tau}, \tag{2.23}$$

$$\tilde{\phi}(\tau) = Tr\left[\alpha Q^\top Q \int_0^\tau \tilde{\psi}(s) ds\right] + a\tau. \tag{2.24}$$

- ***Linearization approach (See the proof of Proposition 1.5)***

$$\tilde{\psi}(\tau) = D(\tau)^{-1} E(\tau), \quad E(0) = 0, D(0) = I_d \tag{2.25}$$

*with*

$$\begin{pmatrix} E(\tau) & D(\tau) \end{pmatrix} = \begin{pmatrix} E(0) & D(0) \end{pmatrix} exp\left\{\tau \begin{pmatrix} M & 2Q^\top Q \\ B & -M^\top \end{pmatrix}\right\}. \tag{2.26}$$

$$\tilde{\phi}(\tau) = \frac{\alpha}{2} Tr\left[log\left(D(\tau)\right) + M^\top \tau\right] + a\tau. \tag{2.27}$$

As in Chapter 1, by $\psi'$ we denote a solution to the algebraic Riccati equation (3.3). A direct consequence of the result above is the following:

COROLLARY 2.1.

$$\lim_{\tau \to \infty} \tilde{\psi}(\tau) = \psi'. \tag{2.28}$$

PROOF. Since $\lambda\left(M - 2Q^\top Q\psi'\right) < 0$ by assumption, we know that the integral in the solution for $\tilde{\psi}$ is convergent, moreover, from Lemma 2.1, we know that $e^{\left(M - 2Q^\top Q\psi'\right)\tau} \searrow 0$ as $\tau \to \infty$, hence the proof is complete. □

This last corollary tells us that the function $\tilde{\psi}$ tends to a stability point of the Riccati ODE. This allows us to claim that, as $\tau \to \infty$, we have $\mathcal{R}\left(\tilde{\psi}(\tau)\right) \searrow 0$.

## 3. Yield Curve Shapes

In this section we perform an investigation on the shapes of the yield curve produced by the Wishart short rate model. We will derive a set of sufficient conditions ensuring that certain shapes are attained. We will work with the general diffusion model and show how to replicate, normal, inverse or humped curves. In the appendix we will repeat the same analysis in a simpler version of the model where there will be further limitations on the possible shapes one can obtain. We use the standard dotted notation to represent derivatives w.r.t. time dimensions.

**3.1. Monotonicity of $\tilde{\psi}(\tau)$ and asymptotic behavior of the Yield Curve.** Here we report a result concerning the monotonicity of the function $\tilde{\psi}(\tau)$, which may be found in Buraschi et al. (2008). First we recall a result from control theory (see Brockett (1970)).

PROPOSITION 2.4. *(Matrix variation of constants formula) If $\mathbf{\Phi}_1(t,t_0)$ is the transition matrix of $\dot{x}(t) = A_1(t)x(t)$ and $\mathbf{\Phi}_2(t,t_0)$ is the transition matrix for $\dot{x}(t) = A_2^\top(t)x(t)$, then the solution of*

$$\dot{X}(t) = A_1(t)X(t) + X(t)A_2(t) + F(t), \tag{3.1}$$

*with the initial state vector $X(t_0)$, is given by:*

$$X(t) = \mathbf{\Phi}_1(t,t_0)X(t_0)\mathbf{\Phi}_2^\top(t,t_0) + \int_0^t \mathbf{\Phi}_1(t,t_0)F(s)\mathbf{\Phi}_2^\top(t,t_0)ds. \tag{3.2}$$

PROPOSITION 2.5. *Let $X = (X_s)_{t\le s\le T}$ be the stochastic process defined by the dynamics* (2.12)*. Then $\tilde{\psi}(\tau)$ is monotonically increasing in $\tau$, i.e., for $\tau_2 \ge \tau_1$, we have that $\tilde{\psi}(\tau_2) \succeq \tilde{\psi}(\tau_1)$.*

PROOF. First, we differentiate (2.17), so as to obtain the following:

$$\ddot{\tilde{\psi}}(\tau) = \dot{\tilde{\psi}}(\tau)M + M^\top\dot{\tilde{\psi}}(\tau) - 2\dot{\tilde{\psi}}Q^\top Q(\tau)\tilde{\psi}(\tau) - 2\tilde{\psi}(\tau)Q^\top Q\dot{\tilde{\psi}}(\tau). \tag{3.3}$$

Next we define $V(\tau) = M - 2Q^\top Q\tilde{\psi}(\tau)$. Then we may write:

$$\ddot{\tilde{\psi}}(\tau) = \dot{\tilde{\psi}}(\tau)V(\tau) + V^\top(\tau)\dot{\tilde{\psi}}(\tau), \tag{3.4}$$

which is solved by:

$$\dot{\tilde{\psi}}(\tau) = \mathbf{\Phi}(\tau,0)\dot{\tilde{\psi}}(0)\mathbf{\Phi}^\top(\tau,0), \tag{3.5}$$

for a state transition matrix $\mathbf{\Phi}(\tau,0)$ of the system matrix $V(\tau)$, solving $\dot{\mathbf{\Phi}}(\tau,0) = V^\top(\tau)\mathbf{\Phi}(\tau,0)$, $\mathbf{\Phi}(0,0) = I_d$. Substitution of the initial condition $\dot{\tilde{\psi}}(0) = B$ yields (see Remark 2.2 and recall from (1.12) that we have $\dot{\tilde{\psi}}(\tau) = \mathcal{R}(\tilde{\psi}(\tau))$):

$$\dot{\tilde{\psi}}(\tau) = \mathbf{\Phi}(\tau,0)B\mathbf{\Phi}^\top(\tau,0). \tag{3.6}$$

This derivative is positive semidefinite for $B \in S_d^+$, upon integration on the interval $(\tau_1,\tau_2)$ we obtain:

$$\tilde{\psi}(\tau_2) - \tilde{\psi}(\tau_1) = \int_{\tau_1}^{\tau_2}\dot{\tilde{\psi}}(u)du. \tag{3.7}$$

Therefore $\tilde{\psi}$ is an increasing function of $\tau$. □

As a consequence, we can derive the following useful corollary concerning the function $\mathcal{R}$, defined in Remark 2.2.

COROLLARY 2.2. $\forall\tau \in [0,\infty)$ *we have:*

$$\mathcal{R}\left(\tilde{\psi}(\tau)\right) \in S_d^+ \text{ for } B \in S_d^+, \tag{3.8}$$

$$\mathcal{R}\left(\tilde{\psi}(\tau)\right) \in S_d^{++} \text{ for } B \in S_d^{++}. \tag{3.9}$$

PROOF. From equation (3.6) we know that:

$$\dot{\tilde{\psi}}(\tau) = \mathbf{\Phi}(\tau,0)B\mathbf{\Phi}^\top(\tau,0) \tag{3.10}$$

But $\mathcal{R}\left(\tilde{\psi}(\tau)\right) = \dot{\tilde{\psi}}(\tau)$. This shows that $\mathcal{R}\left(\tilde{\psi}(\tau)\right)$ is a congruent transformation of B. According to Sylvester's law of inertia the signs of the eigenvalues are unchanged under congruent transformations, hence the claim. □

We now proceed to show the behavior of the yield curve as $\tau \to 0$ and $\tau \to \infty$. First of all we provide the following:

DEFINITION 2.2. *The zero coupon yield* $Y(\tau, X_t) : \mathbb{R}_{\geq 0} \times S_d^+ \to \mathbb{R}_{\geq 0}$ *is defined as:*

$$Y(\tau, X_t) = \frac{\tilde{\phi}(\tau)}{\tau} + \frac{Tr\left[\tilde{\psi}(\tau) X_t\right]}{\tau}. \tag{3.11}$$

*For fixed* $X_t$ *we call the function* $Y(\tau) = Y(\tau, X_t)$ *the* ***yield curve****.*

Then we show the following:

PROPOSITION 2.6. *Let* $X = (X_s)_{t \leq s \leq T}$ *be the stochastic process defined by the dynamics* (2.12)*. Then the following relations hold true:*

$$\lim_{\tau \to 0} Y(\tau) = r_t, \tag{3.12}$$

$$\lim_{\tau \to \infty} Y(\tau) = \mathcal{F}(\psi'), \tag{3.13}$$

*with* $\mathcal{F}$ *as in* (2.19).

PROOF. We start with (3.12). By using l'Hospital rule we may write the following:

$$\lim_{\tau \to 0} \frac{\tilde{\phi}(\tau)}{\tau} = \lim_{\tau \to 0} \mathcal{F}(\tilde{\psi}(\tau)) = \lim_{\tau \to 0} Tr\left[b\tilde{\psi}(\tau)\right] + a = a, \tag{3.14}$$

$$\begin{aligned}
\lim_{\tau \to 0} \frac{Tr\left[\tilde{\psi}(\tau) X_t\right]}{\tau} &= \lim_{\tau \to 0} Tr\left[\mathcal{R}(\tilde{\psi}(\tau)) X_t\right] \\
&= \lim_{\tau \to 0} Tr\left[\left(\tilde{\psi}(\tau) M + M^\top \tilde{\psi}(\tau) - 2\tilde{\psi}(\tau) Q^\top Q \tilde{\psi}(\tau) + B\right) X_t\right] \\
&= Tr\left[B X_t\right]. \qquad (3.15)
\end{aligned}$$

Putting together the two terms we obtain the result since $r_t = a + Tr\left[B X_t\right]$. To show (3.13), we notice that since $\tilde{\psi}(\tau) \to \psi'$ as $\tau \to \infty$, we have that $\mathcal{R}(\tilde{\psi}(\tau)) \to 0$. So using again l'Hospital rule and recalling (2.19) we have:

$$\lim_{\tau \to \infty} Y(\tau) = \lim_{\tau \to \infty} \mathcal{F}(\tilde{\psi}(\tau)) = \mathcal{F}(\psi'), \tag{3.16}$$

so that basically the one dimensional result in Keller-Ressel and Steiner is confirmed also in the present setting. □

**3.2. Yield Curve Shapes in the General Diffusion Model.** In this section we present a set of sufficient conditions, ensuring that the Wishart short rate model produces certain yield curve shapes. Due to the more general structure of the state space, arguments inspired by the scalar case allow us to derive only sufficient conditions, ensuring the attainability of certain yield curve shapes. The absence of a total order relation on $S_d^+$ does not allow us to rule out other possibilities (i.e. more complex shapes). In spite of this we believe that this result is interesting since, e.g. in a calibration setting, we are then able to put ex-ante some constraints to ensure that the model reproduces the yield curve shape observed on the market.

3.2.1. *Statement of the main result.* We begin with a definition:

DEFINITION 2.3. *Let the Yield curve be defined as in Definition 2.2. We say that the yield curve is:*

- ***normal*** *if $Y$ is a strictly increasing function of $\tau$,*
- ***inverse*** *if $Y$ is a strictly decreasing function of $\tau$,*
- ***humped*** *if $Y$ has a local maximum and no minimum on $[0, \infty)$.*

We will see that for our particular choice of the model, the arguments employed in Keller-Ressel and Steiner (2008) may be easily extended, with some adjustments due to the different state space. As in their setting, we report this lemma.

LEMMA 2.3. *A strictly convex or a strictly concave real function on $\mathbb{R}$ intersects an affine function in at most two points. In the case of two intersection points $p_1 < p_2$, the convex function lies strictly below the affine function on the interval $(p_1, p_2)$; if the function is concave it lies strictly above the affine function on $(p_1, p_2)$.*

Before we present the main result, we would like to recall some facts concerning the solution of a class of matrix equations (Wimmer (2009) Theorem 1.29):

THEOREM 2.1. *Let $A \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{m\times m}$, then we have the following:*

- *$\forall\, D \in \mathbb{C}^{m\times n}$ the equation*

$$XA + BX = D \tag{3.17}$$

*has a unique solution if and only if $\alpha + \beta \neq 0$, $\forall \alpha \in \sigma(A), \beta \in \sigma(B)$.*

- *If $\Re(\alpha + \beta) < 0$ $\forall \alpha \in \sigma(A), \beta \in \sigma(B)$ then the following improper integral is convergent and solves the equation:*

$$X = -\int_0^\infty e^{Bs} D e^{As} ds. \tag{3.18}$$

If $B = A^\top$ and $D = -Q$ for $Q \in S_d^+$, we have the following corollary:

COROLLARY 2.3. $\Re\left(\lambda\left(A\right)\right) < 0 \leftrightarrow A^\top X + XA = -Q \quad X, Q \in S_d^+$

Now, we report our main result on the shapes of the yield curve:

THEOREM 2.2. *Consider a short rate model in which the risk neutral dynamics of the short rate is driven by the process $X_t$, defined by the dynamics* (2.12). *Let $B \in S_d^{++}$. Define $M^\star := M - 2Q^\top Q\psi'$ and:*

$$b_{norm} := \alpha \int_0^\infty e^{M^\star s} Q^\top Q e^{M^{\star\top} s} ds, \qquad b_{inv} := \alpha \int_0^\infty e^{Ms} Q^\top Q e^{M^\top s} ds. \tag{3.19}$$

*Then the following holds:*

- *The yield curve is normal if $X_t \prec b_{norm}$.*
- *The yield curve is inverse if $X_t \succ b_{inv}$.*
- *The yield curve is humped if $b_{norm} \prec X_t \prec b_{inv}$.*

3.2.2. *Proof of Theorem 2.2.* We define the function $\mathcal{H}(\tau) : \mathbb{R}_{\geq 0} \to \mathbb{R}$ by

$$\mathcal{H}(\tau) := Y(\tau, X_t)\tau = \tilde{\phi}(\tau) + Tr\left[\tilde{\psi}(\tau)X_t\right]. \tag{3.20}$$

Recalling the system of equations satisfied by $\tilde{\phi}$ and $\tilde{\psi}$, given by equations (2.16) and (2.17), we can compute the derivatives of this function. For the first derivative we have:

$$\begin{aligned}\mathcal{H}'(\tau) &= Tr\left[\alpha Q^\top Q\tilde{\psi}(\tau)\right] + a \\ &+ Tr\left[\left(\tilde{\psi}(\tau)M + M^\top\tilde{\psi}(\tau) - 2\tilde{\psi}(\tau)Q^\top Q\tilde{\psi}(\tau)\right)X_t\right],\end{aligned} \tag{3.21}$$

whereas for the second we have:

$$\begin{aligned}\mathcal{H}''(\tau) = Tr\Big[&\alpha Q^\top Q\mathcal{R}\left(\tilde{\psi}(\tau)\right) \\ &+ \left(\mathcal{R}\left(\tilde{\psi}(\tau)\right)M + M^\top\mathcal{R}\left(\tilde{\psi}(\tau)\right) - 2\mathcal{R}\left(\tilde{\psi}(\tau)\right)Q^\top Q\tilde{\psi}(\tau)\right. \\ &\left. - 2\tilde{\psi}(\tau)Q^\top Q\mathcal{R}\left(\tilde{\psi}(\tau)\right)\right)X_t\Big].\end{aligned} \tag{3.22}$$

Now, notice the following:

$$\begin{aligned}Tr\left[M^\top\mathcal{R}\left(\tilde{\psi}(\tau)\right)X_t\right] &= Tr\left[\mathcal{R}\left(\tilde{\psi}(\tau)\right)X_tM^\top\right], \\ Tr\left[-2\mathcal{R}\left(\tilde{\psi}(\tau)\right)Q^\top Q\tilde{\psi}(\tau)X_t\right] &= Tr\left[-2\tilde{\psi}(\tau)Q^\top Q\mathcal{R}\left(\tilde{\psi}(\tau)\right)X_t\right].\end{aligned}$$

The second equality being justified since the matrices involved are symmetric. This means that we may rewrite $\mathcal{H}''(\tau)$ as follows:

$$\begin{aligned}\mathcal{H}''(\tau) &= Tr\left[\mathcal{R}\left(\tilde{\psi}(\tau)\right)\left(\overbrace{\alpha Q^\top Q + MX_t + X_tM^\top - 4Q^\top Q\tilde{\psi}(\tau)X_t}^{=:k(\tau)}\right)\right] \\ &= Tr\left[\mathcal{R}\left(\tilde{\psi}(\tau)\right)k(\tau)\right].\end{aligned} \tag{3.23}$$

Finally, we can equivalently work with a symmetrization of the function $k(\tau)$:

$$\begin{aligned}\mathcal{H}''(\tau) &= Tr\left[\mathcal{R}\left(\tilde{\psi}(\tau)\right)k(\tau)\right] \\ &= Tr\left[\mathcal{R}\left(\tilde{\psi}(\tau)\right)\frac{k(\tau) + k^\top(\tau)}{2}\right] \\ &= Tr\left[\mathcal{R}\left(\tilde{\psi}(\tau)\right)\left(\overbrace{\alpha Q^\top Q + \left(M - 2Q^\top Q\tilde{\psi}(\tau)\right)X_t + X_t\left(M^\top - 2\tilde{\psi}(\tau)Q^\top Q\right)}^{=:\tilde{k}(\tau)}\right)\right] \\ &= Tr\left[\mathcal{R}\left(\tilde{\psi}(\tau)\right)\tilde{k}(\tau)\right].\end{aligned} \tag{3.24}$$

To start the derivation of the sufficient conditions for the shapes of the yield curve we look at $\mathcal{H}''(0)$. A sufficient condition for $\mathcal{H}''(0) = 0$ is $\tilde{k}(0) = 0$, i.e.:

$$MX_t + X_tM^\top = -\alpha Q^\top Q. \tag{3.25}$$

The solution to this equation is given, according to Theorem 2.1, by:

$$b_{inv} = \alpha\int_0^\infty e^{Ms}Q^\top Qe^{M^\top s}ds. \tag{3.26}$$

Recalling that the eigenvalues of $M$ are negative (see Assumption 2.1) we have that:

$$\text{if } X_t \succ b_{inv} \text{ then } \mathcal{H}''(0) < 0, \tag{3.27}$$

$$\text{if } X_t \prec b_{inv} \text{ then } \mathcal{H}''(0) > 0. \tag{3.28}$$

Recall that $b_{inv}$ above is defined as the solution of the Lyapunov equation resulting from $\tilde{k}(\tau)$ when $\tau = 0$, by noting that $\tilde{\psi}(0) = 0$. Then consider the equation:

$$\alpha Q^\top Q + \left(M - 2Q^\top Q\tilde{\psi}(\tau)\right) X_t + X_t \left(M^\top - 2\tilde{\psi}(\tau)Q^\top Q\right) = 0. \tag{3.29}$$

By recalling that $\tilde{\psi}(\tau) \nearrow \psi'$ as $\tau \to \infty$, we have exactly a solution at infinity if $X_t$ solves:

$$\alpha Q^\top Q + \left(M - 2Q^\top Q\psi'\right) X_t + X_t \left(M^\top - 2\psi' Q^\top Q\right) = 0. \tag{3.30}$$

The solution to the equation above is:

$$b_{norm} = \alpha \int_0^\infty e^{\left(M - 2Q^\top Q\psi'\right)s} Q^\top Q e^{\left(M^\top - 2\psi' Q^\top Q\right)s} ds. \tag{3.31}$$

In the sequel we will prove that this is the crucial level ensuring the presence of a zero for $\mathcal{H}''(\tau)$. The next lemma establishes an order relation between $b_{inv}$ and $b_{norm}$.

LEMMA 2.4. $b_{inv} \succ b_{norm}$.

PROOF. Consider the function $\tilde{k}(\tau)$. It is easy to see that:

- if $X_t = b_{inv}$, then $\tilde{k}(\tau) = -2Q^\top Q\tilde{\psi}(\tau)b_{inv} - 2b_{inv}\tilde{\psi}(\tau)Q^\top Q$,
- if $X_t = b_{norm}$, then

$$\tilde{k}(\tau) = 0 - 2Q^\top Q\left(\tilde{\psi}(\tau) - \psi'\right) b_{norm} - b_{norm}\left(\tilde{\psi}(\tau) - \psi'\right) Q^\top Q2 \in S_d^+, \tag{3.32}$$

since $\psi' \succ \tilde{\psi}(\tau) \quad \forall \tau$.

This shows that $\tilde{k}(\tau)|_{X_t = b_{norm}} \succ \tilde{k}(\tau)|_{X_t = b_{inv}}$. Notice now that, by expliciting the dependence of $\tilde{k}(\tau)$ on $X_t$ we have

$$\tilde{k}(\tau, X_t) = \alpha Q^\top Q + \left(M - 2Q^\top Q\tilde{\psi}(\tau)\right) X_t + X_t \left(M^\top - 2\tilde{\psi}(\tau)Q^\top Q\right), \tag{3.33}$$

where $M - 2Q^\top Q\tilde{\psi}(\tau)$ has negative eigenvalues by Assumption 2.1. It thus follows that $b_{inv} \succ b_{norm}$. □

Let us now prove the following:

LEMMA 2.5. *If* $X_t \succ b_{inv}$, *then* $\mathcal{H}''(\tau) < 0$, $\forall \tau \in (0, \infty)$.

PROOF. Let $X_t = b_{inv} + C$, $C \in S_d^+$, then:

$$\begin{aligned} \tilde{k}(\tau) &= -2Q^\top Q\tilde{\psi}(\tau)b_{inv} - b_{inv}\tilde{\psi}(\tau)Q^\top Q2 \\ &\quad + \left(M - 2Q^\top Q\tilde{\psi}(\tau)\right) C + C\left(M^\top - \tilde{\psi}(\tau)Q^\top Q2\right) \end{aligned} \tag{3.34}$$

This means that $\tilde{k}$ is symmetric with negative eigenvalues (see Assumption 2.1), $\forall \tau \in [0, \infty)$, but then, being $\mathcal{R}(\tilde{\psi}(\tau)) \in S_d^{++}$, it follows that $\mathcal{H}''(\tau) < 0$. □

Next, we prove an existence result for $\tau^\star \in (0, \infty)$ s.t. $\mathcal{H}''(\tau^\star) = 0$.

LEMMA 2.6. *If* $b_{inv} \succ X_t \succ b_{norm}$, *then* $\exists \tau^\star \in (0, \infty)$ *s.t.* $\mathcal{H}''(\tau^\star) = 0$.

PROOF. Let $X_t = b_{norm} + C$, $C \in S_d^+$ s.t. $X_t \prec b_{inv}$. Then:

$$\tilde{k}(\tau) = \alpha Q^\top Q + \left(M - 2Q^\top Q\tilde{\psi}(\tau)\right)(b_{norm} + C) + (b_{norm} + C)\left(M^\top - 2\tilde{\psi}(\tau)Q^\top Q\right). \tag{3.35}$$

But this allows us to claim that:

$$\lim_{\tau\to\infty} \tilde{k}(\tau) = \left(M - 2Q^\top Q\psi'\right)C + C\left(M^\top - 2\psi' Q^\top Q\right). \tag{3.36}$$

This matrix is symmetric with negative eigenvalues. Recall that since $b_{inv} \succ X_t \succ b_{norm}$, we have that $\tilde{k}(0) \in S_d^{++}$. Then we observe that $\tilde{k}(\tau)$ is a continuous function of $\tau$, meaning that there must exist a $\tau'$ s.t., for $\tau > \tau'$, the eigenvalues of $\tilde{k}(\tau)$ are negative. We can now look at $\mathcal{H}''(\tau)$. We recall that:

$$\mathcal{H}''(\tau) = Tr\left[\mathcal{R}\left(\tilde{\psi}(\tau)\right)\tilde{k}(\tau)\right].$$

From Corollary 2.2 we have $\mathcal{R}\left(\tilde{\psi}(\tau)\right) \in S_d^{++}$. We notice also that $\mathcal{H}''$ is a continuous function of $\tau$. Furthermore, we have $\mathcal{H}''(0) > 0$ because $b_{inv} \succ X_t$, see (3.28). From the previous discussion regarding $\tilde{k}(\tau)$ we have that, for $\tau > \tau'$, the second derivative is of the form:

$$\mathcal{H}''(\tau) = Tr\left[\mathcal{R}\left(\tilde{\psi}(\tau)\right)(-\mathcal{K})\right], \tag{3.37}$$

where $\mathcal{K}$ is a symmetric matrix with negative eigenvalues. This means that the second derivative will be negative. By recalling the positiveness of the starting value and the continuity property w.r.t. $\tau$, thanks to the mean value theorem, we can argue that there must exist a $\tau^\star$ s.t. $\mathcal{H}''(\tau) = 0$ as we wanted. □

Along the same lines we can prove the following:

LEMMA 2.7. *If $X_t \prec b_{norm}$, then $\nexists \tau^\star \in (0, \infty)$ s.t. $\mathcal{H}''(\tau^\star) = 0$.*

We finally prove a a result allowing us to conclude that the zero of $\mathcal{H}''(\tau)$ is unique.

LEMMA 2.8. *$\tilde{k}(\tau)$ is monotonically decreasing i.e., for $\tau_2 > \tau_1$, we have $\tilde{k}(\tau_2) - \tilde{k}(\tau_1) \in S_d^-$.*

PROOF. Differentiate $\tilde{k}(\tau)$, so as to obtain:

$$\dot{\tilde{k}}(\tau) = -2Q^\top Q\mathcal{R}\left(\tilde{\psi}(\tau)\right)X_t - 2X_t\mathcal{R}\left(\tilde{\psi}(\tau)\right)Q^\top Q. \tag{3.38}$$

Then we can write:

$$\tilde{k}(\tau_2) - \tilde{k}(\tau_1) = -2\int_{\tau_1}^{\tau_2}\left(Q^\top Q\mathcal{R}\left(\tilde{\psi}(s)\right)X_t + X_t\mathcal{R}\left(\tilde{\psi}(s)\right)Q^\top Q\right)ds. \tag{3.39}$$

Since the RHS is a symmetric matrix with negative eigenvalues, we have the claim. □

Now, since $\tilde{k}(\tau)$ is monotonically decreasing, we obtain that if there exists a value for $\tau$ s.t. $\tilde{k}(\tau) = 0$, then this point in time must be unique.

By relying on Lemmas 2.4, 2.5, 2.6, 2.7 and 2.8 and the condition on the sign of $\mathcal{H}''(0)$ in (3.28) and (3.27), we can argue the following:

- if $X_t \prec b_{norm}$, then $\mathcal{H}$ is strictly convex on $(0, \infty)$.
- if $b_{norm} \prec X_t \prec b_{inv}$, then $\mathcal{H}$ is strictly convex on $(0, \tau^\star)$ and strictly concave on $(\tau^\star, \infty)$.
- if $X_t \succ b_{inv}$, then $\mathcal{H}$ is strictly concave on $(0, \infty)$.

We use these findings on the convexity of $\mathcal{H}$ to determine our conclusions on the convexity of the yield curve. We consider the equation

$$\mathcal{H}(\tau) = c\tau, \qquad \tau \in [0, \infty), \tag{3.40}$$

for some fixed $c \in \mathbb{R}$. Since $\mathcal{H}(0) = 0$ (see (3.20)), this equation has at least one solution, i.e. $\tau_0 = 0$. Now if $X_t \succ b_{inv}$, then $\mathcal{H}$ is strictly concave on $[0, \infty)$, and according to Lemma 2.3, equation (3.40) has at most one additional solution, $\tau_1$. When the solution exists, $\mathcal{H}(\tau)$ crosses $c\tau$ from above at $\tau_1$.
If $X_t \prec b_{norm}$, then $\mathcal{H}(\tau)$ is strictly convex on $[0, \infty)$ and again has at most one additional solution $\tau_2$. If the solution exists, $\mathcal{H}(\tau)$ crosses $c\tau$ from below at $\tau_2$.
In the final case, i.e. if $b_{norm} \prec X_t \prec b_{inv}$, there exists a $\tau^\star$, the zero of $\mathcal{H}''(\tau)$, such that $\mathcal{H}(\tau)$ is strictly convex on $(0, \tau^\star)$ and strictly concave on $(\tau^\star, \infty)$. This implies that there can exist at most two additional solutions $\tau_1, \tau_2$ to (3.40), with $\tau_1 < \tau^\star < \tau_2$, such that $c\tau$ is crossed from below at $\tau_1$ and from above at $\tau_2$. By definition, every solution to (3.40), $\tau_0 = 0$ excluded, is also a solution to:

$$Y(\tau, X_t) = c, \qquad \tau \in (0, \infty), \tag{3.41}$$

with $X_t$ fixed. The properties of crossing from above/below are preserved since $\tau$ is positive. This means that:

- if $X_t \succ b_{inv}$, then (3.41) has at most a single solution, i.e. every horizontal line is crossed by the yield curve in at most a single point. If it is crossed, it is crossed from above, hence we conclude that $Y(\tau, X_t)$ is a strictly decreasing function of $\tau$, meaning that the yield curve is inverse.
- $X_t \prec b_{norm}$, then again (3.41) has at most a single solution. If the solution is crossed, it is crossed from below, hence we conclude that $Y(\tau, X_t)$ is a strictly increasing function of $\tau$, meaning that the yield curve is normal.
- In the last case, i.e. if $b_{norm} \prec X_t \prec b_{inv}$, then we have at most two additional solutions. If they are crossed, the first is crossed from below and the second from above. This allows us to conclude that the yield curve is humped.

In Figure (1) we plot a visualization of the results of the theorem.

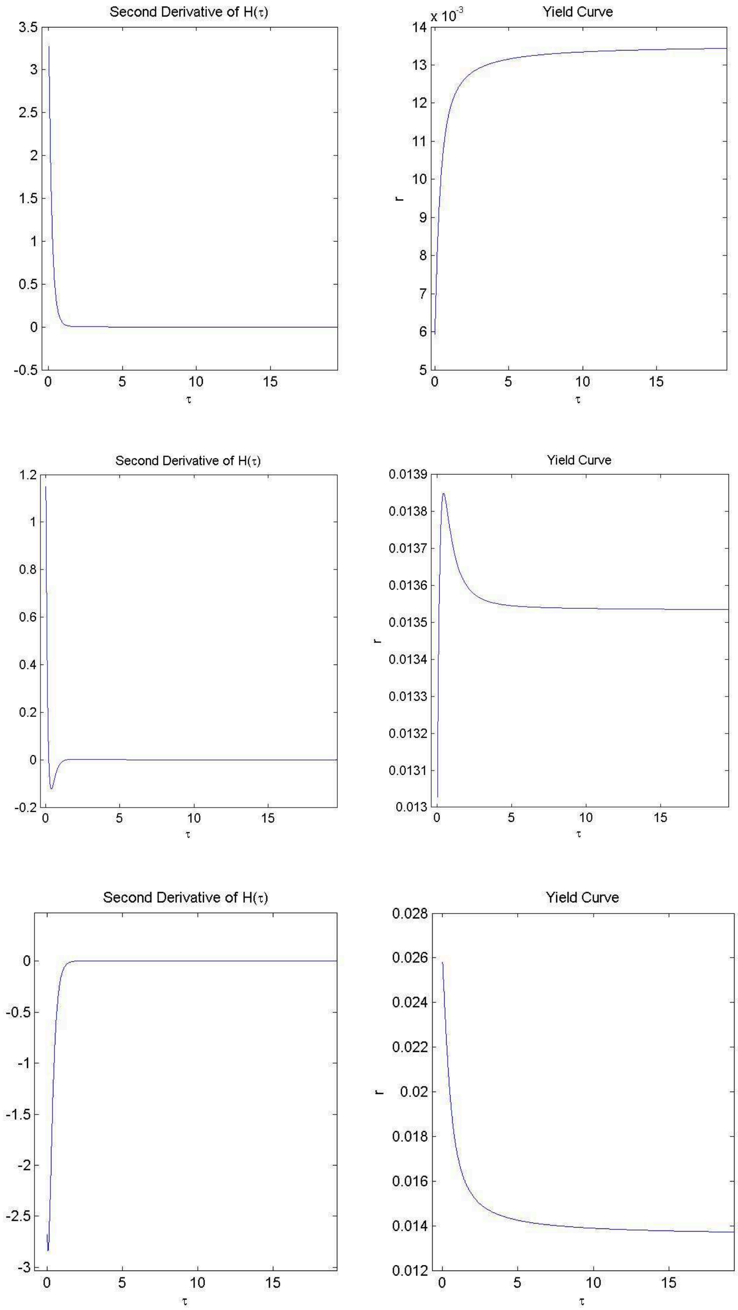


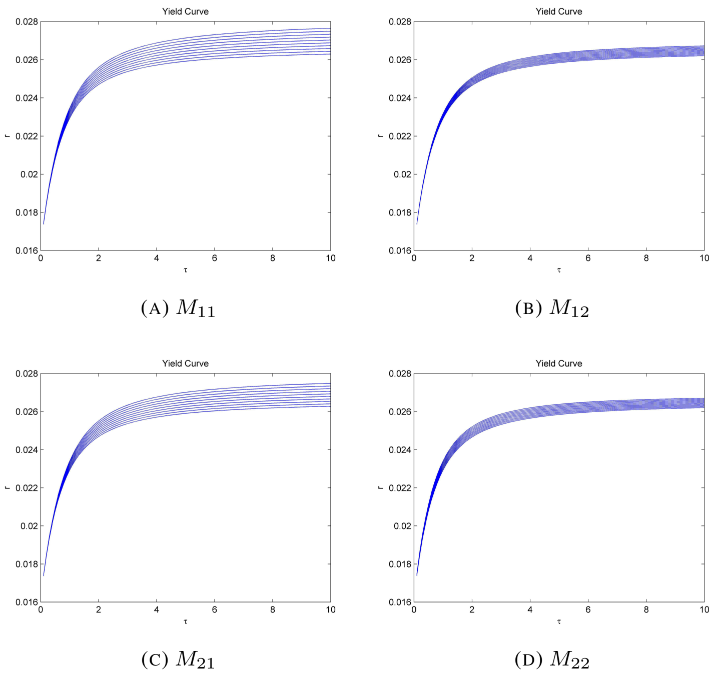

FIGURE 1. Yield curve shapes for different values of $X_t$

## 4. Discussion on the parameters

In this section we provide some intuition on the impacts of the parameters on the shape of the yield curve. As a starting point we use the following values for the parameters of the model:

$$\begin{aligned} X_0 &= \begin{pmatrix} 0.21 & 0.003 \\ 0.003 & 0.7 \end{pmatrix}, \quad M = \begin{pmatrix} -1.4 & 0.1 \\ 0.1 & -1.3 \end{pmatrix}, \\ Q &= \begin{pmatrix} 1 & 0.2 \\ 0.3 & 0.5 \end{pmatrix}, \quad \alpha = 3.1, \\ B &= \begin{pmatrix} 0.01 & 0.005 \\ 0.005 & 0.02 \end{pmatrix}. \end{aligned} \tag{4.1}$$

With these values, a numerical implementation shows that the yield curve is normal. In the following experiments we will perturbate single elements of the matrices and look at the impact on the yield curve.

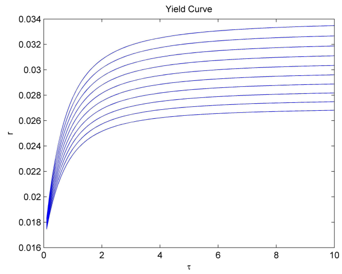



(A) $M_{11}$ (B) $M_{12}$

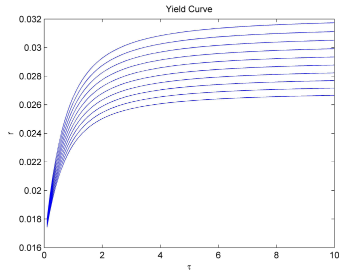
(C) $M_{21}$ (D) $M_{22}$

FIGURE 2. In this figure we show the effect of a perturbation of the single elements of the matrix $M$. We use a sequence of numbers $\eta = 0.01 : 0.01 : 0.1$ and add $\eta_i$ to one of the elements of $M$ while leaving the other elements unchanged. When we add $\eta_i$ to the elements on the main diagonal the YC is shifted upwards. The same happens with off-diagonal elements.

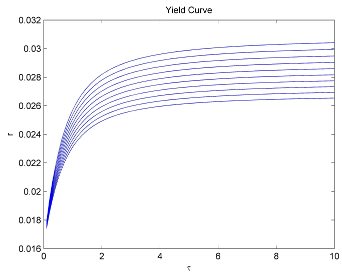
In Figure (2), we perturbate the matrix $M$, by introducing a sequence $\eta = 0.01 : 0.01 : 0.1.$ and we add the values of $\eta_i$ to the single entries of the matrix. It turns out that in all cases we have an upward shift of the yield curve.



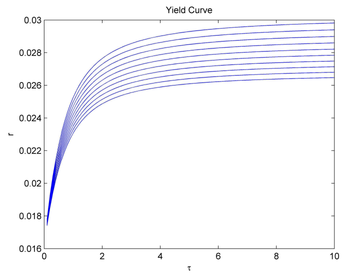

(A) $Q_{11}$




(B) $Q_{12}$




(C) $Q_{21}$




(D) $Q_{22}$

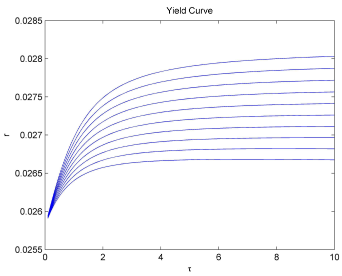
FIGURE 3. In this figure we show the effect of a perturbation of the single elements of the matrix $Q$. We use a sequence of numbers $\eta = 0.01 : 0.01 : 0.1$ and add $\eta_i$ to one of the elements of $Q$ while leaving the other elements unchanged. When we add $\eta_i$ to the elements on the main diagonal the YC is shifted upwards. The same happens with off-diagonal elements.

The impact that we obtain when we perturbate the elements of the matrix $Q$ is similar: we obtain again an upward shift. We would like to perform a more interesting experiment. To this end, we will consider now a larger (w.r.t. the partial order relation on $S_d^+$) value for $X_0$, more precisely:

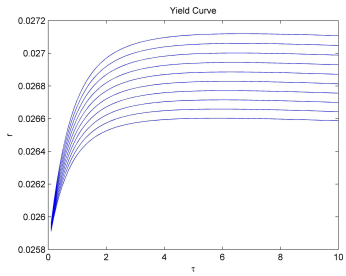
$$X_0 = \begin{pmatrix} 0.3780 & 0.0054 \\ 0.0054 & 1.2600 \end{pmatrix}. \tag{4.2}$$

With this starting value, we have that the yield curve is humped. We perform the same perturbations as before and notice that this time the impact is more varied. When we look at the impact of perturbations of $M$ we notice shapes ranging from nearly normal ($M_{11}$ and $M_{21}$) till humped shapes ($M_{12}$ and $M_{22}$). Anyhow we notice that the impact on the shape of the yield curve is quite strong.

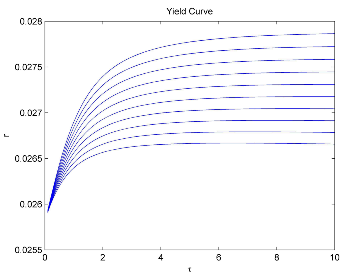
Finally, we work with $Q$. It turns out that the effect of the diffusion matrix is very relevant. Recall that with our starting value for the factor process, we have a humped curve. Figure (5) clearly shows that by performing perturbation on this matrix we are able to recover normal, inverse, or even humped curves. $Q$ seems to be best suited in determining large impacts on the shape of the yield curve, whereas $M$ seems to be suitable for smaller adjustments. Another fact that should be kept in mind, is that the interplay between the parameters is influenced by the different choice of the starting value of the process.



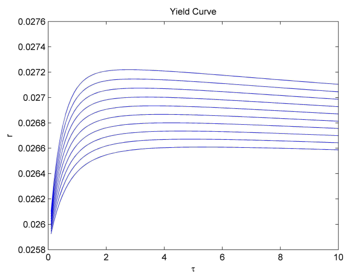

(A) $M_{11}$




(B) $M_{12}$

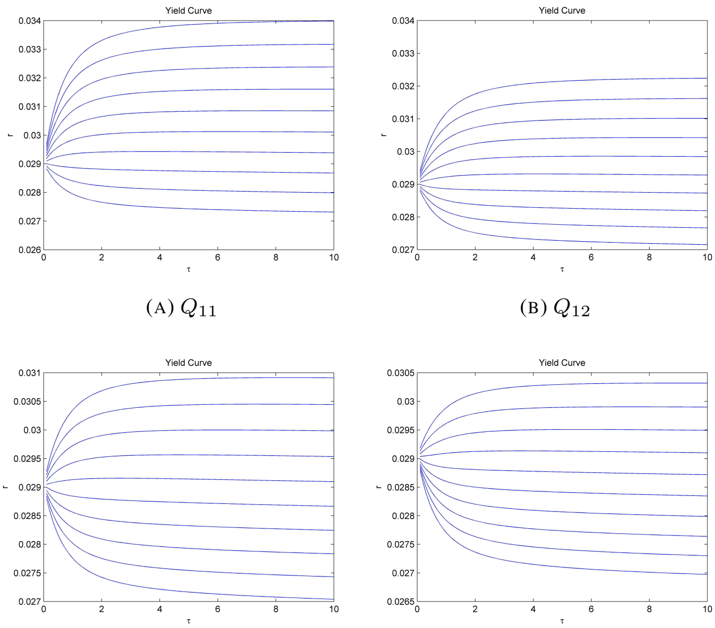



(C) $M_{21}$




(D) $M_{22}$

FIGURE 4. We increase the starting value of the process so as to get a humped yield curve. In this figure we show the effect of a perturbation of the single elements of the matrix $M$. We use a sequence of numbers $\eta = 0.01 : 0.01 : 0.1$ and add $\eta_i$ to one of the elements of $M$ while leaving the other elements unchanged. When we add $\eta_i$ to the elements on the main diagonal the YC is shifted upwards. The same happens with off-diagonal elements.




(A) $Q_{11}$ (B) $Q_{12}$

(C) $Q_{21}$ (D) $Q_{22}$

FIGURE 5. We increase the starting value of the process so as to get a humped yield curve. In this figure we show the effect of a perturbation of the single elements of the matrix $Q$. We use a sequence of numbers $\eta = 0.01 : 0.01 : 0.1$ and add $\eta_i$ to one of the elements of $Q$ while leaving the other elements unchanged. When we add $\eta_i$ to the elements on the main diagonal the YC is shifted upwards. The same happens with off-diagonal elements.

## 5. Conclusions

In this chapter we provided a new pricing formula for a zero coupon bond when the short rate is driven by a Wishart process. The formula turns out to be very explicit. Moreover, we proved a set of sufficient conditions ensuring that the Wishart short rate model produces certain yield curve shapes. In particular we are able to ensure that the yield curve is normal, inverse or humped. In the appendix it is also shown that when we consider a very simple Wishart process it is not possible to replicate inverted curves. We believe that this set of sufficient conditions may be of interest in a calibration setting, since we now know the constraints we should impose in order to replicate the shape of the yield curve that we observe on the market. A further line of research may be given by the study of yield curve shapes in the fully general model that was presented in Section 1.

## 6. Appendix: the Basic Diffusion Model

6.0.3. *Wishart Processes and the Cameron-Martin formula.* Now we introduce a simpler example of a diffusion process on $S_d^+$. This process was first analyzed by Bru (1991), (eq. 3.2). We assume here that $b$ is a scalar satisfying the relation $b \geq d+1$. $W_t$ will denote a matrix Brownian motion, i.e. a matrix whose entries are independent Brownian motions, and $\sqrt{X_t}$ will denote the square root in matrix sense. The dynamics are given by the following matrix SDE:

$$dX_t = bI_d dt + \sqrt{X_t}dW_t + dW_t^\top \sqrt{X_t}, \tag{6.1}$$

which corresponds to the dynamics (2.12) for $M = 0$ and $Q = I_d$. Under this modelling assumptions, the system of Riccati ODE's defining the price of the zero coupon bond looks as follows:

$$\frac{\partial \tilde{\phi}}{\partial \tau} = Tr\left[b\tilde{\psi}(\tau)\right] + a, \quad \tilde{\phi}(0) = 0, \tag{6.2}$$

$$\frac{\partial \tilde{\psi}}{\partial \tau} = -2\tilde{\psi}(\tau)^2 + B, \quad \tilde{\psi}(0) = 0. \tag{6.3}$$

We now report the pricing formula for a zero coupon bond in this setting, which constitutes a particular case of Proposition 2.2.

PROPOSITION 2.7. *Let the short rate be given as in* (1.1)*, for* $X \in WIS_d(x_0, b, 0, I_d)$*. Set* $\tau = T - t$*. Then the price of a zero coupon bond is:*

$$\mathbb{E}^{\mathbb{Q}}\left[\exp\left\{-Tr\left[\int_t^T a + BX_s ds\right]\right\}|\mathcal{F}_t\right] = e^{-\tilde{\phi}(\tau,B) - Tr\left[\tilde{\psi}(\tau,B)X_t\right]}, \tag{6.4}$$

*where:*

$$\tilde{\phi} = \frac{b}{2}\log\left(\det\cosh\sqrt{2B}\tau\right) + a\tau, \tag{6.5}$$

$$\tilde{\psi} = \frac{\sqrt{2B}}{2}\tanh\sqrt{2B}\tau. \tag{6.6}$$

Equipped with an explicit formula for the function $\tilde{\psi}$ and recalling (2.11), we immediately obtain the following:

LEMMA 2.9.

$$\lim_{\tau\to\infty}\tilde{\psi}(\tau, B) = \frac{\sqrt{2B}}{2}. \tag{6.7}$$

If we now recall the ODE satisfied by $\tilde{\psi}$, we realize that $\frac{\sqrt{2B}}{2}$ is a stability point of this ODE. So, in the basic diffusion model, the function $\tilde{\psi}$ tends asymptotically to the stability point of the Riccati ODE.

6.0.4. *Yield Curve Shapes in the Basic Model.* We repeat the analysis that we already performed for the general diffusion model. We will see that due to the absence of flexibility in the drift, it is not possible to replicate inverse yield curves.

THEOREM 2.3. *Consider a short rate model where the risk neutral dynamics of the short rate is driven by the process* $X_t$*, defined by the dynamics* (6.1)*. Let* $B \in S_d^{++}$*. Define:*

$$X^\star = \int_0^\infty e^{-\sqrt{2B}s} b e^{-\sqrt{2B}s} ds \tag{6.8}$$

*Then the following holds:*

- *The yield curve cannot be inverse.*
- *If* $X_t \prec X^\star$ *then we can reproduce a normal yield curve.*
- *If* $X_t \succ X^\star$ *then we can reproduce a humped yield curve.*

PROOF. In order to prove our claim, we follow step by step the approach in Keller-Ressel and Steiner (2008). First of all we define the function $\mathcal{H}(\tau) : \mathbb{R}_{\geq 0} \to \mathbb{R}$ by

$$\mathcal{H}(\tau) := Y(\tau, X_t)\tau = \tilde{\phi}(\tau) + Tr\left[\tilde{\psi}(\tau)X_t\right]. \tag{6.9}$$

The study of the second derivative of this function will allow us to derive the results for the yield curve. Recalling the system of equations satisfied by $\tilde{\phi}$ and $\tilde{\psi}$, given by equations (6.2) and (6.3), we can compute the derivatives of this function. For the first derivative we have:

$$\mathcal{H}'(\tau) = Tr\left[b\tilde{\psi}(\tau)\right] + a + Tr\left[\left(-2\tilde{\psi}(\tau)^2 + B\right)X_t\right], \tag{6.10}$$

whereas for the second we have:

$$\begin{aligned}\mathcal{H}''(\tau) &= Tr\left[b\mathcal{R}\left(\tilde{\psi}(\tau)\right)\right]\\ &+ Tr\left[\left(-2\tilde{\psi}(\tau)\mathcal{R}\left(\tilde{\psi}(\tau)\right) - 2\mathcal{R}\left(\tilde{\psi}(\tau)\right)\tilde{\psi}(\tau)\right)X_t\right].\end{aligned} \tag{6.11}$$

Since all matrices involved are symmetric, we may commute inside the trace operator, and then we obtain:

$$\mathcal{H}''(\tau) = Tr\left[\mathcal{R}\left(\tilde{\psi}(\tau)\right)\left(bI_d - 4\tilde{\psi}(\tau)X_t\right)\right]. \tag{6.12}$$

Define now $k(\tau) := bI_d - 4\tilde{\psi}(\tau)X_t$. The previous expression is equivalent to:

$$Tr\left[\mathcal{R}\left(\tilde{\psi}(\tau)\right)\frac{k(\tau) + k(\tau)^\top}{2}\right] = Tr\left[\mathcal{R}\left(\tilde{\psi}(\tau)\right)\tilde{k}(\tau)\right], \tag{6.13}$$

where $\tilde{k}(\tau) = bI_d - 2\tilde{\psi}(\tau)X_t - 2X_t\tilde{\psi}(\tau)$. We are interested in the convexity of the function $\mathcal{H}$, hence in the zeros of $\mathcal{H}''$. Due to the structure of the state space and the presence of only a partial order relation between matrices, using the arguments employed in the scalar case allows us to derive only sufficient conditions. A sufficient condition for $\mathcal{H}''(\tau) = 0$ is $\tilde{k}(\tau) = 0$. Now, we try to establish if the function has a zero for some value of $X_t$. We notice that:

$$\tilde{k}(0) = bI_d, \tag{6.14}$$

meaning that $\tilde{k}(0) \in S_d^{++} \forall X_t$. We will see that this implies that it is not possible to reproduce strictly decreasing i.e. inverted yield curves in the present setting.
We are also interested in the solution of the equation:

$$bI_d - 2\tilde{\psi}(\tau)X_t - 2X_t\tilde{\psi}(\tau) = 0. \tag{6.15}$$

Starting from this equation we would like to argue about different shapes of the yield curve as a function of $X_t$. Recall that $\tilde{\psi}(\tau) \nearrow \frac{1}{2}\sqrt{2B}$ as $\tau \to \infty$. This implies that we have exactly a zero at infinity if $X_t$ solves:

$$bI_d - \sqrt{2B}X_t - X_t\sqrt{2B} = 0. \tag{6.16}$$

By Theorem 2.1 the solution to this equation is given by:

$$X^\star = \int_0^\infty e^{-\sqrt{2B}s} b e^{-\sqrt{2B}s} ds. \tag{6.17}$$

We argue that this level is the critical level which ensures the presence of a zero of $\mathcal{H}''$. Let us now prove the following:

LEMMA 2.10. *$\tilde{k}(\tau)$ is monotonically decreasing, i.e., for $\tau_2 > \tau_1$, $\tilde{k}(\tau_2) - \tilde{k}(\tau_1) \in S_d^-$.*

PROOF. The argument is similar the the one employed for $\tilde{\psi}(\tau)$. We differentiate $\tilde{k}(\tau)$ so as to obtain:

$$\dot{\tilde{k}}(\tau) = -2\left(\mathcal{R}\left(\tilde{\psi}(\tau)\right)X_t + X_t\mathcal{R}\left(\tilde{\psi}(\tau)\right)\right). \tag{6.18}$$

Then we can write:

$$\begin{aligned}\tilde{k}(\tau_2) - \tilde{k}(\tau_1) &= \int_{\tau_1}^{\tau_2} \dot{\tilde{k}}(s)ds \\ &= -2\int_{\tau_1}^{\tau_2} \left(\mathcal{R}\left(\tilde{\psi}(\tau)\right)X_t + X_t\mathcal{R}\left(\tilde{\psi}(\tau)\right)\right)ds, \end{aligned}\tag{6.19}$$

which shows that the RHS is symmetric with negative eigenvalues, hence the claim. □

We will use the following result, which is proved e.g. in Zhang (1999):

LEMMA 2.11. *Let A,B be two hermitian matrices s.t.* $\lambda_i(A) \in [0,a]$ *and* $\lambda_i(B) \in [0,b]$*, then* $\lambda_i(AB) \in [0,ab]$.

This will be useful to prove the next claim:

LEMMA 2.12. *If* $X_t \succ X^\star$*, then* $\exists \tau^\star \in (0,\infty)$ *such that* $\mathcal{H}(\tau^\star)'' = 0$.

PROOF. Let $X_t = X^\star + \mathcal{C}$, for $\mathcal{C} \in S_d^+$, so that $X_t \succ X^\star$. We look at the asymptotic behavior of $\tilde{k}(\tau)$. We have:

$$\begin{aligned}\tilde{k}(\tau) &= bI_d - 2\tilde{\psi}(\tau)X_t - 2X_t\tilde{\psi}(\tau) \\ &= bI_d - 2\tilde{\psi}(\tau)X^\star - 2X^\star\tilde{\psi}(\tau) - 2\tilde{\psi}(\tau)\mathcal{C} - 2\mathcal{C}\tilde{\psi}(\tau),\end{aligned}\tag{6.20}$$

then we have:

$$\lim_{\tau\to\infty}\tilde{k}(\tau) = -2\left(\frac{\sqrt{2B}}{2}\right)\mathcal{C} - 2\mathcal{C}\left(\frac{\sqrt{2B}}{2}\right). \tag{6.21}$$

According to Lemma 2.11, this matrix has negative eigenvalues. Recall that $\tilde{k}(0) \in S_d^{++}$. Now we have that $\tilde{k}(\tau)$ is a continuous function of $\tau$, which by Lemma 2.10 is monotonically decreasing. This means that there must exist a $\tau'$ such that, for $\tau > \tau'$, the eigenvalues of $\tilde{k}(\tau)$ are negative.
Finally, we look at $\mathcal{H}''(\tau)$. Recall that:

$$\mathcal{H}''(\tau) = Tr\left[\mathcal{R}\left(\tilde{\psi}(\tau)\right)\tilde{k}(\tau)\right]. \tag{6.22}$$

By assumption, we have that $\mathcal{R}\left(\tilde{\psi}(\tau)\right) \in S_d^{++}$. We notice also that $\mathcal{H}''(\tau)$ is a continuous function of $\tau$. Furthermore, since (see (6.20)) $\tilde{k}(0) = bI_d \in S_d^{++}$ we have that $\mathcal{H}''(0) > 0$ as $Tr\left[\mathcal{R}\left(\tilde{\psi}(0)\right)\tilde{k}(0)\right] = bTr\left[BI_d\right]$. From the previous discussion on $\tilde{k}(\tau)$ we have that for $\tau > \tau'$ the second derivative is of the form:

$$Tr\left[\mathcal{R}\left(\tilde{\psi}(\tau)\right)(-\mathcal{K})\right], \tag{6.23}$$

where $\mathcal{K}$ is a symmetric matrix with positive eigenvalues. This means that the second derivative will be negative. By recalling the positiveness of the starting values and the continuity property w.r.t. $\tau$, thanks to the mean-value theorem, we can argue that there must exist a $\tau^\star$ s.t. $\mathcal{H}''(\tau^\star) = 0$ □

Along the same lines we get a second useful claim:

LEMMA 2.13. *If* $X_t \prec X^\star$*, then* $\nexists\, \tau^\star \in [0,\infty)$ *s.t.* $\mathcal{H}(\tau^\star)'' = 0$.

In our investigation for a sufficient condition for $\mathcal{H}''(\tau) = 0$, we obtained an existence result for a value $\tau^\star$ satisfying the condition. We are also able to show that this $\tau^\star$ is unique. Now, since $\tilde{k}(\tau)$ is monotonically decreasing, we obtain that if there exists a value for $\tau$ s.t. $\tilde{k}(\tau) = 0$, then this point in time must be unique.

By relying on lemmas 2.12, 2.13, 2.10, and the fact that $\mathcal{H}''(0) > 0$, we can argue the following:

- if $X_t \succ X^\star$, then $\mathcal{H}$ is strictly convex on $(0, \tau^\star)$ and strictly concave on $(\tau^\star, \infty)$
- if $X_t \prec X^\star$, then $\mathcal{H}$ is strictly convex on $(0, \infty)$
- the function $\mathcal{H}$ can not be strictly concave $\forall \tau \in [0, \infty)$: since $\mathcal{H}''$ is continuous and $\mathcal{H}''(0) > 0$, we have that $\exists \epsilon > 0$ s.t. $\mathcal{H}''(\tau) \geq 0$ for $\tau \in [0, \epsilon)$.

We use these findings on the convexity of $\mathcal{H}$ to determine our conclusions on the convexity of the yield curve. We consider the equation

$$\mathcal{H}(\tau) = c\tau, \qquad \tau \in [0, \infty), \tag{6.24}$$

for some fixed $c \in \mathbb{R}$. Since $\mathcal{H}(0) = 0$ (see (6.9)), this equation has at least one solution, i.e. $\tau_0 = 0$.

Now if $X_t \prec X^\star$, then $\mathcal{H}$ is strictly convex and then by Lemma 2.3 there exists at most one additional solution $\tau_1$ to (6.24) on $(0, \infty)$. If the solution exists, then $c\tau$ is crossed from below at $\tau_1$.

In the other case, i.e. if $X_t \succ X^\star$, we have that there exists a $\tau^\star$, the zero of $\tilde{k}(\tau)$, s.t. $\mathcal{H}$ is strictly convex on $(0, \tau^\star)$ and strictly concave on $(\tau^\star, \infty)$. Now there can exist at most two additional solutions $\tau_1$ and $\tau_2$ to (6.24), with $\tau_1 < \tau^\star < \tau_2$, s.t. $c\tau$ is crossed from below at $\tau_1$ and from above at $\tau_2$.

Finally, since $\mathcal{H}$ is never strictly concave, there exists no additional solution $\tau_1$ to (6.24) on $(0, \infty)$ s.t. $c\tau$ is crossed from above at $\tau_1$.

By definition of $\mathcal{H}$, we have that every solution to (6.24), excluding $\tau_0 = 0$, is also a solution to

$$Y(\tau, X_t) = c, \qquad \tau \in (0, \infty), \tag{6.25}$$

with $X_t$ fixed. The properties of crossing from above/below are preserved since $\tau$ is positive. This means that:

- if $X_t \prec X^\star$, then equation (6.25) has at most a single solution, or equivalently that every horizontal line, if it is crossed, is crossed from below by the yield curve. This shows that the yield curve $Y(\tau)$ is an increasing function of $\tau$, hence, according the terminology in Definition (2.3), the yield curve is normal.
- If $X_t \succ X^\star$, then equation (6.25) has at most two solutions $\tau_1$ and $\tau_2$, and if these two solutions are crossed then $\tau_1$ is crossed from below and $\tau_2$ is crossed from above, meaning that the yield curve is first increasing and then decreasing, i.e. the yield curve is humped.
- Since there exists no solution to (6.24) s.t. $c\tau$ is crossed from above, then there exists no solution to (6.25) s.t. every horizontal line is crossed from above. So the function $Y(\tau)$ is not strictly decreasing, i.e. the yield curve can not be inverse.

□

In summary, the results above show that the absence of the matrix $M$ implies a restriction in the set of yield curves since inverted curves can not be replicated by the model. The following images are a visualization of the results of the theorem:

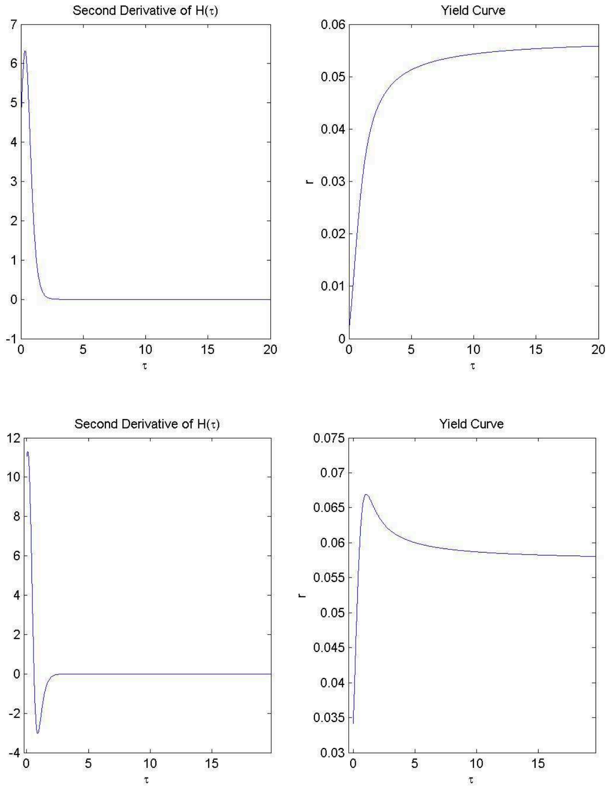



FIGURE 6. Yield curve shapes for different values of $X_t$ in the basic diffusion model.

CHAPTER 3

# A flexible matrix Libor model with smiles

## 1. Introduction

In this chapter we present a unified framework for the valuation of caps, floors and swaptions. These instruments are the most common derivative securities which are traded in a fixed income desk of a financial institution (see e.g. Brigo and Mercurio (2006)). Practitioners usually price these products by relying on Black-Scholes like formulae, which were first presented in Black (1976). The market convention of pricing caps and swaptions using the Black formula is based on an application of the Black and Scholes (1973) formula for stock options by assuming that the underlying interest rates are lognormally distributed. Remarkably, the use of this kind of formulae had no theoretical justification, since they involved a procedure in which the discount factor and the Libor rates were assumed to be independent in order to write the pricing formula as a product of a bond price and the expected payoff. The systematic use of this market practice ignited the interest of academics aiming at providing a coherent theoretical background.

In a series of articles, Miltersen et al. (1997), Brace et al. (1997), Jamshidian (1997) and Musiela and Rutkowski (1997), provided these theoretical foundations, introducing the Libor and Swap Market Model. Following these papers a stream of contributions appeared, trying to extend the basic model to the case where the volatility of the underlying factor is stochastic. The most famous proposals on this side can be found e.g. in Andersen and Brotherton-Ratcliffe (2001), Wu and Zhang (2006), Joshi and Rebonato (2003), Andersen and Andreasen (2002), Piterbarg (2005a), Piterbarg (2005b). Other approaches explored different dynamics for the driving process with respect to the CEV or displaced-diffusion considered before for the Libor rate: for example Glasserman and Kou (2003), Eberlein and Özkan (2005) introduced jump and more general Lévy processes, allowing for discountinuous sample paths of the driving process. Another interesting approach is the one of Brigo and Mercurio (2003) based on a mixture of lognormals. A common pitfall consists in the problematic form of the variance-covariance matrix of the yields, which makes it difficult to perform a PCA analysis.

A typical problem in the previous approaches is that once the closed form solution for cap prices is found, in order to obtain an analogous result for swaptions it is customary to assume that the underlying (which is a coupon bond) behaves like a scalar process (typically again geometric Brownian motion). This results in inconsistencies between the so-called *Libor* and *Swap Market Models*. Even more important, by assuming that the coupon bond is driven by a scalar process, we do not take into account the correlation effects among the different coupons, a key feature of a swaption which may be viewed also as a correlation product. This last remark is of paramount importance for practitioners (see e.g. the introduction of Collin-Dufresne and Goldstein (2002)).

In this chapter we consider a new approach based on the stochastic discount factor methodology, where instead of modeling directly the Libor rate, one concentrates on quotients of traded assets (i.e. bonds). It has been first introduced by Constantinides (1992) and then developed by Gouriéroux and Sufana (2011) in a spot interest rate framework and by Keller-Ressel et al. (2009) in a Libor perspective. They use affine processes on the state space $\mathbb{R}^d_{\geq 0}$ as driving processes and provide a full characterisation of the model, which allows them to provide closed form solutions for caps and swaptions up to Fourier integrals. This approach is very interesting and easily overcomes many difficulties which are to be faced in the computation of Radon-Nikodym derivatives.

We provide a straightforward extension of this approach, by considering affine processes on the state space $S_d^{++}$, the set of positive definite symmetric matrices. This state space may seem akward at first sight, but the processes belonging to this family admit a characterization in terms of ODE's which resembles the one found for standard affine models, an example being given by the famous Duffie and Kan (1996) model. In fact, in Cuchiero et al. (2009) the authors extend to the set $S_d^+$ (the set of positive semidefinite symmetric matrices) the classification of affine processes performed by Duffie et al. (2003) for the state space $\mathbb{R}^d_{\geq 0} \times \mathbb{R}^{n-d}$ introduced by Duffie and Kan (1996). What is more, the state space $S_d^{++}$ leads to stochastic factors which are non trivially correlated. The most famous example of process defined in the set $S_d^{++}$ is the Wishart process, originally introduced by Bru (1991) and then extensively applied in Finance by Gourieroux and Sufana (2003), Gourieroux and Sufana (2005), Gouriéroux and Sufana (2011), Da Fonseca et al. (2008), Da Fonseca et al. (2007b), Da Fonseca et al. (2009), Da Fonseca and Grasselli (2011), among the others.

The interesting feature of our framework is the possibility to obtain semi-closed form solutions for the pricing of swaptions in a multifactor setting, which is a well known challenging problem. In fact the exercise probability involves a multi-dimensional inequality. There have been many approaches to simplify the problem: for example, Singleton and Umantsev (2002) suggest an approximation of the exercise boundary with a linear function of the state variables. However, the most efficient approach seems to be the one of Collin-Dufresne and Goldstein (2002) which heavily uses the affine structure of the model and is based on the Edgeworth expansion for the characteristic function in terms of the cumulants. Since the cumulants decay very quickly the Edgeworth expansion for the exercise probability turns out to be very accurate and fast.

The chapter is organised as follows. In Section 2 we introduce our framework by recalling some useful definitions and results on affine processes. Section 3 investigates the case of the state space $S_d^{++}$ and presents the technical results. In Section 4 we focus on the pricing problem of the relevant derivatives. Caps and Floors are briefly treated since their pricing is now quite standard within the FFT methodology, while we devote more attention to the pricing of swaptions by adopting the approach of Collin-Dufresne and Goldstein (2002). Section 5 illustrates the flexibilty of our framework through a numerical exercise. Section 6 concludes the chapter, and we gather in the technical Appendices proofs and some remarks useful for implementation.

## 2. Affine Processes on the set $S_d^{++}$ of strictly positive definite symmetric matrices

**2.1. General results and notations.** To outline the setup we will consider affine processes taking values in the interior of the cone $S_d^+$. We will use the notations $\psi_t(u) = \phi(t,u)$ and $\phi_t(u) = \phi(t,u)$ so as to be consistent with Keller-Ressel et al. (2009). We will be employing a property of the functions defining the Laplace transform, that we report after the following

DEFINITION 3.1. *(Cuchiero et al. (2009), Definition 2.1) Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space with the filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfying the usual assumptions. A Markov process $\Sigma = (\Sigma_t)_{t\geq 0}$ with state space $S_d^+$, transition probability $p_t(\Sigma_0, A) = \mathbb{P}(\Sigma_t \in A)$ for $A \in S_d^+$, and transition semigroup $(P_t)_{t\geq 0}$ acting on bounded functions $f$ on $S_d^+$ is called affine process if:*

(1) *it is stochastically continuous, that is, $\lim_{s\to t} p_s(\Sigma_0, \cdot) = p_t(\Sigma_0, \cdot)$ weakly on $S_d^+$ $\forall t$, $x \in S_d^+$, and*
(2) *its Laplace transform has exponential-affine dependence on the initial state:*

$$P_t e^{-Tr[u\Sigma_0]} = \mathbb{E}\left[ e^{-Tr[u\Sigma_t]} \middle| \mathcal{F}_0 \right] = \int_{S_d^+} e^{-Tr[u\xi]} p_t(\Sigma_0, d\xi) = e^{-\phi_t(u) - Tr[\psi_t(u)\Sigma_0]}, \tag{2.1}$$

*$\forall t$ and $\Sigma_0, u \in S_d^+$, for some function $\phi : \mathbb{R}_{\geq 0} \times S_d^+ \to \mathbb{R}_{\geq 0}$ and $\psi : \mathbb{R}_{\geq 0} \times S_d^+ \to S_d^+$.*

Having applications in mind, we will consider affine processes which are *solvable* in the sense of Grasselli and Tebaldi (2008) (who investigated affine processes on the more general symmetric cone state space domain): this means that the state space that we will consider is the interior of $S_d^+$, namely the cone of *strictly positive definite* symmetric matrices, denoted by $S_d^{++}$[1]. Solvability is important, in fact it ensures that the Riccati Ordinary Differential Equation associated to the Laplace transform (2.1) through the usual Feynman-Kac argument has a regular globally integrable flow: this will be crucial in order to outline our methodology (see e.g. the proof of Theorem 3.1 in the sequel).

The next property closes our survey on affine processes. It will be needed when we prove that the structure of the model is preserved under changes of measure.

LEMMA 3.1. *(Cuchiero et al. (2009) Lemma 3.2) Let $\Sigma$ be an affine process on $S_d^+$, then the functions $\phi$ and $\psi$ satisfy the following property:*

$$\phi_{t+s}(u) = \phi_t(u) + \phi_s(\psi_t(u)),$$
$$\psi_{t+s}(u) = \psi_s(\psi_t(u)).$$

**2.2. Examples.** The previous general framework may be quite abstract at a first sight, mostly because of the high technical level required to properly introduce the notion of admissibility and existence for affine processes (see Chapter 2). In this subsection we provide some examples in order to show to the unfamiliar reader some concrete applications. We start with the most important one, which will also constitute our main object of study in the numerical illustrations.

2.2.1. *The Wishart process.* We suppose that the process $\Sigma$ is governed by the following (matrix) SDE:

$$d\Sigma_t = (\Omega\Omega^\top + M\Sigma_t + \Sigma_t M^\top)dt + \sqrt{\Sigma_t} dW_t Q + Q^\top dW_t^\top \sqrt{\Sigma_t}, \tag{2.2}$$

[1] By analogy, the set of negative (resp. strictly negative) definite symmetric $d \times d$ matrices will be denoted by $S_d^-$ (resp. $S_d^{--}$).

which was first studied by Bru (1991) and whose solution is known as Wishart process. We assume $M, Q$ invertible and $M$ negative definite in order to ensure stationarity of the process. Moreover we require $\Omega\Omega^\top = \kappa Q^\top Q$ for a real parameter $\kappa \geq d+1$ in order to grant solvability (or equivalently in order to grant that $Det(\Sigma_t) > 0$ with probability 1). Under the solvability assumption Grasselli and Tebaldi (2008) showed that the Riccati ODE corresponding to the characteristic function can be linearized and therefore admits a closed form solution. This is important in view of possible applications since in this case the functions $\phi$ and $\psi$ in definition (3.1) are explicitly known:

PROPOSITION 3.1. *Consider the process $\Sigma = (\Sigma_t)_{0\leq t\leq T}$ which solves the SDE* (2.2). *Then the conditional Laplace transform is given by:*

$$\mathbb{E}\left[e^{-Tr[u\Sigma_T]}\middle|\mathcal{F}_t\right] = e^{-\phi_\tau(u)-Tr[\psi_\tau(u)\Sigma_t]}, \tag{2.3}$$

*where $\tau := T-t$. The functions $\phi_\tau(u)$ and $\psi_\tau(u)$ satisfy the following system of ODE's:*

$$\frac{\partial\psi}{\partial\tau} = \psi_\tau(u)M + M^\top\psi_\tau(u) - 2\psi_\tau(u)Q^\top Q\psi_\tau(u), \quad \psi_0(u) = u, \tag{2.4}$$

$$\frac{\partial\phi}{\partial\tau} = Tr\left[\kappa Q^\top Q\psi_\tau(u)\right], \quad \phi_0(u) = 0 \tag{2.5}$$

*which is solved by*

$$\psi_\tau(u) = (u\psi_{12,\tau}(u) + \psi_{22,\tau}(u))^{-1}(u\psi_{11,\tau}(u) + \psi_{21,\tau}(u)), \tag{2.6}$$

*where*

$$\begin{pmatrix} \psi_{11,\tau}(u) & \psi_{12,\tau}(u) \\ \psi_{21,\tau}(u) & \psi_{22,\tau}(u) \end{pmatrix} = exp\left\{\tau\begin{pmatrix} M & 2Q^\top Q \\ 0 & -M^\top \end{pmatrix}\right\} \tag{2.7}$$

*and*

$$\phi_\tau(u) = \frac{\kappa}{2}Tr\left[log\left(u\psi_{12,\tau}(u) + \psi_{22,\tau}(u)\right) + M^\top\tau\right]. \tag{2.8}$$

PROOF. See Grasselli and Tebaldi (2008). □

Notice that the proposition above also reports the linearization approach in Chapter 1. The Wishart process constitutes the matrix analogue of the square root (Bessel) process. In fact we have that the matrix $M$ can be thought of as a mean reversion parameter: this is evident from the Lyapunov equation defining the long-run matrix $\Sigma_\infty$, which is given by

$$-\kappa Q^\top Q = M\Sigma_\infty + \Sigma_\infty M^\top. \tag{2.9}$$

The second way to appreciate the analogies w.r.t the square root process is to look at the dynamics of the entries of the matrix process $\Sigma$. Concentrating on the main diagonal, in the $2\times 2$ case we have:

$$\begin{aligned} d\Sigma_{11} &= \left(\kappa\left(Q_{11}^2 + Q_{21}^2\right) + 2\left(M_{11}\Sigma_{11} + M_{12}\Sigma_{12}\right)\right)dt \\ &\quad + 2\sigma_t^{11}\left(Q_{11}dW_t^{11} + Q_{21}dW_t^{12}\right) + 2\sigma_t^{12}\left(Q_{11}dW_{21} + Q_{21}dW_t^{22}\right) \end{aligned} \tag{2.10}$$

$$\begin{aligned} d\Sigma_{22} &= \left(\kappa\left(Q_{22}^2 + Q_{12}^2\right) + 2\left(M_{21}\Sigma_{12} + M_{22}\Sigma_{22}\right)\right)dt \\ &\quad + 2\sigma_t^{12}\left(Q_{12}dW_t^{11} + Q_{22}dW_t^{12}\right) + 2\sigma_t^{22}\left(Q_{12}dW^{21} + Q_{22}dW^{22}\right) \end{aligned} \tag{2.11}$$

where we set

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} := \sqrt{\Sigma}. \tag{2.12}$$

We refer to Da Fonseca et al. (2007a) for additional insights on the behavior of the Wishart process when aggregating its parameters.

2.2.2. *The pure jump OU process.* The procedure we adopt in this chapter is general, meaning that we can consider different examples of processes lying in the cone of positive definite matrices. In particular, we may consider the matrix subordinators proposed by Barndorff-Nielsen and Stelzer (2007), or jump-diffusions like in Leippold and Trojani (2010). In what follows we provide some examples with the calculations of the function $\phi_\tau$ and $\psi_\tau$.

Let us consider the SDE

$$d\Sigma_t = M\Sigma_t + \Sigma_t M^\top + dL_t, \tag{2.13}$$

where $M \in GL(d)$ is assumed as usual to be negative definite in order to grant stationarity, and $L_t$ is a pure jump process (compound Poisson Process) with constant intensity $\lambda$ and jump distribution $\nu$ with support on $S_d^{++}$. The strong solution to this equation is given by:

$$\Sigma_t = e^{Mt}\Sigma_0 e^{M^\top t} + \int_0^t e^{M(t-s)} dL_s e^{M^\top (t-s)}. \tag{2.14}$$

We are interested in the computation of the Laplace transform of this family of processes:

$$\mathbb{E}\left[ e^{Tr[u\Sigma_T]} \middle| \mathcal{F}_t \right] = e^{-\phi_\tau(u) - Tr[\psi_\tau(u)\Sigma_0]}. \tag{2.15}$$

The functions $\phi_\tau$ and $\psi_\tau$ solve the following (matrix) ODE's:

$$\frac{\partial \psi_\tau}{\partial \tau} = \psi_\tau(u)M + M^\top \psi_\tau(u) \quad \psi_0(u) = u \tag{2.16}$$

$$\frac{\partial \phi_\tau}{\partial \tau} = -\lambda \int_{S_d^+\setminus\{0\}} \left( e^{-Tr[\psi_\tau(u)\xi]} - 1 \right) \nu(d\xi) \quad \phi_0(u) = 0. \tag{2.17}$$

The solution for the first ODE is given by:

$$\psi_\tau(u) = e^{M^\top \tau} u e^{M\tau}, \tag{2.18}$$

so we can compute the Laplace transform by quadrature:

$$\frac{\partial \phi_\tau}{\partial \tau} = -\lambda \int_{S_d^+\setminus\{0\}} \left( e^{-Tr\left[ e^{M^\top s} u e^{Ms} \xi \right]} - 1 \right) \nu(d\xi). \tag{2.19}$$

In the following we provide explicit computations by assuming some particular distribution $\nu(\cdot)$ for the jump size. The proofs of this formulae may be found in Gupta and Nagar (2000). For the sake of clarity, we specify that the Wishart distribution that we consider in the next paragraphs are the classical distributions arising in the context of multivariate statistics, which differ from the generalization of the same distribution which is obtained when we consider the stochastic process in Chapter 1.

**Wishart Distribution.** Let $J$ be the jump size. Consider the case $J \sim Wis_d(n, \mathcal{Q})$. Then we have

$$\phi_\tau(u) = -\lambda \int_0^\tau \det\left( I_d + 2e^{M^\top s} u e^{Ms} \mathcal{Q} \right)^{-\frac{n}{2}} ds + \lambda\tau. \tag{2.20}$$

**Non-Central Wishart Distribution.** Let be $J \sim Wis_d(n, \mathcal{Q}, \mathcal{M})$, then we have

$$\begin{aligned}\phi_\tau(u) = &-\lambda \int_0^\tau \det(\mathcal{Q})^{-\frac{n}{2}} \det\left(2e^{M^\top s}ue^{Ms} + \mathcal{Q}^{-1}\right)^{-\frac{n}{2}} \times \\ &\exp\left\{Tr\left[-\frac{1}{2}\mathcal{Q}^{-1}\mathcal{M}\mathcal{M}^\top + \frac{1}{2}\mathcal{Q}^{-1}\mathcal{M}\mathcal{M}^\top\mathcal{Q}^{-1}\left(2e^{M^\top s}ue^{Ms} + \mathcal{Q}^{-1}\right)\right]\right\} ds \\ &+ \lambda\tau. \end{aligned} \tag{2.21}$$

**Beta type I distribution.** Let be $J \sim \beta_d^I(a, b)$, then

$$\phi_\tau(u) = -\lambda \int_0^\tau {}_1F_1(a; a+b; -e^{M^\top s}ue^{Ms})ds + \lambda\tau. \tag{2.22}$$

**Beta type II distribution.** Let be $J \sim \beta_d^{II}(a, b)$, then

$$\phi_\tau(u) = -\lambda \int_0^\tau \frac{\Gamma_d(a+b)}{\Gamma_d(b)} \Psi(a; -b + \frac{1}{2}(d+1); e^{M^\top s}ue^{Ms})ds + \lambda\tau, \tag{2.23}$$

where ${}_mF_n$, $\Gamma_d(a)$, and $\Psi(a; b; R)$ denote respectively the hypergeometric function of matrix argument, the multivariate Gamma function and the confluent hypergeometric function, see e.g. Gupta and Nagar (2000).

## 3. A Libor model on $S_d^{++}$

In order to outline the general framework for Libor models, we start by considering a filtered measurable space $(\Omega, \mathcal{F}, \mathcal{F}_t)$ and a family of probability measures $(\mathbb{P}_{T_k})_{1 \le k \le N}$. Under the measure $\mathbb{P}_{T_N}$ we introduce a stochastic process $\Sigma$ taking values on the cone state space $S_d^{++}$. At this stage the process may be a diffusion, a pure jump or a jump-diffusion process taking values on $S_d^{++}$. Consider a discrete tenor structure $0 = T_0 \le T_1 \le ... \le T_N = T$. We recall that the Libor rate is defined via quotients of bonds:

$$L(t, T_k) := \frac{1}{\delta}\left(\frac{B(t, T_{k-1})}{B(t, T_k)} - 1\right), \tag{3.1}$$

where $\delta$ is assumed to be constant and $\delta = T_k - T_{k-1}$. The relation between the Libor rate and the forward price is given by:

$$F(t, T_{k-1}, T_k) = 1 + \delta L(t, T_k). \tag{3.2}$$

We proceed in full analogy with Keller-Ressel et al. (2009) by extending their results to processes taking values on the cone of positive definite matrices. The intuition is simple: in order to build up a Libor model with positive rates, quotients of bonds should be strictly greater than one. On the other hand, a no-arbitrage argument (see e.g. Geman et al. (1995)) implies that quotients of bonds must be martingales under the forward risk neutral measure indexed by the maturity of the denominator, so that the key ingredient in the approach of Keller-Ressel et al. (2009) consists in the possibility of contructing a family of martingales that stay greater than one up to a bounded time horizon. This will be possible thanks to the affine structure of the model, since in this framework bond prices are exponentially affine in the positive (definite) factors, as well as their quotients.

**3.1. Martingales strictly greater than one.** Let us first define the set

$$\mathcal{I}_T := \left\{ u \in S_d : \mathbb{E}\left[e^{-Tr[u\Sigma_T]}\right] < \infty, \forall \Sigma_0 \in S_d^{++} \right\}.$$

By the affine property of the process $\Sigma$ we have

$$\mathbb{E}\left[e^{-Tr[u\Sigma_t]}\right] = e^{-\phi_t(u) - Tr[\psi_t(u)\Sigma_0]},$$

$$\phi : [0,T] \times \mathcal{I}_T \to \mathbb{R},$$

$$\psi : [0,T] \times \mathcal{I}_T \to S_d. \tag{3.3}$$

Within this setting we are able to construct martingales that stay greater than one up to a bounded time horizon $T$.

THEOREM 3.1. *Let $\Sigma$ be an affine process, and let $u \in \mathcal{I}_T \cap S_d^{--}$, then the process $M^u$ defined by*

$$M_t^u = exp\left\{-\phi_{T-t}(u) - Tr\left[\psi_{T-t}(u)\Sigma_t\right]\right\} \tag{3.4}$$

*is a martingale and $M_t^u > 1$ a.s. $\forall t \in [0,T]$.*

PROOF. See Appendix. □

Equipped with this positivity result, we can proceed by considering a tenor structure of non negative Libor rates $L(0,T_k)$ for $k = \{1,...,N-1\}$. Standard arbitrage arguments (see e.g. Geman et al. (1995)) imply that discounted traded assets, in our case bonds, are martingales under the terminal martingale measure:

$$\frac{B(*,T_k)}{B(*,T_N)} \in \mathcal{M}\left(\mathbb{P}_{T_N}\right) \quad \forall k \in \{1,...,N-1\}, \tag{3.5}$$

where $\mathcal{M}\left(\mathbb{P}_{T_N}\right)$ denotes the set of martingales with respect to the forward risk neutral probability $\mathbb{P}_{T_N}$. The idea in Keller-Ressel et al. (2009) is then to model quotients of bond prices using the martingales $M^u$ defined as follows:

$$\frac{B(t,T_1)}{B(t,T_N)} = M_t^{u_1} \tag{3.6}$$

$$\vdots$$

$$\frac{B(t,T_{N-1})}{B(t,T_N)} = M_t^{u_{N-1}} \tag{3.7}$$

$\forall t \in [0,T_1], ..., t \in [0,T_{N-1}]$ respectively. As a consequence, the initial values of the martingales $M_0^{u_k}$ must satisfy the relation

$$M_0^{u_k} = exp\left\{-\phi_T(u_k) - Tr\left[\psi_T(u_k)\Sigma_0\right]\right\} = \frac{B(0,T_k)}{B(0,T_N)}, \tag{3.8}$$

for all $k \in \{1,...,N-1\}$, so that it is possibile to set $u_N = 0$ as we have $M_0^{u_N} = 1$.

In the following proposition, we show that it is possible to fit (basically) any initial term structure of bond rates. The state space we are considering offers a wide range of possibilities to perform this task. However, since we are interested in applications, we adopt the simplest choice directly coming from the scalar case and we focus on the particular (but realistic) case where all Libor rates are positive.

PROPOSITION 3.2. *Let $L(0,T_1),...,L(0,T_N)$ be a tenor structure of positive initial Libor rates, and let $\Sigma$ be an affine process on $S_d^{++}$. Define*

$$\gamma_\Sigma := \sup_{u\in\mathcal{I}_T\cap S_d^{--}} \mathbb{E}\left[e^{-Tr[u\Sigma_T]}\right]. \tag{3.9}$$

*If $\gamma_\Sigma > \frac{B(0,T_1)}{B(0,T_N)}$ then there exists a strictly increasing sequence of matrices (i.e. $u_k \prec u_{k+1}$ if and only if $u_k - u_{k+1} \in S_d^{--}$) $u_1 \prec u_2 \prec ... \prec u_{N-1} \prec 0$ in $\mathcal{I}_T \cap S_d^{--}$ and $u_N = 0$ such that*

$$M_0^{u_k} = \frac{B(0,T_k)}{B(0,T_N)}, \qquad \forall k \in \{1,...,N\}. \tag{3.10}$$

*Conversely, let the bond prices be given by (3.6)-(3.7) and satisfy the initial condition (3.8). Then the Libor rates $L(t,T_k)$ are positive a.s. $\forall t \in [0,T_k]$ and $k \in \{1,...,N-1\}$.*

PROOF. See Appendix. □

**3.2. A fully-affine arbitrage-free model.** If we look at the definition of the Libor rate we realize that it is quite natural to require quotients of bonds to be driven by an exponentially affine function of the state: in fact, in this case also bond prices as well as forward prices will be affine functions. This is also in line with the previous approaches of Constantinides (1992) and Gouriéroux and Sufana (2011) based on the stochastic discount factor. In other words, the approach of Keller-Ressel et al. (2009) is able to provide a fully affine structure[2]. In order to prove the affine structure or our model, we first show that under (3.6)-(3.7), forward prices are of exponential-affine form under any forward measure. To do this, first we notice that in this framework quotients of bonds are exponentially affine in the state factors, so that also forward prices will be: for $k = 1,...,N-1$

$$\begin{aligned}\frac{B(t,T_k)}{B(t,T_{k+1})} &= \frac{B(t,T_k)}{B(t,T_N)}\frac{B(t,T_N)}{B(t,T_{k+1})} = \frac{M_t^{u_k}}{M_t^{u_{k+1}}}\\ &= exp\left\{-\phi_{T_N-t}(u_k) + \phi_{T_N-t}(u_{k+1})\right\}\\ &\quad exp\left\{Tr\left[\left(-\psi_{T_N-t}(u_k) + \psi_{T_N-t}(u_{k+1})\right)\Sigma_t\right]\right\}\\ &=: exp\left\{A_{T_N-t}(u_k,u_{k+1}) + Tr\left[B_{T_N-t}(u_k,u_{k+1})\Sigma_t\right]\right\}.\end{aligned} \tag{3.11}$$

With this result, we are able to show very easily that the model is arbitrage free, that is forward prices are martingales with respect to their corresponding forward measures (see Geman et al. (1995)):

$$\frac{B(*,T_k)}{B(*,T_N)} \in \mathcal{M}\left(\mathbb{P}_{T_N}\right). \tag{3.12}$$

This comes from the fact that forward measures are related one another via the quotients of the martingales $M^u$:

$$\frac{\partial \mathbb{P}_{T_k}}{\partial \mathbb{P}_{T_{k+1}}}|_{\mathcal{F}_t} = \frac{F(t,T_k,T_{k+1})}{F(0,T_k,T_{k+1})} = \frac{B(0,T_{k+1})}{B(0,T_k)}\frac{M_t^{u_k}}{M_t^{u_{k+1}}}, \tag{3.13}$$

$\forall k \in \{1,...,N\}$. Then $L(*,T_k)$ is a martingale under the forward measure $\mathbb{P}_{T_{k+1}}$ since the successive densities from $\mathbb{P}_{T_{k+1}}$ to $\mathbb{P}_{T_N}$ yield a telescoping product and $\mathbb{P}_{T_N}$ martingale (see Keller-Ressel et al. (2009)). More precisely:

$$1 + \delta L(*,T_k) = \frac{B(*,T_k)}{B(*,T_{k+1})} = \frac{M^{u_k}}{M^{u_{k+1}}} \in \mathcal{M}\left(\mathbb{P}_{T_{k+1}}\right) \tag{3.14}$$

[2]This is the reason why we will be able to apply the approach by Collin-Dufresne and Goldstein (2002), who originally started by an affine short rate in order to price swaptions: in fact, also in their framework bond prices are affine functions of the state variables.

since

$$\frac{M^{u_k}}{M^{u_{k+1}}}\prod_{l=k+1}^{N-1}\frac{M^{u_l}}{M^{u_{l+1}}}=M^{u_k}\in\mathcal{M}\left(\mathbb{P}_{T_N}\right). \tag{3.15}$$

Also, the density between the $\mathbb{P}_{T_k}$-forward measure and the terminal forward measure $\mathbb{P}_{T_N}$ is given by the martingale $M^{u_k}$ as indicated by (3.6)-(3.7):

$$\frac{\partial\mathbb{P}_{T_k}}{\partial\mathbb{P}_{T_N}}|_{\mathcal{F}_t}=\frac{B(0,T_N)}{B(0,T_k)}\frac{B(t,T_k)}{B(t,T_N)}=\frac{B(0,T_N)}{B(0,T_k)}M_t^{u_k}=\frac{M_t^{u_k}}{M_0^{u_k}}. \tag{3.16}$$

In this arbitrage-free model with positive Libor rates, the affine structure is preserved: that is, it is possible to extend to the state space $S_d^{++}$ the analogous result of Keller-Ressel et al. (2009).

PROPOSITION 3.3. *Let the bond structure be defined through (3.6)-(3.7), where the process $M^{u_\cdot}$ is given by (3.4). Then forward prices are exponentially affine in the state variable $\Sigma$ under any forward measure.*

PROOF. The result comes directly from formula (6.23) in Keller-Ressel et al. (2009) once the scalar product is replaced by the trace operator. □

## 4. Pricing of Derivatives

We now focus on the pricing problem for vanilla options like Caps&Floors and for exotic options like swaptions in the affine Libor model on $S_d^{++}$ introduced in the previous section. We shall see that the pricing of Caps and Floors may be performed using standard Fourier pricing techniques as in Keller-Ressel et al. (2009), whereas, for the case of swaptions, we will resort to a quasi closed form solution. In fact, since the moments of the underlying affine process are known through its characteristic function, we can expand the exercise probability via an Edgeworth developement, as shown in Collin-Dufresne and Goldstein (2002). This approach will lead to an efficient approximation: what is more, it will avoid the numerical problems underlying the computation of the exercise probability in Keller-Ressel et al. (2009).

**4.1. Caps and Floors.** A Cap may be thought of as a portfolio of call options on the successive Libor rates, named Caplets, whereas Floors are portfolios of put options named floorlets. These options are usually settled *in arrears*, which means that the caplet with maturity $T_k$ is settled at time $T_{k+1}$. The tenor lenght $\Delta T$ is assumed to be constant. Since the two products are equivalent, we will focus on Caps. A Cap with nominal capital $N=1$ has a payoff given by the following:

$$\Delta T\left(L(T_k,T_k)-K\right)^+ \qquad k=1,...,N-1 \tag{4.1}$$

We rewrite the payoff of caplets as in Keller-Ressel et al. (2009):

$$\begin{aligned}\Delta T\left(L(T_k,T_k)-K\right)^+ &= \left(1+\Delta TL(T_k,T_k)-(1+\Delta TK)\right)^+\\ &=\left(\frac{M_{T_k}^{u_k}}{M_{T_k}^{u_{k+1}}}-\mathcal{K}\right)^+,\end{aligned} \tag{4.2}$$

with $\mathcal{K}:=1+\Delta TK$.

Thus we see that the caplet is equivalent to an option on the forward price. In order to avoid the computation of expectations involving a joint distribution, each single caplet is priced under the corresponding forward measure:

$$\mathbb{C}\left(T_k, K\right) = B(0, T_{k+1})\mathbb{E}^{\mathbb{P}_{T_{k+1}}}\left[\left(\frac{M^{u_k}_{T_k}}{M^{u_{k+1}}_{T_k}} - \mathcal{K}\right)^+\right]$$

$$= B(0, T_{k+1})\mathbb{E}^{\mathbb{P}_{T_{k+1}}}\left[\left(e^Y - \mathcal{K}\right)^+\right], \tag{4.3}$$

with:

$$Y := log\left(\frac{M^{u_k}_{T_k}}{M^{u_{k+1}}_{T_k}}\right) = A_{T_N - T_k}(u_k, u_{k+1}) + Tr\left[B_{T_N - T_k}(u_k, u_{k+1})\Sigma_{T_k}\right], \tag{4.4}$$

for $A_{T_N-T_k}(u_k, u_{k+1})$, $B_{T_N-T_k}(u_k, u_{k+1})$ defined as in (3.11). The pricing problem can be solved via Fourier techniques through the Carr and Madan (1999) methodology. Hence we have the following proposition, whose standard proof is omitted.

PROPOSITION 3.4. *Let $\alpha > 0$. The price of a caplet with strike K and maturity $T_k$ is given by the formula:*

$$\mathbb{C}\left(T_k, K\right) = B(0, T_{k+1})\frac{exp\left\{-\alpha c\right\}}{2\pi}$$

$$\times \int_{-\infty}^{+\infty} e^{-ivc}\frac{\mathbb{E}^{\mathbb{P}_{T_{k+1}}}\left[e^{i(v-(\alpha+1)i)\left(A_{T_N-T_k}(u_k,u_{k+1})+Tr\left[B_{T_N-T_k}(u_k,u_{k+1})\Sigma_{T_k}\right]\right)}\right]}{\left(\alpha+iv\right)\left(1+\alpha+iv\right)}dv, \tag{4.5}$$

*where:*

$$c = log\left(1 + \Delta TK\right),$$

$$A_{T_N-T_k}(u_k, u_{k+1}) = -\phi_{T_N-T_k}(u_k) + \phi_{T_N-T_k}(u_{k+1}),$$

$$B_{T_N-T_k}(u_k, u_{k+1}) = -\psi_{T_N-T_k}(u_k) + \psi_{T_N-T_k}(u_{k+1}).$$

In other words, pricing a Cap involves the computation of the moment generating function of the Wishart process, which can be efficiently performed through the linearization of the associated Riccati ODEs as explained in Proposition 3.1. The parameter $\alpha > 0$ represents the damping factor introduced by the Carr and Madan (1999) methodology. We report in the Appendix B the explicit expression of the characteristic function involved in the pricing procedure.

**4.2. Swaptions.** The payoff of a receiver (resp. payer) swaption may be seen as a call (resp. put) on a coupon bond with strike price equal to one. We consider a receiver swaption that starts at $T_i$ with maturity $T_m$, $(i < m \leq N)$. The time-$T_i$ value is given by:

$$\mathbb{S}_{T_i}(K, T_i, T_m) = \left(\sum_{k=i+1}^{m} c_k B(T_i, T_k) - 1\right)^+ \tag{4.6}$$

where

$$c_k = \begin{cases} \Delta TK & \text{if } i+1 \leq k \leq m-1, \\ 1 + \Delta TK & \text{if } k = m. \end{cases} \tag{4.7}$$

Unfortunately, we face some difficulties if we try to adopt the Fourier technique that we employed to price a caplet. To see this we look at the proof of Proposition 7.2. in Keller-Ressel et al. (2009), which requires the computation of the Fourier transform of the payoff[3]:

(4.8)

$$\tilde{f}(z) = \int_{\mathbb{R}^{\frac{d(d+1)}{2}}} e^{Tr\left[iz\Sigma_{T_i}\right]} \left( \sum_{k=i+1}^{m} c_k e^{A_{T_N-T_i}(u_k,u_i)+Tr\left[B_{T_N-T_i}(u_k,u_i)\Sigma_{T_i}\right]} - 1 \right)^+ dvech(\Sigma_{T_i}),$$

where for a symmetric matrix $A$, $vech(A)$ stands for the vector in $\mathbb{R}^{d(d+1)/2}$ consisting in the columns of the upper-diagonal part of $A$ including the diagonal. The problem is given by the presence of the positive part in the payoff function. To get rid of it, we should be able to find a value $\tilde{\Sigma}$ such that

$$\sum_{k=i+1}^{m} c_k e^{A_{T_N-T_i}(u_k,u_i)+Tr\left[B_{T_N-T_i}(u_k,u_i)\tilde{\Sigma}\right]} = 1, \tag{4.9}$$

that is we should solve a single equation in $d(d+1)/2$ unknowns (the elements of $\tilde{\Sigma}$), which is highly non trivial when $d > 1$. Thus, pricing swaptions is challenging when we consider multiple factor affine models: this is a well known problem, see e.g. Jamshidian (1989) and Collin-Dufresne and Goldstein (2002). This is why Keller-Ressel et al. (2009) investigate the case $d = 1$, that is a Libor model driven by a (univariate) CIR process like in Jamshidian (1987). However, swaptions are essentially correlation products which are sensitive to changes in the movements of the yield curve, so that $d$ should be necessarily greater than one in order to take into account the multivariate nature of the yield curve.

In our approach we propose to follow the procedure suggested by Collin-Dufresne and Goldstein (2002) in order to approximate the exercise probabilities for the swaption. We define the $T_i$-price of a coupon bond, for $i < m \leq N$, as follows:

$$CB(T_i) = \sum_{k=i+1}^{m} c_k B(T_i, T_k). \tag{4.10}$$

Let us derive the general form of the pricing formula for a receiver swaption, for $0 = T_0 = t < T_i$:

$$\begin{aligned}
\mathbb{S}_0(K, T_i, T_m) &= \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_0^{T_i} r_s ds}\left(CB(T_i) - 1\right)^+\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_0^{T_i} r_s ds}\left(CB(T_i)\mathbf{1}_{(CB(T_i)>1)} - \mathbf{1}_{(CB(T_i)>1)}\right)\right] \\
&= \sum_{k=i+1}^{m} c_k \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_0^{T_k} r_s ds}\mathbf{1}_{(CB(T_i)>1)}\right] \\
&- \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_0^{T_i} r_s ds}\mathbf{1}_{(CB(T_i)>1)}\right].
\end{aligned}$$

[3] $B(T_i, T_k) = \frac{B(T_i,T_k)}{B(T_i,T_N)} \frac{B(T_i,T_N)}{B(T_i,T_i)} = \frac{M_{T_i}^{u_k}}{M_{T_i}^{u_i}} = \exp\left\{A_{T_N-T_i}(u_k, u_i) + Tr\left[B_{T_N-T_i}(u_k, u_i)\Sigma_{T_i}\right]\right\}$

We switch to the forward measure as follows:

$$\begin{aligned}
\mathbb{S}_0(K,T_i,T_N) &= \sum_{k=i+1}^{m} c_k B(0,T_k)\mathbb{E}^{\mathbb{Q}}\left[\frac{e^{-\int_0^{T_k} r_s ds}}{B(0,T_k)}\mathbf{1}_{(CB(T_i)>1)}\right] \\
&- B(0,T_i)\mathbb{E}^{\mathbb{Q}}\left[\frac{e^{-\int_0^{T_i} r_s ds}}{B(0,T_i)}\mathbf{1}_{(CB(T_i)>1)}\right] \\
&= \sum_{k=i+1}^{m} c_k B(0,T_k)\mathbb{E}^{\mathbb{P}_{T_k}}\left[\mathbf{1}_{(CB(T_i)>1)}\right] \\
&- B(0,T_i)\mathbb{E}^{\mathbb{P}_{T_i}}\left[\mathbf{1}_{(CB(T_i)>1)}\right] \\
&= \sum_{k=i+1}^{m} c_k B(0,T_k)\mathbb{P}_{T_k}\left[(CB(T_i)>1)\right] \\
&- B(0,T_i)\mathbb{P}_{T_i}\left[(CB(T_i)>1)\right].
\end{aligned}$$

The exercise probabilities $\mathbb{P}_{T_k}\left[(CB(T_i)>1)\right]$ and $\mathbb{P}_{T_i}\left[(CB(T_i)>1)\right]$ do not admit in general a closed form expression, so that we adapt to our setting the Edgeworth expansion procedure proposed by Collin-Dufresne and Goldstein (2002). Intuitively, the moments of the coupon bonds admit a simple closed-form expression in our affine framework, and these moments uniquely identify the cumulants of the distribution. One can expand the characteristic function in terms of the cumulants and compute the exercise probabilities by Fourier inversion.

Using the notation of Collin-Dufresne and Goldstein (2002) (formula (5)) for the $q-th$ power of a coupon bond we notice that, for $i<m\le N$:

$$\begin{aligned}
(CB(T_i))^q &= (c_{i+1}B(T_i,T_{i+1})+...+c_m B(T_i,T_m))^q \\
&= \sum_{j_1,\dots,j_q=i+1}^{m}\left(c_{j_1}\cdot...\cdot c_{j_q}\right)\times\left(B(T_i,T_{j_1})\cdot...\cdot B(T_i,T_{j_q})\right). \qquad (4.11)
\end{aligned}$$

Now in our framework we have (see also formula (7.9) in Keller-Ressel et al. (2009))

$$B(T_i,T_{j_l}) = \frac{M_{T_i}^{u_{j_l}}}{M_{T_i}^{u_i}} \qquad (4.12)$$

for $l=1,...,q$, meaning that we can rewrite the $q-th$ power of the coupon-bond as follows:

$$(CB(T_i))^q = \sum_{j_1,\dots,j_q=i+1}^{m}\left(c_{j_1}\cdot...\cdot c_{j_q}\right)\times\left(\frac{M_{T_i}^{u_{j_1}}}{M_{T_i}^{u_i}}\cdot...\cdot\frac{M_{T_i}^{u_{j_q}}}{M_{T_i}^{u_i}}\right). \qquad (4.13)$$

Recall, from (3.4), that we have

$$M_{T_i}^{u_{j_l}} = exp\left\{-\phi_{T_N-T_i}(u_{j_l}) - Tr\left[\psi_{T_N-T_i}(u_{j_l})\Sigma_{T_i}\right]\right\}, \qquad (4.14)$$

for $l=1,...,q$ and

$$M_{T_i}^{u_i} = exp\left\{-\phi_{T_N-T_i}(u_i) - Tr\left[\psi_{T_N-T_i}(u_i)\Sigma_{T_i}\right]\right\}. \qquad (4.15)$$

In conclusion, the $q-th$ moment under $\mathbb{P}_{T_k}$ has the following expression:

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}_{T_k}}\left[CB(T_i)^q\right]\\
&=\sum_{j_1,\dots,j_q=i+1}^{m}\left(c_{j_1}\cdot\ldots\cdot c_{j_q}\right)\times\mathbb{E}^{\mathbb{P}_{T_k}}\left[\left(\frac{M_{T_i}^{u_{j_1}}}{M_{T_i}^{u_i}}\cdot\ldots\cdot\frac{M_{T_i}^{u_{j_q}}}{M_{T_i}^{u_i}}\right)\right]\\
&=\sum_{j_1,\dots,j_q=i+1}^{m}\left(c_{j_1}\cdot\ldots\cdot c_{j_q}\right)\times\\
&\mathbb{E}^{\mathbb{P}_{T_k}}\left[exp\left\{\sum_{l=1}^{q}\Big(-\phi_{T_N-T_i}(u_{j_l})-Tr\left[\psi_{T_N-T_i}(u_{j_l})\Sigma_{T_i}\right]\Big)\right.\right.\\
&\left.\left.+q\Big(\phi_{T_N-T_i}(u_i)+Tr\left[\psi_{T_N-T_i}(u_i)\Sigma_{T_i}\right]\Big)\right\}\right]\\
&=\sum_{j_1,\dots,j_q=i+1}^{m}\left(c_{j_1}\cdot\ldots\cdot c_{j_q}\right)\times exp\left\{\left(-\sum_{l=1}^{q}\phi_{T_N-T_i}(u_{j_l})\right)+q\phi_{T_N-T_i}(u_i)\right\}\\
&\times\mathbb{E}^{\mathbb{P}_{T_k}}\left[exp\left\{Tr\left[\left(\left(-\sum_{l=1}^{q}\psi_{T_N-T_i}(u_{j_l})\right)+q\psi_{T_N-T_i}(u_i)\right)\Sigma_{T_i}\right]\right\}\right],
\end{aligned}\tag{4.16}$$

where the functions $\phi$ and $\psi$ are as usual the solutions of Riccati ODE's of the form (2.4), (2.5). Once the first $m$ moments under the corresponding forward measures are exactly determined, we can estimate the exercise probabilities $\mathbb{P}_{T_k}\left[(CB(T_0)>1)\right]$ under each forward measure as in Collin-Dufresne and Goldstein (2002), by relying on a cumulant expansion on $\mathbb{P}_{T_k}\left[CB(T_0)\right]$.

## 5. The Wishart Libor Model

The aim of this section is to illustrate a specific choice for the driving process $\Sigma$. As in the general setup, we specify the process under the terminal probability measure $\mathbb{P}_{T_N}$. The example we choose is the Wishart process, which was already presented in section 2.2.1:

$$d\Sigma_t=(\Omega\Omega^\top+M\Sigma_t+\Sigma_tM^\top)dt+\sqrt{\Sigma_t}dW_t^{T_N}Q+Q^\top dW_t^{T_N\top}\sqrt{\Sigma_t}.\tag{5.1}$$

Here $W_t^{T_N}$ denotes a matrix Brownian motion, i.e. a $d\times d$ matrix of independent Brownian motions under the $\mathbb{P}_N$-forward probability measure. In the sequel we will write $W_t$ for notational simplicity.
In this section we show the impact of the relevant parameters on the implied volatility surface generated by vanilla options for a Libor model driven by a Wishart process. In order to investigate some complex movements of the implied volatility surface, we first compute the covariation between the Libor rate and its volatility: this covariation is a crucial quantity allowing for the so called skew effect on the smile, in perfect analogy with the leverage effect for vanilla options in the equity market.

**5.1. The skew of vanilla options.** In order to compute the covariation between the Libor rate and its volatility, we need to derive the dynamics of the Libor rate in the Wishart model. This may be done along the following steps: using the shorthand

$$B_k:=B_{T_N-t}(u_k,u_{k+1})=-\psi_{T_N-t}(u_k)+\psi_{T_N-t}(u_{k+1}),\tag{5.2}$$

recall that we have:

$$1 + \delta L(t, T_k, T_{k+1}) = \frac{B(t, T_k)}{B(t, T_{k+1})} = e^{A_k + Tr[B_k \Sigma_t]}. \tag{5.3}$$

In differential form, after dividing both sides by $L(t, T_k, T_{k+1})$ we have

$$\begin{aligned} &\frac{dL(t, T_k, T_{k+1})}{L(t, T_k, T_{k+1})} \\ &= \frac{1 + \delta L(t, T_k, T_{k+1})}{L(t, T_k, T_{k+1})} \left( [...]dt + Tr\left[ B_k d\Sigma_t \right] \right). \end{aligned} \tag{5.4}$$

To preserve analytical tractability, we freeze the coefficients and approximate as follows:

$$\frac{1 + \delta L(t, T_k, T_{k+1})}{L(t, T_k, T_{k+1})} \approx \frac{1 + \delta L(0, T_k, T_{k+1})}{L(0, T_k, T_{k+1})} =: C. \tag{5.5}$$

PROPOSITION 3.5. *Under the assumption of frozen coefficients (5.5), the conditional infinitesimal correlation between the Libor rate and its volatility cannot be negative and is given by*

$$\begin{aligned} &d \left\langle L(t, T_k, T_{k+1}), vol(L(t, T_k, T_{k+1})) \right\rangle \\ &= \frac{Tr\left[ B_k Q^\top Q B_k Q^\top Q B_k \Sigma \right] dt}{\sqrt{Tr\left[ Q B_k \Sigma B_k^\top Q^\top \right]} \sqrt{Tr\left[ \Sigma B_k Q^\top Q B_k Q^\top Q B_k Q^\top Q B_k \right]}}. \end{aligned} \tag{5.6}$$

PROOF. See Appendix. □

From the previous formula we realize that the matrix $Q$ is responsible for the shape of the skew. We also have an indirect impact of the mean reversion speed matrix $M$ coming from the term $B_k$ which is the difference of two solutions of the Riccati ODE's (2.4) and (2.5). The presence of $\Sigma$ suggests that in the present framework the skew is stochastic. What is more, it can only have positive sign.

**5.2. Numerical illustration with diagonal parameters.** The dynamics above show that the Wishart specification provides a very rich structure of the model. In order to get an understanding of the impact of different parameters we will look first at the case where all matrices are diagonal, which basically corresponds to a model driven by a two factor square root process (see e.g. Da Fonseca and Grasselli (2011)).

We use the following set of parameters as a benchmark:

$$\Sigma_0 = \begin{pmatrix} 3.75 & 0 \\ 0 & 3.45 \end{pmatrix}, \quad M = \begin{pmatrix} -0.3125 * 1.0e-003 & 0 \\ 0 & -0.5000 * 1.0e-003 \end{pmatrix},$$
$$Q = \begin{pmatrix} 0.034 & 0 \\ 0 & 0.0420 \end{pmatrix}, \quad \beta = 3.$$

The impact of the Gindikin parameter $\kappa$ is quite easy to understand: the process acts by influencing the overall level of the surface. This is due to the fact that the higher $\beta$ the lower the probability that the process $\Sigma$ approaches 0. It is interesting to note that there is not only a level impact, but also a curvature effect, as we can see in Figure (1).

Figure (1) here

Let us now look at the parameters along the diagonals of the matrices $M$ and $Q$. The following claims may be easily checked by looking at the SDE's satisfied by the elements of $\Sigma$ (see also Da Fonseca et al. (2007a)). Note that we assumed all eigenvalues of $M$ to lie in the negative real line.

- For $|M_{11}| \nearrow (\searrow)$ the surface is shifted downwards (upwards).
- For $|M_{22}| \nearrow (\searrow)$ the surface is shifted downwards (upwards).

The impact is more evident for OTM caplets with short maturities. This is due to the fact that as the process decreases (in matrix sense) the probability that caplets with short maturities are exercised is lowered more than the analogous probability for longer term caplets.

Figure (2) here

We then consider the impact of $Q_{11}, Q_{22}$. We have the following:

- As $Q_{11} \nearrow (\searrow)$ the surface is shifted upwards (downwards). In particular if we multiply $Q_{22}$ by a constant $c > 1$, then the increment in the short term is higher for OTM than for ITM caplets. If $c < 1$ then the decrease is higher for short term OTM caplets, which is intuitive, given the discussion above.
- The same impacts, with different magnitudes, is observed also for $Q_{22}$.

Figure (3) here

**5.3. The term structure of ATM implied volatilities for caplets.**

5.3.1. *Diagonal parameters.* We proceed to consider the term structure of caplet implied volatilities. When the matrix $\Sigma_0$ is diagonal, the impact of the elements of $Q$ is the same: an increase in the absolute value of any element of $Q$ will result in a steeper term structure of ATM caplet volatilities.

Figure (4) here

Considering a model where $\Sigma_0$ is a full matrix does not influence this result in a significant way.

5.3.2. *More complex adjustments: impact of off-diagonal elements.* In order to appreciate the flexibility of the Wishart framework, we focus now on the impact of the off-diagonal elements. We introduce off-diagonal elements in $M$ and $Q$ and look at the relative change in the short term smile (4 months) and the long term smile (32 months). We introduce a fully populated matrix $\Sigma_0$ and look at the impact of $M_{12}$ and $M_{21}$. Our experiments show that there is a symmetry between the sign of $\Sigma_{0,12}$ (the initial value of $\Sigma_{12}$) and $M_{12}$, $M_{21}$. More precisely, the implied volatility changes are as in Table 1.

| | $\Sigma_{0,12} > 0$ | $\Sigma_{0,12} < 0$ |
|---|---|---|
| $M_{12} > 0$ | Increase | Decrease |
| $M_{12} < 0$ | Decrease | Increase |
| $M_{21} > 0$ | Increase | Decrease |
| $M_{21} < 0$ | Decrease | Increase |

TABLE 1. Implied volatility changes: relation between $\Sigma_{0,12}$ and $M_{12}, M_{21}$.

The reason for this symmetry is to be looked for in the drift part of the dynamics of the single elements of the matrix process $\Sigma$.

Next we look at the impact of $Q_{12}$ and $Q_{21}$. To this end we model $Q$ as a symmetric matrix and set $Q_{21} = Q_{12} = \rho\sqrt{Q_{11}Q_{22}}$ for a real parameter $\rho$. Also in this case we recognize two main shapes of the adjustment that we denote by $S_1, S_2$.

| | $\Sigma_{0,12} > 0$ | $\Sigma_{0,12} < 0$ |
|---|---|---|
| $\rho > 0$ | $S_1$ | $S_2$ |
| $\rho < 0$ | $S_2$ | $S_1$ |

TABLE 2. Implied volatility changes: relation between $\Sigma_{0,12}$ and $\rho$.

We now proceed to perform other numerical tests which will show that our modelling framework has a certain degree of flexibility. For these tests we set:

$$M = \begin{pmatrix} -0.3125 * 1.0e - 003 & 0 \\ 0 & -0.5000 * 1.0e - 003 \end{pmatrix},$$

$$Q = \begin{pmatrix} 0.02 & \rho\sqrt{Q_{11}Q_{22}} \\ \rho\sqrt{Q_{11}Q_{22}} & 0.0420 \end{pmatrix}, \quad \beta = 3,$$

so basically $M$ is parametrized as before but $Q$ is symmetric and equiped with a parameter $\rho$ which summarizes the information on the off-diagonal elements. We require $\Sigma_0 = \Sigma_\infty$, where $\Sigma_\infty$ is given by the solution of the Lyapunov equation (2.9). After that we perturbate $\Sigma_0$ in order to include off-diagonal elements and set $\Sigma_{0,12} = \Sigma_{0,21} = 2$. We have a good degree of control on the term structure of ATM implied volatilities. In particular, we may have larger percentage shifts in the long-term w.r.t. the short-term ATM implied volatility, or, for $\rho = -0.6$ we may even reproduce a situation where the short term ATM implied volatility increases whereas the long-term ATM implied volatility decreases.

Figure (5) here

If we adopt the same kind of parametrization for the matrix $M$ by introducing a second parameter $\rho_2$, then we have further flexibility because we can impose many different combinations of $\rho, \rho_2$. For example, Figure (9) shows that we are able to isolate an effect on the term structure of ATM implied volatility: in fact we have a moderate change for ITM caplets while OTM caplets are practically unchanged, but the shape of the term structure of ATM implied volatility is modified in a significant way.

Figure (9) here

Finally, just for illustrative purposes we report a prototypical Caplet volatility surface generated by the model.

Figure (7) here

As far as Swaptions are concerned an example of ATM implied volatility surface for different expiries and underlying swap lengths is given below.

Figure (8) here

## 6. The Pure Jump Libor Model

Finally, in this section, we would like to provide a second example for the driving process $\Sigma$, so as to let the reader appreciate the degree of generality of this framework. As in the general setup, we specify the process under the terminal probability measure $\mathbb{P}_{T_N}$. The example we choose is a matrix compound Poisson process, which was already presented in section 2.2.2:

$$d\Sigma_t = M\Sigma_t + \Sigma_t M^\top + dL_t^{\mathbb{P}_{T_N}}, \tag{6.1}$$

All assumptions presented in section 2.2.2 are in order. More specifically, we assume that $L_t^{\mathbb{P}_{T_N}}$ is a compound Poisson process with constant intensity $\lambda$ and jump distribution haveing support in $S_d^{++}$. As a specific example of jump distribution we choose the standard Wishart distribution. By recalling the results in section 2.2.2 we have that the solution for the function $\psi_\tau(u)$ is

$$\psi_\tau(u) = e^{M^\top \tau} u e^{M\tau}, \tag{6.2}$$

whereas for $\phi_\tau(u)$ we have

$$\phi_\tau(u) = -\lambda \int_0^\tau \det\left(I_d + 2e^{M^\top s} u e^{Ms} \mathcal{Q}\right)^{-\frac{n}{2}} ds + \lambda\tau. \tag{6.3}$$

In concrete pricing applications, the computation of the solution for $\phi_\tau(u)$ implies a numerical integration with respect to the time dimension. This numerical integration has an impact on the performance of the model which turns out to be slower than the Wishart Libor model. For illustrative purposes, we report an example for an implied volatility surface for caplets generated by the compound Poisson Libor model with central Wishart distributed jumps. The mean reversion matrix $M$ and the jump intensity $\lambda$ are given by:

$$M = \begin{pmatrix} -0.0550 & 0 \\ 0 & -0.1760 \end{pmatrix},$$
$$\lambda = 0.1.$$

As far as the jump size distribution is concerned, the parameters are the following:

$$\mathcal{Q} = \begin{pmatrix} 0.27 & 0 \\ 0 & 0.05 \end{pmatrix},$$
$$n = 3.1$$

Finally, the initial state of the process is

$$\Sigma_0 = \begin{pmatrix} 1.875 & 0.6 \\ 0.6 & 1.275 \end{pmatrix}. \tag{6.4}$$

## 7. Conclusions

In this chapter we presented an extension of the approach of Keller-Ressel et al. (2009) to the more general setting of affine processes on positive definite matrices. We showed that their methodology may be adapted to this state space in a straightforward way. What is more, it is possible to efficiently price European swaptions in this multi-factor setting by means of a cumulant expansion due to Collin-Dufresne and Goldstein (2002). In doing so we are in front of a setting which is potentially able to capture correlation effects which can not be described by a single-factor setting. We provided numerical examples for the Wishart Libor model, where the introduction of off-diagonal elements gives rise to new possibilities in the control of the shape of the implied volatility surface. Our contribution may be seen as a starting point for a description of market models in this state space, in consequence we believe that there are many possible directions for future research. An example is given by the problem of calibrating this family of models to real market data. As the structure of the products in the fixed-income market suggests, even in the plain vanilla case, we expect the objective function that should be minimized in the calibration procedure to be quite involved. It will definitely lead to highly non trivial issues when dealing with the *implementation risk* in the spirit of Da Fonseca and Grasselli (2011). For this reason the calibration of our model, being a delicate issue, may constitute an interesting contribution by its own. Once the model is calibrated on vanillas, one could then investigate the performance of the model on more exotic structures, like e.g. Bermudan swaptions and barrier options.

## 8. Appendix A: proofs

**8.1. Proof of Theorem 3.1.** For all $u \in \mathcal{I}_T$ we have

$$\mathbb{E}\left[M_T^u\right] = \mathbb{E}\left[e^{-Tr[u\Sigma_T]}\right] < \infty,$$

and by the affine property we obtain

$$\begin{aligned}\mathbb{E}\left[M_T^u|\mathcal{F}_t\right] &= \mathbb{E}\left[exp\left\{-\phi_{T-T}(u) - Tr\left[\psi_{T-T}(u)\Sigma_T\right]\right\}|\mathcal{F}_t\right]\\ &= \mathbb{E}\left[exp\left\{-Tr\left[u\Sigma_T\right]\right\}|\mathcal{F}_t\right]\\ &= exp\left\{-\phi_{T-t}(u) - Tr\left[\psi_{T-t}(u)\Sigma_t\right]\right\} = M_t^u,\end{aligned}$$

hence the process is a martingale. Now we show that $M_t^u > 1$. Recall that by assumption $u \in \mathcal{I}_T \cap S_d^{--}$ and

$$M_t^u = \mathbb{E}\left[exp\left\{-Tr\left[u\Sigma_T\right]\right\}|\mathcal{F}_t\right],$$

so that if $-Tr\left[u\Sigma_T\right] > 0$ a.s. then we are done. We proceed as in Gourieroux and Sufana (2003) and apply the singular value decomposition to the negative definite matrix $u$, i.e. $u$ may be written as:

$$u = \sum_{i=1}^{n} \lambda_i u_i u_i^\top$$

where $\lambda_i$ are the eigenvalues of $u$ and $u_i$ are the eigenvectors. By assumption $\Sigma_T$ takes values in $S_d^{++}$, hence

$$\begin{aligned}-Tr\left[u\Sigma_T\right] &= -Tr\left[\sum_{i=1}^{n} \lambda_i u_i u_i^\top \Sigma_T\right]\\ &= -\sum_{i=1}^{n} \lambda_i Tr\left[u_i u_i^\top \Sigma_T\right]\\ &= -\sum_{i=1}^{n} \lambda_i u_i^\top \Sigma_T u_i > 0 \qquad (8.1)\end{aligned}$$

as we wanted.

**8.2. Proof of Proposition 3.2.** We follow closely the proof in Keller-Ressel et al. (2009). By assumption, initial Libor rates are strictly positive, then

$$\frac{B(0,T_1)}{B(0,T_N)} > \frac{B(0,T_2)}{B(0,T_N)} > ... > \frac{B(0,T_N)}{B(0,T_N)} = 1. \qquad (8.2)$$

Recall that we have

$$\mathbb{E}\left[e^{-Tr[u_1\Sigma_T]}\right] = M_0^{u_1} = exp\left\{-\phi_T(u_1) - Tr\left[\psi_T(u_1)\Sigma_0\right]\right\} = \frac{B(0,T_1)}{B(0,T_N)}. \qquad (8.3)$$

By the definition of $\gamma_\Sigma$ in (3.9), we have that if $\gamma_\Sigma = \infty$ then we are done, else we can claim that there exists an $\epsilon > 0$ such that $\gamma_\Sigma - \epsilon > \frac{B(0,T_1)}{B(0,T_N)}$. Then we can find a matrix $\tilde{u}$ s.t.

$$\mathbb{E}\left[e^{-Tr[\tilde{u}\Sigma_T]}\right] > \gamma_\Sigma - \epsilon > \frac{B(0,T_1)}{B(0,T_N)}. \tag{8.4}$$

In analogy with Keller-Ressel et al. (2009) we introduce the function

$$\begin{aligned} f &: [0,1] \to \mathbb{R}_{\geq 0} \\ \xi &\to \mathbb{E}\left[e^{-Tr[\xi\tilde{u}\Sigma_T]}\right] \end{aligned} \tag{8.5}$$

and we want to show that $f$ is continuous. First, since $\Sigma \in S_d^{++}$ and $u \in S_d^{--}$ we have that if $u \prec v$ then $-Tr\left[u\Sigma_T\right] > -Tr\left[v\Sigma_T\right]$, hence by monotone convergence we can conclude that $f$ is increasing. We now introduce an increasing sequence $(a_n)_{n\in\mathbb{N}} \nearrow 1$ and apply Fatou's lemma to obtain

$$\liminf_{n\to\infty} \mathbb{E}\left[e^{-Tr[a_n\tilde{u}\Sigma_T]}\right] \geq \mathbb{E}\left[\liminf_{n\to\infty} e^{-Tr[a_n\tilde{u}\Sigma_T]}\right] = \mathbb{E}\left[e^{-Tr[\tilde{u}\Sigma_T]}\right],$$

implying that $f$ is lower semi-continuous. Since $f$ is also increasing we have that $f$ is continuous. Now $f(0) = 1$ and $f(1) > \frac{B(0,T_1)}{B(0,T_N)}$, hence there exist some numbers $0 = \xi_N < \xi_{N-1} < ... < \xi_1 < 1$ such that

$$f\left(\xi_k\right) = M_0^{\xi_k\tilde{u}} = \frac{B(0,T_k)}{B(0,T_N)}, \qquad \forall k \in \{1, ..., N\}.$$

By setting $u_k = \xi_k\tilde{u}$ (for $k = 1, ..., N-1$) we obtain a sequence of matrices $u_k \prec u_{k+1}, u_k - u_{k+1} \in S_d^{--}$ which allows us to fit the initial tenor structure of Libor rates as desired. Finally, we apply Proposition 1 and Lemma 3.2 (ii) in Cuchiero et al. (2009) in order to obtain the last sentence of the Proposition 3.2.

**8.3. Proof of Proposition 3.5.** In this section we proceed as in the proof of Proposition 4.1 in Da Fonseca et al. (2008). Recall that $W_t$ is a shorthand for $W_t^{T_N}$. From (5.4) it follows that

$$\begin{aligned} \frac{dL(t,T_k,T_{k+1})}{L(t,T_k,T_{k+1})} &= C\left((...)dt + 2\sqrt{Tr\left[QB_k\Sigma B_k^\top Q^\top\right]}\left(\frac{Tr\left[QB_k\sqrt{\Sigma}dW_t\right]}{\sqrt{Tr\left[QB_k\Sigma B_k^\top Q^\top\right]}}\right)\right) \\ &:= C\left((...)dt + 2\sqrt{Tr\left[QB_k\Sigma B_k^\top Q^\top\right]}d\tilde{W}_t\right), \end{aligned} \tag{8.6}$$

where $C$ was defined in (5.5) and the scalar noise driving the factor process may be derived as follows:

$$\begin{aligned} dTr\left[QB_k\Sigma_tB_kQ^\top\right] &= \left(Tr\left[QB_k\beta Q^\top QB_kQ^\top\right] + 2Tr\left[QB_kM\Sigma_tB_kQ^\top\right]\right)dt \\ &+ 2Tr\left[QB_k\sqrt{\Sigma_t}dW_tQB_kQ^\top\right] \\ &= (...)dt + 2\sqrt{Tr\left[\Sigma_tB_kQ^\top QB_kQ^\top QB_kQ^\top QB_k\right]}\frac{Tr\left[QB_kQ^\top QB_k\sqrt{\Sigma_t}dW_t\right]}{\sqrt{Tr\left[\Sigma B_kQ^\top QB_kQ^\top QB_kQ^\top QB_k\right]}} \\ &:= (...)dt + 2\sqrt{Tr\left[\Sigma_tB_kQ^\top QB_kQ^\top QB_kQ^\top QB_k\right]}dZ_t. \end{aligned} \tag{8.7}$$

The covariation between the noise of the Libor rate and its volatility is then given by

$$
\begin{aligned}
\left\langle d\tilde{W}_t, dZ_t \right\rangle &= \left\langle \frac{Tr\left[QB_k\sqrt{\Sigma_t}dW_t\right]}{\sqrt{Tr\left[QB_k\Sigma_t B_k^\top Q^\top\right]}}, \frac{Tr\left[QB_kQ^\top QB_k\sqrt{\Sigma_t}dW_t\right]}{\sqrt{Tr\left[\Sigma_t B_k Q^\top QB_kQ^\top QB_kQ^\top QB_k\right]}} \right\rangle \\
&= \frac{\left(\sum_{p,q,r,s} Q_{pq}B_{qr}\sqrt{\Sigma}_{rs}dW_{sp}\right)\left(\sum_{a,b,c,d,e,f,g} Q_{ab}B_{bc}Q^\top_{cd}Q_{de}B_{ef}\sqrt{\Sigma}_{fg}dW_{ga}\right)}{\sqrt{Tr\left[QB_k\Sigma B_k^\top Q^\top\right]}\sqrt{Tr\left[\Sigma B_kQ^\top QB_kQ^\top QB_kQ^\top QB_k\right]}} \\
&= \frac{\sum_{a,b,c,d,e,f,g,q,r} B_{fe}Q^\top_{ed}B_{cb}Q^\top_{ba}Q_{aq}B_{qr}\sqrt{\Sigma}_{rg}\sqrt{\Sigma}_{gf}dt}{\sqrt{Tr\left[QB_k\Sigma B_k^\top Q^\top\right]}\sqrt{Tr\left[\Sigma B_kQ^\top QB_kQ^\top QB_kQ^\top QB_k\right]}} \\
&= \frac{Tr\left[B_kQ^\top QB_kQ^\top QB_k\Sigma_t\right]dt}{\sqrt{Tr\left[QB_k\Sigma_t B_k^\top Q^\top\right]}\sqrt{Tr\left[\Sigma_t B_kQ^\top QB_kQ^\top QB_kQ^\top QB_k\right]}}.
\end{aligned}
$$

Now we turn on the positivity of the skew. With the notation in the proof of Proposition 3.2, from $\xi_k > \xi_{k+1}$ we have $u_k \prec u_{k+1}$ and then $B_k \in S_d^+$. In all terms in the numerator and the denominator we recognize congruent transformations of matrices in $S_d^+$ which leave the eigenvalues unchanged. The self-duality of $S_d^+$ allows us to claim that all traces are positive, hence we are done.

## 9. Appendix B: the characteristic function

In order to price caplets, we need to have a more explicit form for the characteristic function appearing in Proposition 3.4. Once we have this expression we can plug in the functions $\phi_\tau(u)$ and $\psi_\tau(u)$ to obtain a closed form solution. The pricing problem will be then solved via FFT. Recall that we are considering the following expectation:

$$
\begin{aligned}
\varphi(v) &= \mathbb{E}^{\mathbb{P}_{T_{k+1}}}\left[e^{i(v-(\alpha+1)i)\left(A_k+Tr\left[B_k\Sigma_{T_k}\right]\right)}\right] \\
&= e^{i(v-(\alpha+1)i)A_k}\mathbb{E}^{\mathbb{P}_{T_{k+1}}}\left[exp\left\{Tr\left[\underbrace{i\left(v-(\alpha+1)\,i\right)B_k}_{u}\Sigma_{T_k}\right]\right\}\right] \qquad (9.1)
\end{aligned}
$$

where

$$
\begin{aligned}
A_k &:= -\phi_{T_N-T_k}(u_k) + \phi_{T_N-T_k}(u_{k+1}); \\
B_k &:= -\psi_{T_N-T_k}(u_k) + \psi_{T_N-T_k}(u_{k+1}). \qquad (9.2)
\end{aligned}
$$

As we computed the shape of the function $\phi_\tau(u)$ and $\psi_\tau(u)$ under the $\mathbb{P}_{T_N}$-forward measure, we need to switch from the $\mathbb{P}_{T_{k+1}}$ to the $\mathbb{P}_{T_N}$-forward measure:

$$
\begin{aligned}
& e^{i(v-(\alpha+1)i)A_k}\mathbb{E}^{\mathbb{P}_{T_{k+1}}}\left[e^{Tr\left[u\Sigma_{T_k}\right]}\right] \\
&= e^{i(v-(\alpha+1)i)A_k}\mathbb{E}^{\mathbb{P}_{T_N}}\left[\frac{\partial\mathbb{P}_{T_{k+1}}}{\partial\mathbb{P}_{T_N}}e^{Tr\left[u\Sigma_{T_k}\right]}\right] \\
&= e^{i(v-(\alpha+1)i)A_k}\mathbb{E}^{\mathbb{P}_{T_N}}\left[\frac{M_{T_k}^{u_{k+1}}}{M_0^{u_{k+1}}}e^{Tr\left[u\Sigma_{T_k}\right]}\right], \qquad (9.3)
\end{aligned}
$$

where the last equation follows from (3.16). Let us focus on the expectation which becomes:

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{P}_{T_N}}\Bigg[exp\Big\{ - \phi_{T_N-T_k}(u_{k+1}) - Tr\left[\psi_{T_N-T_k}(u_{k+1})\Sigma_{T_k}\right] \\
&+ \phi_{T_N}(u_{k+1}) + Tr\left[\psi_{T_N}(u_{k+1})\Sigma_0\right] + Tr\left[u\Sigma_{T_k}\right]\Big\}\Bigg] \\
&= exp\Big\{ - \phi_{T_N-T_k}(u_{k+1}) + \phi_{T_N}(u_{k+1}) + Tr\left[\psi_{T_N}(u_{k+1})\Sigma_0\right]\Big\} \\
&\times \mathbb{E}^{\mathbb{P}_{T_N}}\left[e^{Tr\left[\left(-\psi_{T_N-T_k}(u_{k+1})+u\right)\Sigma_{T_k}\right]}\right] \\
&= exp\Big\{ - \phi_{T_N-T_k}(u_{k+1}) + \phi_{T_N}(u_{k+1}) + Tr\left[\psi_{T_N}(u_{k+1})\Sigma_0\right] \\
&- \phi_{T_k}\Big( - \psi_{T_N-T_k}(u_{k+1}) + u\Big) - Tr\left[\psi_{T_k}\Big( - \psi_{T_N-T_k}(u_{k+1}) + u\Big)\Sigma_0\right]\Big\}.
\end{aligned} \tag{9.4}
$$

Now, recalling the previous terms in front of the expectation in (9.3), we obtain the final expression which is

$$
\begin{aligned}
exp\Bigg\{ & i(v-(\alpha+1)i)\Big(\overbrace{-\phi_{T_N-T_k}(u_k)+\phi_{T_N-T_k}(u_{k+1})}^{A_k}\Big) \\
& - \phi_{T_N-T_k}(u_{k+1}) + \phi_{T_N}(u_{k+1}) \\
& -\phi_{T_k}\Big( - \psi_{T_N-T_k}(u_{k+1}) + \underbrace{i(v-(\alpha+1)i)\Big(\overbrace{-\psi_{T_N-T_k}(u_k)+\psi_{T_N-T_k}(u_{k+1})}^{B_k}\Big)}_{u}\Big) \\
& + Tr\left[\psi_{T_N}(u_{k+1})\Sigma_0\right] \\
& - Tr\left[\psi_{T_k}\Bigg( - \psi_{T_N-T_k}(u_{k+1}) + \underbrace{i(v-(\alpha+1)i)\overbrace{\Big( - \psi_{T_N-T_k}(u_k)+\psi_{T_N-T_k}(u_{k+1})\Big)}^{B_k}}_{u}\Bigg)\Sigma_0\right]\Bigg\}.
\end{aligned}
$$

## 10. Figures

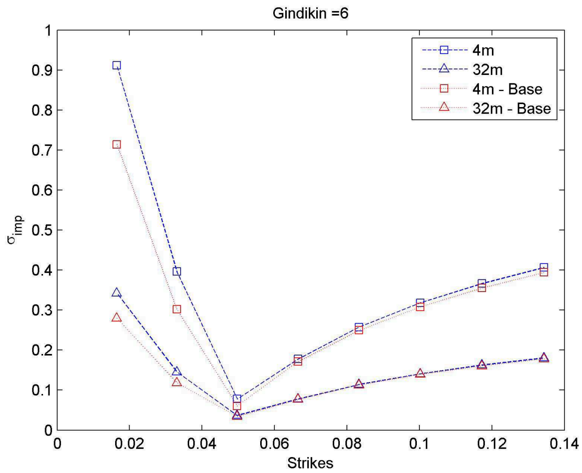



FIGURE 1. Doubling $\beta$ with respect to the basic case causes an upward shift of the surface. The plot represents the two smiles (4 months and 32 months) for the basic ($\beta = 3$) and the modified case ($\beta = 6$).

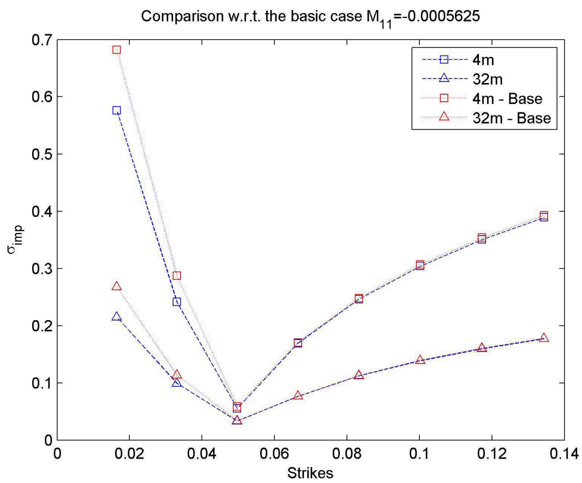



FIGURE 2. Impact of $M_{11}$. $M_{11}$ is negative and the present image shows the effects on the two smiles (4 months and 32 months) we obtain when we multiply it by a constant $c = 1.8$

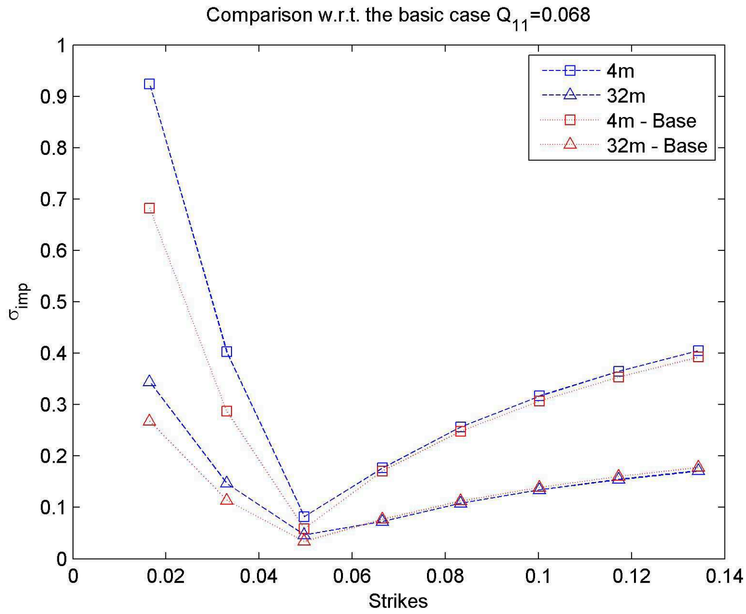



FIGURE 3. Impact of $Q_{11}$. $Q_{11}$ is positive and the present image shows the effects on the two smiles (4 months and 32 months) we obtain when we multiply it by a constant $c = 2$

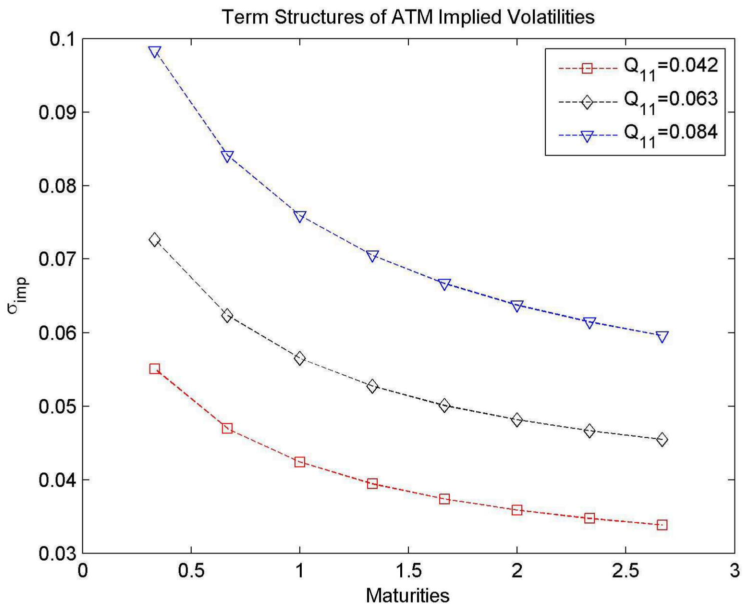



FIGURE 4. Impact of $Q$ on the term structures of ATM implied volatilities. Here we consider $Q_{11}$ and multiply its value by a constant $c = 1, 1.5, 2$ so as to get the values in the legend.

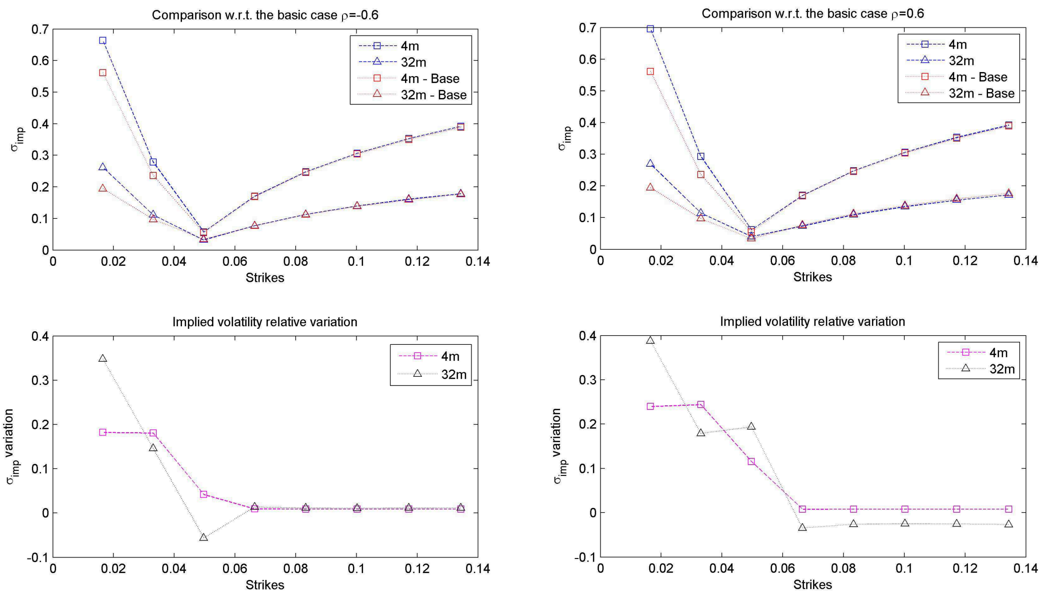



FIGURE 5. The images above highlight the flexibility of the Wishart Libor model. We are able to impose different patterns to the term structure of ATM implied volatility. On the top we have the smiles and on the bottom we observe the relative changes of the smiles, i.e. for every point of the smiles we calculate the quantity $\left(\sigma^{imp}_{final} - \sigma^{imp}_{initial}\right) / \sigma^{imp}_{initial}$. Notice in particular the situation on the left side, where we observe around 5% (ATM) an increase of the short term smile and a decrease on the long term.

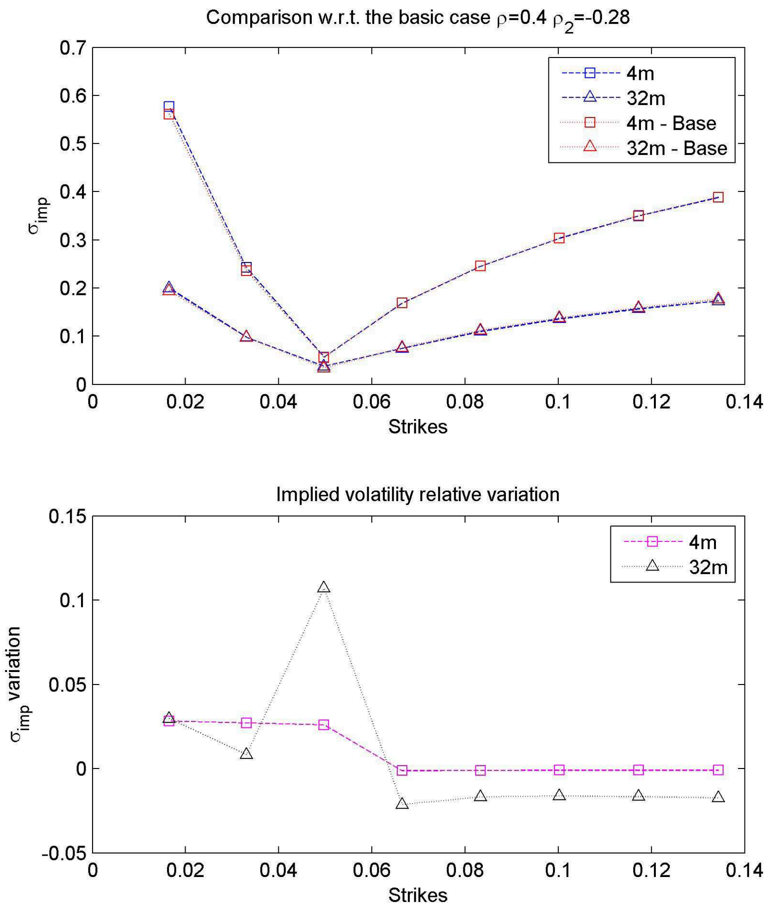



FIGURE 6. Impact on the implied volatility surface when both $M$ and $Q$ are parametrized as symmetric matrices. Notice the level around 5%, corresponding to ATM. This shows that if we parametrize both $M$ and $Q$ via $\rho, \rho_2$ we have a flexible setting which is controlled just by two parameters that allow us to perform different combinations. In particular $\rho$ and $\rho_2$ have opposite impacts in the present example ($\rho > 0$ whereas $\rho_2 < 0$), meaning that we have a good degree of control.

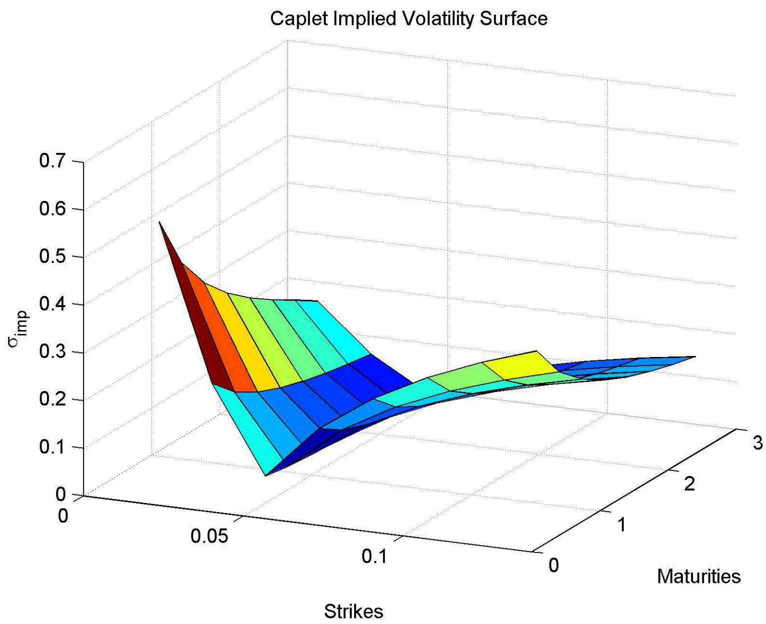



FIGURE 7. Caplet Implied Volatility Surface generated by the Wishart Libor model

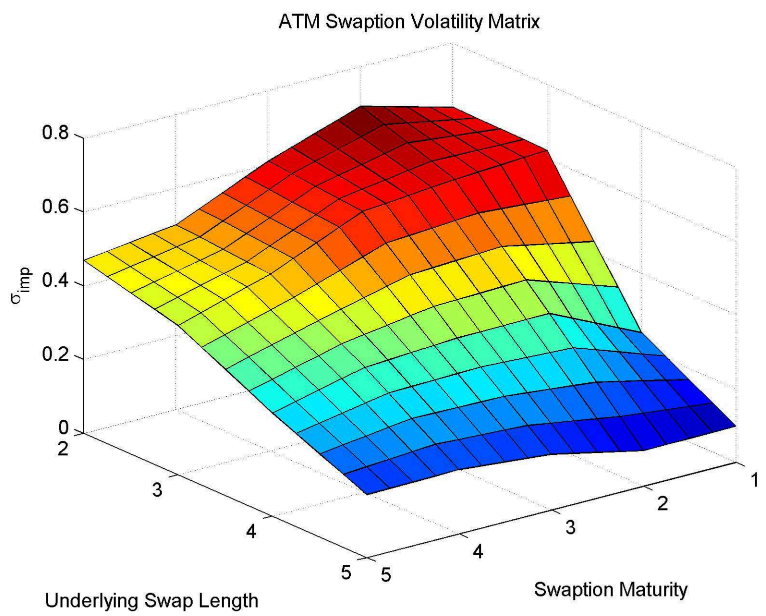



FIGURE 8. ATM Swaption Implied Volatility Surface generated by the Wishart Libor model

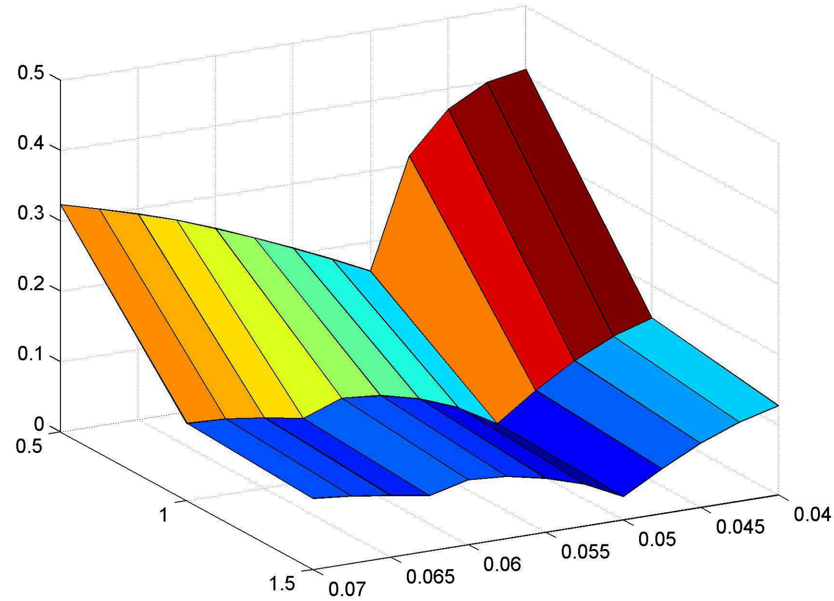



FIGURE 9. Caplet Implied Volatility Surface generated by the compound Poisson Libor model with Wishart distributed jumps

**Part 3**

# Foreign exchange market

CHAPTER 4

# The multi-Heston case

## 1. Introduction

Derivatives with multiple underlying components have attracted in recent times an increasing amount of attention, partly in form of public scrutiny and criticisms. Their popularity extends to the different asset classes traded in the financial markets, ranging from basket options written on equity stocks and foreign exchange rates, to CDO structures on various fixed income instruments and complex hybrid structures combining the different underlying types (see e.g. Wystup (2006), Clark (2011), Esquível et al. (2010)). From a client perspective the appeal of such structures is clear and understandable: diversification reduces the exposure to individual components while the whole structure is cheaper to buy than multiple single-underlying options, see Qu (2010). The correlations between the returns of the different components are the key inputs in the determination of this rebate. Most often, however, these correlations are not observable in the market; a fact that leads to material uncertainties and complications in the valuation and risk management of, even rather standard, multi-component derivatives.

There is nevertheless a notable and mostly overlooked example, in which much can be said from the market about the correlation structure. This is often the case of derivatives with multiple underlying FX rates (see Esquível et al. (2010)). Rather paradoxically, here we face the opposite problem: instead of having little evidence on where correlations trade, the market gives many indications on their structure, potentially more than any standard model, like a copula model, can handle. This information can be extracted from the liquid vanilla markets and the fact that, differently from other asset classes, appropriate multiplication/division of FX rates are still FX rates, and hence directly observable through standard FX spot transactions. To capture the market correlation structure we need a model that is able to price consistently vanilla options across different FX pairs (see e.g. Castagna (2010), Beneder and Elkenbracht-Huizing (2003)). At the same time, given the high dimensionality of the problem, for any possible practical application we seek to retain some kind of analytical tractability. In this paper, we present a first proposal of such a model.

The focus of our analysis is on the simultaneous calibration of several FX vanilla surfaces. This is indeed an important model prerequisite for pricing structures with multiple underlying FX rates. For the sake of simplicity, let us consider as an example the FX market composed by three currencies: say Euro (EUR), US dollar (USD), and Japanese Yen (JPY). All three currency pairs EUR/USD, USD/JPY, and EUR/JPY are liquidly traded, both as FX spot transactions and vanilla FX options. Let $S^{\mathrm{d,f}}(t)$ be the spot exchange rate at time $t$ as the amount of domestic (d) currency for one unit of foreign currency (f). As a consequence of the standard triangular relationship between the FX rates, e.g. $S^{\mathrm{JPY,EUR}}(t) = S^{\mathrm{JPY,USD}}(t)S^{\mathrm{USD,EUR}}(t)$, the price of a multi-dimensional option on any of the two of the three pairs is also *implicitly* sensitive to the volatility of the excluded pair. This fact is easily understood in the case of no-skew. To fix ideas let us consider a basket option on EUR/USD and USD/JPY. The correlation between the two main FX rates $\rho_{\mathrm{EUR/USD-USD/JPY}}$ can be derived from the volatility of the cross via the triangular relation (see e.g. Clark (2011) p. 228)

$$\rho_{\mathrm{EUR/USD-USD/JPY}} = \frac{\sigma^2_{\mathrm{EUR/JPY}} - \sigma^2_{\mathrm{EUR/USD}} - \sigma^2_{\mathrm{USD/JPY}}}{2\sigma_{\mathrm{EUR/USD}}\sigma_{\mathrm{USD/JPY}}}. \tag{1.1}$$

By varying the cross volatility $\sigma_{\mathrm{EUR/JPY}}$ and keeping the main volatilities fixed, one can span over any value of the correlation between -1 and 1, yielding significantly different values for the basket structure. Similarly, if the volatility skew is included, the basket structure becomes sensitive to *all* three volatility smiles. A good pricing model must be able to reproduce all three vanilla markets; theoretically, a very challenging prerequisite.

The easiest way to proceed is to calibrate the volatility smiles of the two main currency pairs using typical unidimensional models and try to obtain the third from the correlation between them. If, however, we stick to a simple mathematical modelling of the correlation, say either constant or time-dependent, we face two problems. Firstly, a simultaneous calibration to all volatility smiles in the triangle is in general difficult to obtain, as a constant correlation does not provide enough flexibility to match the smile of the FX cross. Secondly, the model is not functionally symmetric with respect to which FX pairs we choose to be the main ones and which one the cross. In other terms, the stochastic process of the cross FX rate is functionally different from the processes of the main FX rates it is derived from.
This latter point, although theoretical at first sight, is of key practical importance. In principle we have the freedom of choosing any of the currencies as our numéraire when pricing a multi-currency structure (although some choices are more natural than others, e.g. the payout currency which is specified in the contract). A model that is able to treat symmetrically all currencies must be therefore preferred. The main purpose of this paper is to specify a model with this property in the class of multi-dimensional stochastic volatility models. To achieve our goal, we need a paradigm change in the model specification. We have to forgo the standard approach of leaving the task of matching the volatility skew of the cross FX rate to the correlation between the main exchange rates. Instead of putting the currency pairs at the basis of our model, we start from the observation that any exchange rate may be seen as a ratio between two quantities, the value of the currencies with respect to some universal numéraire, and include this feature in the specification of the model: this reflects the point of view of the Benchmark approach in Heath and Platen (2006a) and Heath and Platen (2006b). In this way, our model does not change qualitatively depending on which perspective is used. The benefits of this property are manifold: *i)* the calibration is universal, i.e., it is independent on the exotic product being priced (due to the symmetry of the numéraire currency), *ii)* the price of the exotic will not vary depending on which perspective is chosen; moreover, *iii)* the model produces symmetric and consistent risk sensitivities with possibly positive impacts on the way multi-dimensional option books can be risk managed.

The model we present in this paper is a multi-factor stochastic volatility model of Heston , see Heston (1993), which extends the standard approach in Garman and Kohlhagen (1983). The Heston dynamics leads to an affine model which is known to retain analytical tractability. We will provide a complete discussion concerning the set of risk neutral measures and the rules to change between them. Rather remarkably, as a consequence of the specific Heston-type dynamics, the model remains functionally invariant, after parameter rescaling, when the risk-neutral measure is changed. This is the key feature that allows us to obtain a universal calibration with reasonable computational effort. We will then test the model on real market data and show how a joint calibration of the volatility smiles of EUR/USD/JPY triangle is possible. In- and out-of-sample calibration tests will be reported to comment on the robustness of the parameter estimation.

The chapter is organized as follows: we present the model in Sect. 2, initially using the perspective given by some kind of universal numéraire. We continue with the basic properties of the model, such as the positivity of the instantaneous covariance matrix and the presence of stochastic skewness, before presenting the invariance of the model and transformation rule of its parameters when the risk neutral measure is changed in Sect. 3. The explicit formulae for the characteristic function and option prices are given in Sect. 4 before turning to the small vol-of-vol expansions of the option prices and implied volatilities in Sect. 5. Finally, the joint calibration to EUR/USD/JPY market volatility smiles is presented in Sect. 6, together with a discussion of the procedure and the results, including the Feller condition and moment explosion.

## 2. Multifactor Heston-based exchange model

We consider a foreign exchange market in which $N$ currencies are traded between each other via standard FX spot and FX vanilla option transactions. We start by considering the value of each of these currencies in units of a universal numéraire. We will see that the discussion is independent on the exact specification of this numéraire. To fix the ideas, one can think of it as a precious metal, e.g. gold, or the market portfolio like in the benchmark approach of Heath and Platen (2006a) and Heath and Platen (2006b).

Let us work in the risk neutral measure defined by the universal numéraire and call $S^{0,i}(t)$ the value at time $t$ of one unit of the currency $i$ in terms of our universal numéraire (note that $S^{0,i}(t)$ can itself be thought as an exchange rate, between the universal numéraire and the currency $i$). We model each of the $S^{0,i}(t)$ via a multi-variate Heston stochastic volatility model Heston (1993) with $d$ independent Cox-Ingersoll-Ross (CIR) components Cox et al. (1985), $\mathbf{V}(t) \in \mathbb{R}^d$, and an equivalent number of driving noises, $\mathbf{Z}(t) \in \mathbb{R}^d$. The dimension $d$ can be chosen according to the specific problem and may reflect a PCA-type analysis. We further assume that these stochastic volatility components are *common* between the different $S^{0,i}(t)$. Formally, we write

$$\frac{dS^{0,i}(t)}{S^{0,i}(t)} = (r^0 - r^i)dt - (\mathbf{a}^i)^\top \mathrm{Diag}(\sqrt{\mathbf{V}(t)})d\mathbf{Z}(t), \quad i = 1, \ldots, N; \tag{2.1}$$

$$dV_k(t) = \kappa_k(\theta_k - V_k(t))dt + \xi_k \sqrt{V_k(t)}dW_k(t), \quad k = 1, \ldots, d; \tag{2.2}$$

where $\kappa_k, \theta_k, \xi_k \in \mathbb{R}$ are standard parameters in a CIR dynamics. $\mathrm{Diag}(\sqrt{\mathbf{V}(t)})$ denotes the diagonal matrix with the square root of the elements of the vector $\mathbf{V}(t)$ in the principal diagonal, this term is multiplied with the linear vector $\mathbf{a}^i \in \mathbb{R}^d$ ($i = 1, \ldots, N$); as a result, the dynamics of the exchange rate is driven by a linear projection of the variance factor $\mathbf{V}(t)$ along a direction parametrized by $\mathbf{a}^i$, namely the total instantaneous variance is $(\mathbf{a}^i)^\top \mathrm{Diag}(\mathbf{V}(t))\mathbf{a}^i dt$. In each monetary area $i$, the money-market account accrues interest based on the deterministic risk free rate $r^i$,

$$dB^i(t) = r^i B^i(t)dt, \quad i = 1, \ldots, N; \tag{2.3}$$

in our gold analogy $r^0$ is the precious metal lease rate. Finally, we assume an orthogonal correlation structure between the stochastic drivers

$$\langle dZ_k(t)dW_h(t)\rangle = \rho_k \delta_{kh} dt, \quad k, h = 1, \ldots, d, \tag{2.4}$$

together with $\langle dZ_k(t)dZ_h(t)\rangle = \delta_{kh}dt$ and $\langle dW_k(t)dW_h(t)\rangle = \delta_{kh}dt$.

This concludes the description of our model.

The idea behind this approach is that each exchange rate is driven by several independent drivers $dZ_k(t)$ ($k = 1, .., d$), each with an independent stochastic variance factor $V_k(t)$, to which $dZ_k(t)$ is partially correlated via $\rho_k$. The vectors $\mathbf{a}^i$ ($i = 1, \ldots, N$) describe by how much each of the different volatilities contribute to the dynamics of $S^{0,i}(t)$ . This correlation structure is responsible for the appearance of non-standard effects in the model, like a stochastic skewness, see subsection 2.3.

All in all, we have introduced a total number of parameters equal to $N_{\mathrm{P}} = Nd + 5d$ ($Nd$ from the vectors $\mathbf{a}^i$ and 5 for each CIR process, $\kappa_k, \theta_k, \xi_k, \rho_k$ and the initial value $V_k(0)$) to describe the volatility skew of $(N^2 - N)/2$ currency pairs. As rule of thumb, assuming that each currency pair can be approximately modeled by a standard one-dimensional Heston model, which is described by 5 parameters, around $5(N^2 - N)/2$ parameters are needed to fit all volatility surfaces; the value of $d$ should be chosen to produce approximately this number of parameters, if not less, to avoid instabilities due to overfitting.

Let us now turn our attention to the exchange rate $S^{i,j}(t)$ between two different currencies, say $i$ and $j$. We set by definition $S^{i,j}(t) := S^{0,j}(t)/S^{0,i}(t)$. By straightforward calculation, we obtain for $i, j = 1, .., N$:

$$\frac{dS^{i,j}(t)}{S^{i,j}(t)} = (r^i - r^j)dt + (\mathbf{a}^i - \mathbf{a}^j)^\top \mathrm{Diag}(\mathbf{V}(t))\mathbf{a}^i dt + (\mathbf{a}^i - \mathbf{a}^j)^\top \mathrm{Diag}(\sqrt{\mathbf{V}(t)})d\mathbf{Z}(t). \tag{2.5}$$

Note that at this stage we are still working under the risk neutral measure defined by the universal numéraire. The additional drift term in (2.5) can be understood as a quanto adjustment between the currency 0 and $i$.

PROPOSITION 4.1. *The dynamics of the exchange rate (2.1) satisfies the triangular relation, namely for all* $t \geq 0$ *and* $i, j, l = 1, .., N$*:*

$$dS^{i,j}(t) = d\left(S^{i,l}(t)S^{l,j}(t)\right). \tag{2.6}$$

PROOF. Straightforward application of Ito's rule. □

This symmetry property is fundamental in order to have a model that yields universal calibration and consistent pricing of exotic options. In the following subsections, we will analyze some additional properties of the model and familiarize with the meaning of the different parameters.

**2.1. Vectors $\mathbf{a}^i$ cannot be the canonical basis.** A rather natural choice would be to set $\mathbf{a}^i$ equal to the canonical basis $\mathbf{e}^i$ (i.e. the $i$-th element of the canonical basis of $\mathbb{R}^N$, $e^i_l = \delta_{li},\ i, l = 1, .., N$), then (2.5), for $i \neq j$, reads:

$$\frac{dS^{i,j}(t)}{S^{i,j}(t)} = (r^i - r^j)dt + V^i(t)dt + \sqrt{V_i(t)}dZ^i(t) - \sqrt{V_j(t)}dZ^j(t), \tag{2.7}$$

which in the 3-currency case leads to the 3-factor Heston model. The problem with this choice is that the covariances (and thus the correlations) between different pairs are forced to be positive,

$$\left\langle \frac{dS^{i,j}(t)}{S^{i,j}(t)}, \frac{dS^{i,l}(t)}{S^{i,l}(t)} \right\rangle = V^i(t)dt \geq 0. \tag{2.8}$$

As we need the flexibility to define negative correlation between FX rates, the vectors $\mathbf{a}^i$ cannot be taken equal to the canonical basis. These additional parameters are needed to describe a multi-dimensional FX market.

**2.2. Stochastic covariance matrix.** By construction, the model (2.1) has a stochastic instantaneous $(N-1)\times(N-1)$ covariance matrix. We prove here that the specification of the model always leads to a positive definite covariance matrix, a fundamental prerequisite for any well-posed multi-dimensional model. For example, we consider the case of three currencies and the associated $2\times 2$ candidate covariance matrix where we fix the point of view of currency $i$:

$$\mathrm{Cov}^i(t) = \begin{pmatrix} \left\langle \frac{dS^{i,j}(t)}{S^{i,j}(t)} \right\rangle & \left\langle \frac{dS^{i,j}(t)}{S^{i,j}(t)}, \frac{dS^{i,l}(t)}{S^{i,l}(t)} \right\rangle \\ \left\langle \frac{dS^{i,j}(t)}{S^{i,j}(t)}, \frac{dS^{i,l}(t)}{S^{i,l}(t)} \right\rangle & \left\langle \frac{dS^{i,l}(t)}{S^{i,l}(t)} \right\rangle \end{pmatrix}, \tag{2.9}$$

where

$$\left\langle \frac{dS^{i,j}(t)}{S^{i,j}(t)}, \frac{dS^{i,l}(t)}{S^{i,l}(t)} \right\rangle = \left(\mathbf{a}^i - \mathbf{a}^j\right)^\top \mathrm{Diag}(\mathbf{V}(t)) \left(\mathbf{a}^i - \mathbf{a}^l\right) dt. \tag{2.10}$$

Variances, i.e., the diagonal terms in the covariance matrix, are clearly positive due to the positiveness of $V_k(t), k = 1,..,d$. To prove that the whole covariance matrix is positive definite, it is sufficient to introduce the $(N-1)\times d$ matrix $B$ defined as follows:

$$B_{j,k} := a^j_k - a^1_k; \quad k = 1,..,d; \quad j = 2,\dots,N, \tag{2.11}$$

where without loss of generality we can take the point of view of currency $i = 1$. With this notation the $(N-1)\times(N-1)$ variance/covariance matrix $\mathrm{Cov}^1(t) = \mathrm{Cov}(t)$ is given by

$$\mathrm{Cov}(t) = B\,\mathrm{Diag}(\mathbf{V}(t))B^\top, \tag{2.12}$$

which is positive semidefinite since for all $\mathbf{x}\in\mathbb{R}^d$:

$$\mathbf{x}^\top \mathrm{Cov}(t)\mathbf{x} = (B^\top\mathbf{x})^\top \mathrm{Diag}(\mathbf{V}(t))B^\top\mathbf{x} \geq 0,$$

where the last inequality follows from the positive definiteness of the matrix $\mathrm{Diag}(\mathbf{V}(t))$.
In conclusion the variance-covariance matrix is well defined, and, as an important side effect, we have the usual bound for the correlations, i.e. all correlations are bounded by one in absolute value.

**2.3. Stochastic Skew.** To shed some additional light on the meaning of the vectors $\mathbf{a}^i$ we calculate the skewness $\varsigma$ for a given exchange rate $S^{i,j}$, defined as the correlation between the log returns and the stochastic variance:

$$\varsigma^{i,j}(t) = \frac{\langle \mathrm{Noise}(\log S^{i,j}), \mathrm{Noise}(\mathrm{Vol}(\log S^{i,j}))\rangle_t}{\sqrt{\langle \mathrm{Noise}(\log S^{i,j})\rangle\rangle_t}\sqrt{\langle(\mathrm{Noise}(\mathrm{Vol}(\log S^{i,j})))\rangle_t}}. \tag{2.13}$$

Differently from standard single factor models, multifactor Heston models produce a stochastic skewness. In fact, by straightforward calculation we obtain:

$$\begin{aligned} \left\langle d\log S^{i,j}, dV_k\right\rangle_t &= \left\langle \sum_{l=1}^d \left(\mathbf{a}^i_l - \mathbf{a}^j_l\right)\sqrt{V_l}dZ_l, \xi_k\sqrt{V_k}\rho_k dZ_k \right\rangle_t \\ &= \left(a^i_k - a^j_k\right)\xi_k V_k(t)\rho_k dt \qquad i,j = 1,..,N;\ k = 1,...,d. \end{aligned} \tag{2.14}$$

Combining this term with

$$\left\langle d\log S^{i,j}, d\log S^{i,j}\right\rangle_t = \left(\mathbf{a}^i - \mathbf{a}^j\right)^\top \mathrm{Diag}(\mathbf{V}(t))\left(\mathbf{a}^i - \mathbf{a}^j\right) dt$$

and $\langle dV_k, dV_k \rangle_t = \xi_k^2 V_k(t)dt$, we obtain

$$\varsigma^{i,j}(t) = \frac{\sum_{k=1}^d \left(a_k^i - a_k^j\right) \xi_k V_k(t)\rho_k}{\sqrt{\sum_{k=1}^d \xi_k^2 V_k(t)}\sqrt{\left(\mathbf{a}^i - \mathbf{a}^j\right)^\top \text{Diag}(\mathbf{V}(t))\left(\mathbf{a}^i - \mathbf{a}^j\right)}}. \tag{2.15}$$

This quantity is stochastic due to the presence of the variance factors $V_k(t)$, $k = 1, .., d$ Christoffersen et al. (2009). The vectors $\mathbf{a}^i$ are directly related to the amount of skewness for each of the different exchange rates.

## 3. Numéraire invariance

Up to now we have worked under the risk neutral measure defined by our (rather unspecified) universal numéraire. In practical pricing applications, it is more convenient to change the numéraire to any of the currencies included in our FX multi-dimensional system. Without loss of generality, let us consider the risk neutral measure defined by the $i$-th money market account $B^i(t)$ and derive the dynamical equations for the standard FX rate $S^{i,j}(t)$, its inverse $S^{j,i}(t)$, and a generic cross $S^{j,l}(t)$ for $i, j, l = 1, .., N$.

Under the assumptions of the fundamental theorem of asset pricing (cfr. e.g. Björk (2004), chapters 13 and 14), investing domestic money into any foreign currency money-market account cannot produce a risk free return different from $r^i$. In other terms, the ratio $S^{i,j}(t)B^j(t)/B^i(t)$ must be a local martingale under the $i$-th risk free measure $\mathbb{Q}^i$, provided that such risk neutral measure exists. Hence,

$$\begin{aligned} d\left(\frac{S^{i,j}(t)B^j(t)}{B^i(t)}\right) &= \frac{S^{i,j}(t)B^j(t)}{B^i(t)}\left((\mathbf{a}^i - \mathbf{a}^j)^\top \text{Diag}(\mathbf{V}(t))\mathbf{a}^i dt + (\mathbf{a}^i - \mathbf{a}^j)^\top \text{Diag}(\sqrt{\mathbf{V}(t)})d\mathbf{Z}(t)\right) \\ &= \frac{S^{i,j}(t)B^j(t)}{B^i(t)}(\mathbf{a}^i - \mathbf{a}^j)^\top \text{Diag}(\sqrt{\mathbf{V}(t)})d\mathbf{Z}^{\mathbb{Q}^i}(t). \end{aligned} \tag{3.1}$$

In the last line we implicitly defined the new Brownian motion vector $\mathbf{Z}^{\mathbb{Q}^i}(t)$ under the measure $\mathbb{Q}^i$ from the constraint of having a $\mathbb{Q}^i$-local martingale and by Girsanov theorem:

$$d\mathbf{Z}(t)^{\mathbb{Q}^i} = d\mathbf{Z}(t) + \text{Diag}(\sqrt{\mathbf{V}(t)})\mathbf{a}^i dt, \quad i = 1, .., N. \tag{3.2}$$

If we denote $\mathbb{Q}^0$ the risk neutral measure associated with the universal numéraire, the Radon-Nikodym derivative corresponding to the change of measure from $\mathbb{Q}^0$ to $\mathbb{Q}^i$ reads

$$\frac{d\mathbb{Q}^i}{d\mathbb{Q}^0}|_t = \exp\left(-\int_0^t \left(\mathbf{a}^i\right)^\top \text{Diag}(\sqrt{\mathbf{V}(s)})d\mathbf{Z}(s) - \frac{1}{2}\int_0^t \left(\mathbf{a}^i\right)^\top \text{Diag}(\mathbf{V}(s))\mathbf{a}^i ds\right), \tag{3.3}$$

where we observe that the shift is performed only on the process $\mathbf{Z}(t)$.
The $\mathbb{Q}^i$ risk neutral dynamics of the exchange rate $S^{i,j}(t)$ becomes

$$dS^{i,j}(t) = S^{i,j}(t)\left((r^i - r^j)dt + (\mathbf{a}^i - \mathbf{a}^j)^\top \text{Diag}(\sqrt{\mathbf{V}(t)})d\mathbf{Z}^{\mathbb{Q}^i}(t)\right), \tag{3.4}$$

as desired.
Given our assumption on the correlation structure in (2.4), we can write the following standard factorization under $\mathbb{Q}^0$

$$dW_k(t) = \rho_k dZ_k(t) + \sqrt{1 - \rho_k^2}dZ_k^\perp(t), \quad k = 1, ..., d. \tag{3.5}$$

where $\mathbf{Z}^\perp(t)$ is a Brownian motion independent of $\mathbf{Z}(t)$. Hence the measure change has also an impact on the variance processes, via the correlations $\rho_k$, $k = 1, .., d$,

$$dW_k^{\mathbb{Q}^i}(t) = dW_k(t) + \rho_k \left(\mathbf{e}^k\right)^\top \text{Diag}(\sqrt{\mathbf{V}(t)})\mathbf{a}^i dt. \tag{3.6}$$

The component of $dW_k(t)$ which is orthogonal to the spot driver $d\mathbf{Z}(t)$ is not affected by the measure change. This choice may seem arbitrary at first sight, but it turns out to be consistent with the market practice. More importantly, Del Baño Rollin (2008) shows that this assumption constitutes a sufficient condition ensuring that the foreign-domestic parity is satisfied. We finally obtain the dynamic equations under the new measure. With an appropriate redefinition of the CIR parameters

$$\begin{aligned}\rho_k^{\mathbb{Q}^i} &= \rho_k, \\ \kappa_k^{\mathbb{Q}^i} &= \kappa_k + \xi_k \rho_k a_k^i, \\ \theta_k^{\mathbb{Q}^i} &= \theta_k \frac{\kappa_k}{\kappa_k^{\mathbb{Q}^i}},\end{aligned}$$

we can recast the variance SDE in its original form

$$dV_k(t) = \kappa_k^{\mathbb{Q}^i}(\theta_k^{\mathbb{Q}^i} - V_k(t))dt + \xi_k \sqrt{V_k(t)} dW_k^{\mathbb{Q}^i}(t). \tag{3.7}$$

We can now show that the Girsanov theorem used in (3.2) has been correctly applied.

PROPOSITION 4.2. *The exponential local martingale (3.3) is a true martingale, that is* $\mathbb{E}^{\mathbb{Q}^0}\left[\frac{d\mathbb{Q}^i}{d\mathbb{Q}^0}|_t\right] = 1.$

PROOF. See the Appendix. □

In the Appendix we derive a proof of the previous result based on the Feller explosion test for diffusions as in Wong and Heyde (2004) and Mijatović and Urusov (2011). In order to get a quick intuition, from (3.3) we can check that the process $\mathbf{V}$ has the same behavior at the boundaries 0 and $+\infty$ under both measures $\mathbb{Q}^0$ and $\mathbb{Q}^i$. In fact the dimension of the process $\mathbf{V}$ and the boundary behavior at infinity are the same (that is the Feller condition is unchanged and the probability of an explosion in finite time is zero). As a consequence, the Radon-Nikodym derivative defined through (3.3) is a (true) martingale, so that the Girsanov change of measure is allowed.

The invariance of the functional form of the model under measure change is an appealing feature of our model; other specifications of the stochastic volatility will almost surely break this symmetry. Financially, it makes sense to enforce mean reversion of the variance, other than mean explosion, yielding a condition on $\kappa_k^{\mathbb{Q}^i} > 0$ or conversely on the original model parameters $\kappa_k$, $\xi_k$, $\rho_k$, and $a_k^i$.

The inverse FX rate follows from Ito calculus

$$\begin{aligned}\frac{dS^{j,i}(t)}{S^{j,i}(t)} &= S^{i,j}(t) d\left(\frac{1}{S^{i,j}(t)}\right) \\ &= [r^j - r^i + (\mathbf{a}^j - \mathbf{a}^i)^\top \mathrm{Diag}(\mathbf{V}(t))(\mathbf{a}^j - \mathbf{a}^i)]dt + (\mathbf{a}^j - \mathbf{a}^i)^\top \mathrm{Diag}(\sqrt{\mathbf{V}(t)}) d\mathbf{Z}^{\mathbb{Q}^i}(t),\end{aligned} \tag{3.8}$$

which includes the self-quanto adjustment. Similarly, the SDE of a generic cross FX rate becomes

$$\begin{aligned}\frac{dS^{j,l}(t)}{S^{j,l}(t)} &= \frac{S^{i,j}(t)}{S^{i,l}(t)} d\left(\frac{S^{i,l}(t)}{S^{i,j}(t)}\right) \\ &= [r^j - r^l + (\mathbf{a}^j - \mathbf{a}^l)^\top \mathrm{Diag}(\mathbf{V}(t))(\mathbf{a}^j - \mathbf{a}^i)]dt + (\mathbf{a}^j - \mathbf{a}^l)^\top \mathrm{Diag}(\sqrt{\mathbf{V}(t)}) d\mathbf{Z}^{\mathbb{Q}^i}(t).\end{aligned} \tag{3.9}$$

The additional drift term is the quanto adjustment as described by the current model choice. By applying Girsanov theorem again, this time switching to the $\mathbb{Q}^j$ risk neutral measure, the term is removed while the CIR parameters become

$$\kappa_k^{\mathbb{Q}^j} = \kappa_k^{\mathbb{Q}^i} + \rho_k \xi_k (a_k^j - a_k^i),$$
$$\theta_k^{\mathbb{Q}^j} = \theta_k^{\mathbb{Q}^i} \frac{\kappa_k^{\mathbb{Q}^i}}{\kappa_k^{\mathbb{Q}^j}}, \tag{3.10}$$

together with the invariant $\rho_k^{\mathbb{Q}^j} = \rho_k^{\mathbb{Q}^i}$ and $\xi_k^{\mathbb{Q}^j} = \xi_k^{\mathbb{Q}^i}$. These are the fundamental transformation rules for the model parameters.

## 4. Option pricing

Together with the invariance of the model specification with respect to the numéraire choice, a second central feature of the model is the availability of a (semi)-analytical solution for *all* vanilla option prices. The pricing formula itself is independent of the option underlying, once we work under the risk neutral measure associated with one of the currencies in the option and the parameters are transformed via (3.10).

Let us consider a call option $C(S^{i,j}(t), K^{i,j}, \tau), i, j = 1, .., N, i \neq j$, on a generic FX rate $S^{i,j}(t) = \exp(x^{i,j}(t))$ with strike $K^{i,j}$, maturity $T$ ($\tau = T - t$ is the time to maturity) and face equal to one unit of the foreign currency. We write for the CIR parameters $\kappa_k = \kappa_k^{\mathbb{Q}^i}, \theta_k = \theta_k^{\mathbb{Q}^i}$ and so on, implicitly assuming that they have been transformed via (3.10) in the $i$-th risk neutral measure $\mathbb{Q}^i$. Being an affine model, the (generalized) characteristic function conditioned on the initial values

$$\phi^{i,j}(\omega, t, \tau, x, \mathbf{V}) = \mathbb{E}_t^{\mathbb{Q}^i}[e^{\mathtt{i}\omega x^{i,j}(T)} | x^{i,j}(t) = x, \mathbf{V}(t) = \mathbf{V}] \tag{4.1}$$

can be derived analytically (here $\mathtt{i} = \sqrt{-1}$). Standard numerical integration methods can then be used to invert the Fourier transform to obtain the probability density at $T$ or the vanilla price via integration against the payoff, with overall little computational effort. By applying standard arguments (see e.g. Lewis (2000), Sepp (2003)) the value of a call option can be expressed in terms of the integral of the product of the Fourier transform of the payoff and the generalized characteristic function of the log-asset price[1]:

$$C(S^{i,j}(t), K^{i,j}, \tau) = e^{-r^i \tau} \frac{1}{2\pi} \int_{\mathcal{Z}} \phi^{i,j}(-\mathtt{i}\lambda, t, \tau, x, \mathbf{V}) \Phi(\lambda) d\lambda, \tag{4.2}$$

where

$$\Phi(\lambda) = \int_{\mathcal{Z}} e^{\mathtt{i}\lambda x} \left(e^x - K^{i,j}\right)^+ dx$$

is the Fourier transform of the payoff function and $\mathcal{Z}$ denotes the strip of regularity of the payoff, that is the admissible domain where the integral in (4.2) is well defined. In other words, the pricing problem is essentially solved once the (conditional) characteristic function of the log-exchange rate is known. We recall the relationship between the characteristic function and the moment generating function. In what follows we will derive the moment generating function $G^{i,j}(\omega, t, \tau, x, \mathbf{V})$ (Laplace tranform) from which the characteristic function is easily derived via a rotation in the complex plane

[1]Here we adopt the pricing method of Lewis (2000) who uses the characteristic function computed with a complex argument, also called generalized characteristic function. The complex argument $\omega$ typically belongs to a strip of regularity for the function $\phi^{i,j}$ in order to be able to integrate the payoff function. On the other hand, this method avoids the introduction of the damping integrating factor required by the methodology of Carr and Madan (1999).

$\phi^{i,j}(\omega, t, \tau, x, \mathbf{V}) = G^{i,j}(\mathrm{i}\omega, t, \tau, x, \mathbf{V})$. It is therefore sufficient to determine the moment generating function, as in the following proposition:

PROPOSITION 4.3. *In the Multi-Heston model the conditional Laplace transform of the log-exchange rate is given by:*

$$G^{i,j}(\omega, t, \tau, x, \mathbf{V}) = \exp\left[\omega x + \left(r^i - r^j\right)\omega(\tau) + \sum_{k=1}^{d}\left(A_k^{i,j}(\tau) + B_k^{i,j}(\tau)V_k\right)\right], \tag{4.3}$$

*where for* $k = 1, .., d$*:*

$$A_k^{i,j}(\tau) = \frac{2\kappa_k\theta_k}{\xi_k^2}\log\frac{\lambda_k^+ - \lambda_k^-}{\lambda_k^+ e^{\lambda_k^-(\tau)} - \lambda_k^- e^{\lambda_k^+(\tau)}}; \tag{4.4}$$

$$B_k^{i,j}(\tau) = \frac{\left(\omega^2 - \omega\right)}{2}\left(a_k^i - a_k^j\right)^2 \frac{1 - e^{-\sqrt{\Delta_k}\tau}}{\lambda_k^+ e^{-\sqrt{\Delta_k}\tau} - \lambda_k^-}; \tag{4.5}$$

$$\Delta_k = \left(-\kappa_k + \omega\left(a_k^i - a_k^j\right)\rho_k\xi_k\right)^2 - \xi_k^2\left(\omega^2 - \omega\right)\left(a_k^i - a_k^j\right)^2; \tag{4.6}$$

$$\lambda_k^{\pm} = \frac{\left(-\kappa_k + \omega\left(a_k^i - a_k^j\right)\rho_k\xi_k\right) \pm \sqrt{\Delta_k}}{2}. \tag{4.7}$$

PROOF. See Appendix. □

In summary the call price is known once the characteristic function $G^{i,j}$ is known explicitly, as in the present framework.

## 5. Expansions

As we will see in the sequel, the calibration of our model can be performed by relying on a standard non-linear least squares procedure, which will be employed to minimize the distance between the model implied volatilities and market ones. Model implied volatilities are extracted from the prices produced by the FFT routine. This procedure is quite demanding from a numerical point of view. An alternative approach consists in fitting implied volatilities via a simpler function, for example by looking at a possible relationship between the prices produced by the model, see (4.2), and the standard Black-Scholes formula for a suitable volatility. The next result states that it is possible to approximate the prices of options under the multi-Heston model, via a suitable expansion of the standard Black-Scholes formula and its derivatives. The proof, which is reported in the appendix, relies on arguments which may be found in Lewis (2000) and Da Fonseca and Grasselli (2011) (we drop all currency indices, it is intended that we are considering the $(i, j)$ FX pair). Define $\tau = T - t$ and let us define the real deterministic functions $\mathcal{B}_k^{(0)}, \mathcal{B}_k^{(1)}, \mathcal{B}_k^{(2)}, \mathcal{B}_k^{(3)}, k = 1, .., d$ as

$$\mathcal{B}_k^{(0)}(\tau) = \left(a_k^i - a_k^j\right)^2 \frac{1 - e^{-\kappa_k\tau}}{\kappa_k}; \tag{5.1}$$

$$\mathcal{B}_k^{(1)}(\tau) = \left(a_k^i - a_k^j\right)^3 \rho_k\xi_k\left(\frac{1}{\kappa_k^2} - \frac{e^{-\kappa_k\tau}}{\kappa_k^2} - \frac{\tau e^{-\kappa_k\tau}}{\kappa_k}\right); \tag{5.2}$$

$$\mathcal{B}_k^{(2)}(\tau) = \left(a_k^i - a_k^j\right)^4 \frac{\xi_k^2}{2\kappa_k^2}\left(\frac{1 - e^{-2\kappa_k\tau}}{\kappa_k} - 2\tau e^{-\kappa_k\tau}\right); \tag{5.3}$$

$$\mathcal{B}_k^{(3)}(\tau) = \left(a_k^i - a_k^j\right)^4 \rho_k^2\xi_k^2\left(\frac{1 - e^{-\kappa_k\tau}}{\kappa_k^3} - \frac{\tau e^{-\kappa_k\tau}}{\kappa_k^2} - \frac{\tau^2 e^{-\kappa_k\tau}}{2\kappa_k}\right) \tag{5.4}$$

and $\mathcal{A}_k^{(0)}, \mathcal{A}_k^{(1)}, \mathcal{A}_k^{(2)}, \mathcal{A}_k^{(3)}, k = 1, .., d$ as

$$\mathcal{A}_k^{(0)}(\tau) = \left(a_k^i - a_k^j\right)^2 \theta_k \left(\tau + \frac{e^{-\kappa_k \tau} - 1}{\kappa_k}\right); \tag{5.5}$$

$$\mathcal{A}_k^{(1)}(\tau) = \left(a_k^i - a_k^j\right)^3 \theta_k \rho_k \xi_k \left(\frac{\tau}{\kappa_k} + 2\frac{e^{-\kappa_k \tau} - 1}{\kappa_k^2} + \frac{\tau e^{-\kappa_k \tau}}{\kappa_k}\right); \tag{5.6}$$

$$\begin{aligned} \mathcal{A}_k^{(2)}(\tau) = \left(a_k^i - a_k^j\right)^4 \theta_k \rho_k^2 \xi_k^2 \Bigg( &\frac{\tau}{\kappa_k} + \frac{e^{-\kappa_k \tau} - 1}{\kappa_k^3} + \frac{\tau e^{-\kappa_k \tau}}{\kappa_k^2} \\ &- \frac{e^{-\kappa_k \tau} - 1}{\kappa_k^2} + \frac{\tau^2 e^{-\kappa_k \tau}}{2\kappa_k} - \frac{\tau e^{-\kappa_k \tau}}{\kappa_k} + \frac{e^{-\kappa_k \tau} - 1}{\kappa_k}\Bigg); \end{aligned} \tag{5.7}$$

$$\mathcal{A}_k^{(3)}(\tau) = \left(a_k^i - a_k^j\right)^4 \theta_k \rho_k^2 \xi_k^2 \left(\frac{\tau}{\kappa_k^2} + 3\frac{e^{-\kappa_k \tau} - 1}{\kappa_k^3} + 2\frac{\tau e^{-\kappa_k \tau}}{\kappa_k^2} + \frac{\tau^2 e^{-\kappa_k \tau}}{2\kappa_k}\right). \tag{5.8}$$

Finally, define the integrated variance as:

$$v = \sigma^2 \tau = \sum_{k=1}^{d} \big(\mathcal{A}_k^{(0)}(\tau) + \mathcal{B}_k^{(0)}(\tau) V_k\big). \tag{5.9}$$

PROPOSITION 4.4. *Assume that all vol-of-vol parameters $\xi_k, k = 1, .., d$ have been scaled by the same factor $\alpha > 0$. Then the call price $C(S(t), K, \tau)$ in the Multifactor Heston-based exchange model can be approximated in terms of the scale factor $\alpha$ by differentiating the Black Scholes formula $C_{\text{BS}}(S(t), K, \sigma, \tau)$ with respect to the log exchange rate $x(t) = \ln S(t)$ and the integrated variance $v = \sigma^2 \tau$:*

$$\begin{aligned} C(S(t), K, \tau) \approx\; & C_{\text{BS}}(S(t), K, \sigma, \tau) \\ & + \alpha \sum_{k=1}^{d} \left(\mathcal{A}_k^{(1)}(\tau) + \mathcal{B}_k^{(1)}(\tau) V_k\right) \partial_{xv}^2 C_{\text{BS}}(S(t), K, \sigma, \tau) \\ & + \alpha^2 \sum_{k=1}^{d} \left(\mathcal{A}_k^{(2)}(\tau) + \mathcal{B}_k^{(2)}(\tau) V_k\right) \partial_{vv}^2 C_{\text{BS}}(S(t), K, \sigma, \tau) \\ & + \alpha^2 \sum_{k=1}^{d} \left(\mathcal{A}_k^{(3)}(\tau) + \mathcal{B}_k^{(3)}(\tau) V_k\right) \partial_{xxv}^3 C_{\text{BS}}(S(t), K, \sigma, \tau) \\ & + \frac{\alpha^2}{2} \left[\sum_{k=1}^{d} \left(\mathcal{A}_k^{(1)}(\tau) + \mathcal{B}_k^{(1)}(\tau) V_k\right)\right]^2 \partial_{xxvv}^4 C_{\text{BS}}(S(t), K, \sigma, \tau). \end{aligned} \tag{5.10}$$

PROOF. See Appendix. □

We can now state another formula, which does not involve the computation of option prices and constitutes an approximation of the implied volatility surface for a short time to maturity. This formula may constitute a useful alternative in order to get a quicker calibration for short maturities. The proof is again provided in detail in the Appendix.

PROPOSITION 4.5. *For a short time to maturity the implied volatility expansion in terms of the vol-of-vol scale factor $\alpha$ in the multifactor Heston-based exchange model is given by:*

$$\sigma_{\mathrm{imp}}^2 \approx \sigma_0^2 + \alpha \left( \sum_{k=1}^{d} \frac{\rho_k \xi_k}{2} \left( a_k^i - a_k^j \right)^4 V_k \right) \frac{m_f}{\sigma_0^2}$$
$$+ \alpha^2 \frac{m_f^2}{12 \left(\sigma_0^2\right)^2} \left[ \sum_{k=1}^{d} \left(1 + 2\rho_k^2\right) \xi_k^2 \left( a_k^i - a_k^j \right)^4 V_k - \frac{15}{4\sigma_0^2} \left( \sum_{k=1}^{d} \rho_k \xi_k \left( a_k^i - a_k^j \right)^3 V_k \right)^2 \right],$$

*where* $\sigma_0^2 = \left(\mathbf{a}^i - \mathbf{a}^j\right)^\top \mathrm{Diag}(\mathbf{V}) \left(\mathbf{a}^i - \mathbf{a}^j\right)$ *and* $m_f = \log\left(\frac{S^{i,j} e^{(r^i - r^j)\tau}}{K^{i,j}}\right)$ *denotes the forward log-moneyness.*

PROOF. See Appendix. □

## 6. Simultaneous calibration of the USD/EUR/JPY triangle

**6.1. Setup.** In this section we show an example of simultaneous calibration to three market volatility surfaces of options: we consider the implied volatility surfaces for USD/EUR, USD/JPY and EUR/JPY as observed in the FX market on a typical day (data from 23rd July 2010), that is $N = 3$ with $i = \mathrm{USD};\ \mathrm{EUR};\ \mathrm{JPY}$. The volatility sample includes expiry dates ranging from 3 days to 5 years. The quotes follow the standard Delta quoting conversion in the FX option market, we have quotes on DN, 25 Delta, 15 Delta, and 10 Delta[2].

We try to fit simultaneously the three volatility surfaces using two stochastic drivers, $d = 2$. This choice yields a total number of parameters $N_P = 16$, comparable to the number of parameters in 3 independent Heston models (15 parameters). This choice should not lead to overfitting instabilities. We work under the USD risk neutral measure to derive the option prices of the pairs EUR/USD and USD/JPY and the EUR measure for the EUR/JPY options, using (4.2). We calibrate the CIR parameters in the USD measure $\kappa_k^{\mathrm{USD}}, \theta_k^{\mathrm{USD}}, \xi_k^{\mathrm{USD}}, \rho_k^{\mathrm{USD}},\ k = 1, 2$. The parameters for the EUR/JPY are transformed to the EUR measure through Eqs. (3.10) and the invariance property of correlation and vol-of-vol parameters.

The calibration is done via a standard non-linear least-squares optimizer that minimizes the total calibration error in terms of the difference between calibrated and target implied volatilities

$$\mathrm{Err} = \sum_n \left( \sigma_{n,\mathrm{market}}^{\mathrm{imp}} - \sigma_{n,\mathrm{model}}^{\mathrm{imp}} \right)^2. \tag{6.1}$$

The use of a norm in price should be avoided as the numerical range for option prices may be large, thus introducing a bias in the optimization (for a more detailed discussion, see Da Fonseca and Grasselli (2011)).

[2] It is important to stress that in the forex market implied volatilities surfaces are expressed in terms of maturity and Delta (see e.g. Wystup and Reiswich (2010), Clark (2011)): the market practice is to quote volatilities for strangles and risk reversals which can then be employed to reconstruct a whole surface of implied volatilities via an interpolation method (see e.g. Wystup and Reiswich (2010), Wystup (2006), Clark (2011)). Once we have the quotes in terms of Delta, to perform the calibration we have to convert Deltas into strike prices. The procedure can be found e.g. in Beneder and Elkenbracht-Huizing (2003).

**6.2. Calibration results.** In Table (1) we report the result of the calibration of the model. We performed the optimization considering different sets of expiries. The expiries considered in the largest sample are the following: 1, 2, 3, 6, 9 months and 1 year. The result for this particular choice of expiries is reported on the first column on the left. Then we repeated the experiment by excluding the largest expiry, 1 year. The result is reported in the second column. We proceed in this way by excluding more and more expiries. The smallest sample is reported in the last column and considers only options expyring in 1 and 2 months.
In Figs. 1, 2 and 3 we plot the market implied volatilities against those produced by the model. The plots refer to the largest sample in Table 1. Market volatilities are denoted by crosses, model volatilities are denoted by circles. The model yields a satisfactory fit of the market points. We are not aware of any other stochastic volatility model that can achieve a simultaneous calibration of a three-currency triangle with this accuracy.

The plots for the calibration on the sub-samples are completely analogous.

**6.3. Parameters stability tests.** In this subsection we comment on the stability of the parameters via two different types of analysis. We first measure the impact on the parameters resulting from the calibration procedure. Secondly, we fit the model parameters to a certain sample and then use these parameters to price an option which is not included in the sample. If the out-of-sample prices are close to the market, the model gives a reasonable description of the joint underlying FX rates dynamics. Moreover, the calibration can be done on a limited set of expiries, reducing the computation effort of the optimizer.

As far as the first analysis is concerned, we show in Table 2 the relative variations computed with respect to the largest sample. With the exception of $\kappa_1$ we can see that there is a good degree of stability of the parameters across the sub-samples. Since $\kappa_1$ is the only value which seems to fluctuate significantly, we perform also a second calibration experiment, where we fix $\kappa_k = 1, k = 1, 2$. The results of this experiment are outlined in Table 3. The relative variation of the parameters can be found in Table 4. We notice that with this choice we get a good degree of stability, the most relevant fluctuation is now around 20% for $\theta_1$[3].

Let us now turn our attention to the out-of-sample exercise. In Tables 6, 7, 8 and 9 we show the difference between the market and the out-of-sample volatility for all sub-samples. The differences are always well below one volatility point, hence satisfactory.

**6.4. Feller condition.** It is well known, see e.g. Feller (1951); Andersen and Piterbarg (2007), that for square root processes 0 represents an attainable state when the Feller condition is not satisfied, that is when $2\kappa_k\theta_k < \xi_k^2$. In our modelling framework we have two volatility factors, hence we can perform the check for each factor. In Table 5 we report the quantity $F_k = 2\kappa_k\theta_k - \xi_k^2, k = 1, 2$. We observe violations of the Feller condition, which constitutes a well-known fact in the FX derivative practice, shared with the standard one-dimensional Heston model, see Clark (2011).

[3]We do not report, for the sake of brevity, the volatility surfaces arising from this last experiment, but the quality of the fit is the same as before.

**6.5. Moment explosions.** Another well established fact is that stochastic volatility models often suffer of pathological moment explosions which might often impact the stability of the pricing tools, see e.g. Andersen and Piterbarg (2007), Keller-Ressel (2011) and Glasserman and Kim (2011). The model dynamics might lead to the explosion of moments, which become infinite in finite time.

Qualitatively, the present model shares with the Heston model the same type of singularities. Let us define the explosion time of the $n$-th moment as:

$$T_\star(n) = \sup\left\{t > 0 : \mathbb{E}\left[(S^{i,j}(t))^n\right] < \infty\right\}. \tag{6.2}$$

In our 2-Heston model, the $n$-th moment $m_n(t) = \mathbb{E}\left[\left(S^{i,j}(t)\right)^n\right]$ is computed by (4.3) as follows:

$$m_n(t) = G^{i,j}(n, 0, t, x, \mathbf{V}) \tag{6.3}$$

$$= (S^{i,j}(0))^n e^{n(r^i - r^j)t} \exp\left\{A_1^{i,j}(t) + B_1^{i,j}(t)V_1(0)\right\} \tag{6.4}$$

$$. \exp\left\{A_2^{i,j}(t) + B_2^{i,j}(t)V_2(0)\right\} \tag{6.5}$$

where the functions $A_k^{i,j}(t), B_k^{i,j}(t)$ are given in Proposition 4.3 with $\omega = n$. Hence, the moments can be decomposed in a product of one-dimensional moment generating functions. As a result, the moment explosion properties follow directly from the known results of the one-dimensional Heston model.

In complete analogy with Keller-Ressel (2011) (Section 6.1), from the Riccati ODE (8.9) for each factor $k = 1, 2$, we define the quantities $\Delta_k(n)$ and $\chi_k(n)$ as follows:

$$\Delta_k(n) = \left(\kappa_k - \omega\left(a_k^i - a_k^j\right)\rho_k\xi_k\right)^2 - (\omega^2 - \omega)\left(a_k^i - a_k^j\right)^2; \tag{6.6}$$

$$\chi_k(n) = -\kappa_k + \omega\left(a_k^i - a_k^j\right)\rho_k\xi_k. \tag{6.7}$$

We can then apply the moment explosion formula and obtain the following

PROPOSITION 4.6. *In the multi-Heston model, $T_\star(n)$ is given by:*

1. $T_\star(n) = +\infty$, *if $\Delta_k(n) \geq 0, \forall k = 1, 2$;*
2. $T_\star(n) = \inf_k \left\{\frac{2}{\sqrt{-\Delta_k(n)}}\left(\arctan\left(\frac{\sqrt{-\Delta_k(n)}}{\chi_k(n)}\right) + \pi\mathbf{1}_{\chi_k(n)<0}\right)\right\}$, *if $\Delta_k(n) < 0$, for some $k = 1, 2$.*

Then we can easily calculate the time of moment explosion for all currency pairs. In Table 14 we consider moments up to order 20.

**6.6. Implied risk-neutral distribution.** As a result of our calibration procedure, we have a set of model parameters allowing us to compute option prices which are in line with market data. The implied risk neutral probability densities of the different FX rates can be derived from option prices (cfr e.g. Breeden and Litzenberger (1978))

$$\varphi_t(S) = e^{r_i(T-t)} \left.\frac{\partial^2 C(S(t), K, T)}{\partial K^2}\right|_{K=S}. \tag{6.8}$$

The second derivative is approximated given a grid of option prices. We show the results in Figs. 10, 11 and 12. Notice that we plot the raw distribution resulting from prices, thus some numerical treatment is required in order to get a smooth distribution. All distributions feature asymmetry and leptokurtosis, features that are standard for returns in financial series.

## 7. Conclusions

We have introduced a new multi-factor stochastic volatility Heston-based model that can provide an accurate joint description of multiple FX vanilla options across different currency pairs. The emphasis in the model specification has been in the preservation of the specific symmetries of FX markets. Differently from other asset classes, appropriate multiplications/divisions and inversions of FX rates are still FX rates. Having a model that is functionally invariant with respect to these operations has immediate benefits: the calibrations are by construction universal, i.e., independent on the choice of the risk-free rate of the investor, leading to consistent prices and risk-sensitivities for multi-dimensional exotic options.

The choice of the CIR dynamics for the stochastic variance is instrumental in achieving this symmetry. We have indeed proven that our model is invariant with respect to the choice of the numéraire once the model parameters are appropriately transformed. The model is always of affine-type independently on which currency is used as risk free, leading to semi-analytical expression for all vanilla options between *any* of two currencies. This property is crucial when it comes to calibrating the model. In a standard global optimization algorithm we can consider together vanilla options in all currency pairs and achieve a simultaneous fit to the different volatility surfaces with reasonable computational effort.

The model shares naturally several stylized facts with the Heston model. The Feller condition is often violated when fitting the model to FX volatility surfaces, a common observation in the practice. Moreover, higher moments of the spot distribution explode at finite time; a property that might lead to complications/instabilities in standard numerical pricing routines. Finally, like any pure stochastic volatility model, our model cannot be expected to deliver a perfect calibration of the vanilla surfaces across all Deltas and tenors, especially in the short end.

Having said that, the main result of the paper is a promising joint calibration of the model to the implied volatilities smile of the EUR/USD/JPY FX triangle. The fit remains satisfactory across the currency pairs, Deltas and tenors which were considered. Several in- and out-of-sample calibration studies in fact have proven the robustness of the calibration, especially once the mean reversion speed $\kappa$ has been fixed. Asymptotic expansions of the implied volatility surface are also included as they shed light on the meaning of the different model parameters and can help speeding up the calibration procedure.

The price to pay in order to obtain a consistent simultaneous calibration to all volatilities surfaces is that the instantaneous volatilities of the currency pairs do not have single dedicated drivers. Their dynamics is rather brought about by a linear combination of several hidden stochastic factors. As in any principal component analysis, it is not easy to assign a financial meaning to each model parameter. As this study has shown, this appealing feature has most likely to be traded away in order to capture the complex phenomenology of the present global and widely interconnected FX markets.

## 8. Appendix A: Proofs

**8.1. Proof of Proposition** (4.2)**.** We will apply Theorem 2.1 from Mijatović and Urusov (2011): using their notation we have $Y_k = V_k$ and $b_k(x) = -\rho_k x_k^{1/2} a_k^i$, where we recall that by independency of the $V_k$ processes the Radon-Nykodim derivative splits in a product and we can use a separability argument.

If the Feller condition holds, that is when $2\kappa_k\theta_k \geq \xi_k^2$, then conditions a) and c) in Theorem 2.1 of Mijatović and Urusov (2011) hold, as from (3.7) we note that the boundary behavior of $V_k$ is the same under both $\mathbb{Q}^0$ and $\mathbb{Q}^i$.

Suppose now that the Feller condition is violated. Condition a) still holds while condition c) does not, but the boundary $l = 0$ is good. To see this let us compute the functions $\rho_k(x), s_k(x)$ (given by formulae (16),(18) in Mijatović and Urusov (2011)) in our framework:

$$\rho_k(x) = \alpha x^{-\frac{2\kappa_k\theta_k}{\xi_k^2}} \exp(\frac{2\kappa_k x}{\xi_k^2}),$$

where $\alpha$ denotes a positive constant, and

$$s_k(x) = \int_c^x \rho(y)dy = \alpha \int_c^x y^{-\frac{2\kappa_k\theta_k}{\xi_k^2}} \exp(\frac{2\kappa_k y}{\xi_k^2})dy,$$

where $c$ denotes another arbitrary constant. We can now confirm using (26) in Mijatović and Urusov (2011), that 0 is good: in fact we have $s_k(0) > -\infty$ (by the assumption $2\kappa_k\theta_k < \xi_k^2$) and the function

$$\frac{(s_k(x) - s_k(0))b_k^2(x)}{\rho_k(x)\sigma_k^2(x)} = \alpha(s_k(x) - s_k(0))x^{\frac{2\kappa_k\theta_k}{\xi_k^2}} \exp(\frac{2\kappa_k x}{\xi_k^2})$$

is locally integrable around 0 since $s_k(x)$ behaves like $x^{-\frac{2\kappa_k\theta_k}{\xi_k^2}+1}$. Hence (26) holds and it follows that the Radon-Nykodim derivative (3.3) is a true martingale.

**8.2. Proof of Proposition** (4.3)**.** Recall that $\phi(\omega, t, \tau, x, \mathbf{V}) = G(\mathtt{i}\omega, t, \tau, x, \mathbf{V})$, with $x^{i,j}(t) = \log S^{i,j}(t)$. The functions $\phi, G$ represent resp. the characteristic function and the moment generating function of the log-exchange rate. In order to determine these quantities, we first need to write the PDE satisfied by $G$. First of all we write the dynamics of $x = x^{i,j}$:

$$\begin{aligned} dx(t) = &\left( \left(r^i - r^j\right) - \frac{1}{2}(\mathbf{a}^i - \mathbf{a}^j)^\top \mathrm{Diag}\left(\mathbf{V}(t)\right)\left(\mathbf{a}^i - \mathbf{a}^j\right) \right) dt \\ &+ \left(\mathbf{a}^i - \mathbf{a}^j\right)^\top \mathrm{Diag}\left(\sqrt{\mathbf{V}(t)}\right) d\mathbf{Z}^{\mathbb{Q}^i}(t). \end{aligned} \tag{8.1}$$

We also compute the following covariation terms for $k = 1, .., d$:

$$\begin{aligned} d\left\langle x, V_k \right\rangle_t &= \left\langle \left(\mathbf{a}^i - \mathbf{a}^j\right)^\top \mathrm{Diag}\left(\sqrt{\mathbf{V}}\right) d\mathbf{Z}^{\mathbb{Q}^i}, \xi_k \sqrt{V_k}\rho_k dZ_k^{\mathbb{Q}^i} \right\rangle_t \\ &= \left\langle \left(a_k^i - a_k^j\right) \sqrt{V_k} dZ_k^{\mathbb{Q}^i}, \xi_k \sqrt{V_k}\rho_k dZ_k^{\mathbb{Q}^i} \right\rangle_t \\ &= \left(a_k^i - a_k^j\right) V_k(t)\xi_k\rho_k dt. \end{aligned} \tag{8.2}$$

The Laplace transform $G$ solves the following backward Kolmogorov equation (Karatzas and Shreve (1991)):

$$\begin{aligned} &\frac{\partial G}{\partial t} + \frac{1}{2}\frac{\partial^2 G}{\partial x^2}\left(\mathbf{a}^i - \mathbf{a}^j\right)^\top \mathrm{Diag}\left(\mathbf{V}\right)\left(\mathbf{a}^i - \mathbf{a}^j\right) \\ &\sum_{k=1}^d \frac{\partial^2 G}{\partial x \partial V_k}\left(a_k^i - a_k^j\right) V_k\xi_k\rho_k + \frac{1}{2}\sum_{k=1}^d \frac{\partial^2 G}{\partial V_k^2}\xi_k^2 V_k \\ &\left( \left(r^i - r^j\right) - \frac{1}{2}\left(\mathbf{a}^i - \mathbf{a}^j\right)^\top \mathrm{Diag}\left(\mathbf{V}\right)\left(\mathbf{a}^i - \mathbf{a}^j\right) \right) \frac{\partial G}{\partial x} \\ &+ \sum_{k=1}^d \frac{\partial G}{\partial V_k}\kappa_k\left(\theta_k - V_k\right) = 0 \end{aligned} \tag{8.3}$$

with terminal condition $G(\omega, T, 0, x, \mathbf{V}) = e^{\omega x}$ with $\omega \in \mathbb{R}$. In order to solve this problem we look for an exponential affine solution of the form:

$$G(\omega, t, \tau, x, \mathbf{V}) = \exp\left(A(t,T) + \sum_{k=1}^{d} B_k(t,T)V_k + C(t,T)x\right), \tag{8.4}$$

for some deterministic functions $A, B_k, C$ that may depend on both $t, T$. Upon substitution of the guess and recognition of the terms we obtain the following system of $d+2$ ODE's:

$$\frac{\partial A}{\partial t} + \sum_{k=1}^{d} B_k(t,T)\kappa_k\theta_k + \left(r^i - r^j\right)C(t,T) = 0; \tag{8.5}$$

$$\begin{aligned}&\frac{\partial B_k}{\partial t} + \frac{1}{2}C^2(t,T)\left(a_k^i - a_k^j\right)^2 + C(t,T)B_k(t,T)\left(a_k^i - a_k^j\right)\rho_k\xi_k \\ &+ \frac{1}{2}B_k^2(t,T)\xi_k^2 - \frac{1}{2}\left(a_k^i - a_k^j\right)^2 C(t,T) - B_k(t,T)\kappa_k = 0;\end{aligned} \tag{8.6}$$

$$\frac{\partial C}{\partial t} = 0, \tag{8.7}$$

with terminal conditions: $A(T,T) = 0, \quad B_k(T,T) = 0, \quad C(T,T) = \omega$ for $k = 1, .., d$. From (8.7) and its terminal condition, we deduce that $C(t,T) = \omega$ for $t \in [0,T]$, so we can rewrite the system as follows:

$$\frac{\partial A}{\partial t} + \sum_{k=1}^{d} \kappa_k\theta_k B_k(t,T) + \left(r^i - r^j\right)\omega = 0; \tag{8.8}$$

$$\begin{aligned}&\frac{\partial B_k}{\partial t} + \frac{1}{2}B_k^2(t,T)\xi_k^2 + \left(-\kappa_k + \omega\left(a_k^i - a_k^j\right)\rho_k\xi_k\right)B_k(t,T) \\ &+ \frac{\omega^2 - \omega}{2}\left(a_k^i - a_k^j\right)^2 = 0, \quad k = 1, .., d.\end{aligned} \tag{8.9}$$

Now for $d = 1, .., d$ we assume that $B_k(t,T)$ can be written by means of a function $E_k(t,T)$ and set:

$$B_k(t,T) = \frac{\frac{\partial}{\partial t}E_k(t,T)}{\frac{\xi_k^2}{2}E_k(t,T)}, \tag{8.10}$$

Now we proceed to solve (8.9) which is a Riccati ODE. Write the following:

$$\frac{\partial B_k}{\partial t} = \frac{\frac{\partial^2}{\partial t^2}E_k(t,T)\frac{\xi_k^2}{2}E_k(t,T) - \left(\frac{\partial}{\partial t}E_k(t,T)\right)^2\frac{\xi_k^2}{2}}{\left(\frac{\xi_k^2}{2}E_k(t,T)\right)^2} \tag{8.11}$$

If we substitute (8.10) and (8.11) into (8.9) we obtain a second order ODE:

$$\frac{\partial^2 E_k}{\partial t^2} + \left(-\kappa_k + \omega\left(a_k^i - a_k^j\right)\rho_k\xi_k\right)\frac{\partial E_k}{\partial t} + \frac{\xi_k^2\left(\omega^2 - \omega\right)}{4}\left(a_k^i - a_k^j\right)^2 E_k(t,T) = 0 \tag{8.12}$$

We look for a solution of the form $E_k(t,T) = e^{\lambda(T-t)}$. Substitution of the guess yields the following:

$$\lambda^2 - \lambda\left(-\kappa_k + \omega\left(a_k^i - a_k^j\right)\rho_k\xi_k\right) + \frac{\xi_k^2\left(\omega^2 - \omega\right)}{4}\left(a_k^i - a_k^j\right)^2 = 0 \tag{8.13}$$

Now define:

$$\Delta_k = \left(-\kappa_k + \omega\left(a_k^i - a_k^j\right)\rho_k\xi_k\right)^2 - \xi_k^2\left(\omega^2 - \omega\right)\left(a_k^i - a_k^j\right)^2 \tag{8.14}$$

$$\lambda_k^{\pm} = \frac{\left(-\kappa_k + \omega\left(a_k^i - a_k^j\right)\rho_k\xi_k\right) \pm \sqrt{\Delta_k}}{2} \tag{8.15}$$

Notice that:

$$\lambda_k^+ - \lambda_k^- = \sqrt{\Delta_k} \tag{8.16}$$

$$\lambda_k^+ \lambda_k^- = \frac{\xi_k^2 \left(\omega^2 - \omega\right)}{4} \left(a_k^i - a_k^j\right)^2 \tag{8.17}$$

Hence:

$$E_k(t,T) = C_+ e^{\lambda_k^+ (T-t)} + C_- e^{\lambda_k^- (T-t)} \tag{8.18}$$

Recalling that $B(T,T) = 0$ we obtain that:

$$\frac{C_+}{C_-} = -\frac{\lambda_k^-}{\lambda_k^+} \tag{8.19}$$

so that

$$E_k(t,T) = -\lambda_k^- e^{\lambda_k^+ (T-t)} + \lambda_k^+ e^{\lambda_k^- (T-t)} \tag{8.20}$$

Recalling the guess for $B_k(t,T)$ with some algebra we get the solution for (8.9), which is:

$$B_k(t,T) = \frac{\left(\omega^2 - \omega\right)}{2} \left(a_k^i - a_k^j\right)^2 \frac{1 - e^{-\sqrt{\Delta_k}(T-t)}}{\lambda_k^+ e^{-\sqrt{\Delta_k}(T-t)} - \lambda_k^-}. \tag{8.21}$$

Equipped with the solution for $B_k(t,T)$ we can now compute $A(t,T)$ as follows:

$$\begin{aligned} A(T,T) - A(t,T) &= \int_t^T \frac{\partial}{\partial u} A(u,T) du \\ A(t,T) &= \int_t^T \sum_{k=1}^d \kappa_k \theta_k B_k(u,T) + \left(r^i - r^j\right) \omega du \\ &= \left(r^i - r^j\right) \omega (T-t) + \sum_{k=1}^d \kappa_k \theta_k \int_t^T B_k(u,T) du \\ &= \left(r^i - r^j\right) \omega (T-t) + \sum_{k=1}^d \frac{2\kappa_k \theta_k}{\xi_k^2} \int_t^T \frac{\frac{\partial}{\partial t} E_k(t,T)}{E_k(t,T)} du, \end{aligned} \tag{8.22}$$

which implies that the solution for $A(t,T)$ is

$$\begin{aligned} A(t,T) &= \left(r^i - r^j\right) \omega (T-t) + \sum_{k=1}^d \frac{2\kappa_k \theta_k}{\xi_k^2} \log \frac{\lambda_k^+ - \lambda_k^-}{\lambda_k^+ e^{\lambda_k^- (T-t)} - \lambda_k^- e^{\lambda_k^+ (T-t)}} \\ &= \left(r^i - r^j\right) \omega (T-t) + \sum_{k=1}^d A_k(t,T), \end{aligned} \tag{8.23}$$

where the functions $A_k(t,T)$ are implicitly defined by the last equality for $d = 1,..,d$. Now we obtain the statement of the proposition once we replace $B_k^{i,j}(\tau) = B_k(t,T), A_k^{i,j}(\tau) = A_k(t,T)$ with $\tau = T - t$.

**8.3. Proof of Proposition** (4.4)**.** The starting point is given by the Riccati ODE (8.9) expressed in terms of time-to-maturity $\tau = T - t$ and perturbed by introducing the vol-of-vol scale parameter $\alpha$:

$$\begin{aligned} \frac{\partial B_k}{\partial \tau} &= \frac{1}{2} B_k^2(\tau) \alpha^2 \xi_k^2 + \left(-\kappa_k + \omega \left(a_k^i - a_k^j\right) \rho_k \alpha \xi_k\right) B_k(\tau) \\ &+ \frac{\omega^2 - \omega}{2} \left(a_k^i - a_k^j\right)^2, \quad k = 1,..,d. \end{aligned} \tag{8.24}$$

We consider the following expansion in terms of $\alpha$: $B_k(\tau) = B_{k,0}(\tau) + \alpha B_{k,1}(\tau) + \alpha^2 B_{k,2}(\tau)$. By plugging in the expansion and upon recognition of terms we obtain the following system of ODE's:

$$\frac{\partial B_{k,0}}{\partial \tau} = -\kappa_k B_{k,0}(\tau) + \frac{\omega^2 - \omega}{2}\left(a_k^i - a_k^j\right)^2; \tag{8.25}$$

$$\frac{\partial B_{k,1}}{\partial \tau} = -\kappa_k B_{k,1}(\tau) + \omega\left(a_k^i - a_k^j\right)\rho_k \xi_k B_{k,0}(\tau); \tag{8.26}$$

$$\frac{\partial B_{k,2}}{\partial \tau} = -\kappa_k B_{k,2}(\tau) + \omega\left(a_k^i - a_k^j\right)\rho_k \xi_k B_{k,1}(\tau) + \frac{1}{2}B_{k,0}^2(\tau)\xi_k^2. \tag{8.27}$$

If we denote $\gamma := \frac{\omega^2-\omega}{2}$ then the solutions are easily computed as:

$$\begin{aligned} B_{k,0}(\tau) &= \underbrace{B_{k,0}(0)}_{=0} e^{-\kappa_k \tau} + e^{-\kappa_k \tau}\int_0^\tau e^{\kappa_k u}\gamma\left(a_k^i - a_k^j\right)^2 du \\ &= \gamma \mathcal{B}_k^{(0)}(\tau); \end{aligned} \tag{8.28}$$

$$\begin{aligned} B_{k,1}(\tau) &= \underbrace{B_{k,1}(0)}_{=0} e^{-\kappa_k \tau} + e^{-\kappa_k \tau}\int_0^\tau e^{\kappa_k u}\omega\left(a_k^i - a_k^j\right)\rho_k \xi_k \gamma \mathcal{B}_k^{(0)}(u) du \\ &= \omega\gamma \mathcal{B}_k^{(1)}(\tau); \end{aligned} \tag{8.29}$$

$$\begin{aligned} B_{k,2}(\tau) &= \underbrace{B_{k,2}(0)}_{=0} e^{-\kappa_k \tau} + e^{-\kappa_k \tau}\int_0^\tau e^{\kappa_k u}\omega^2\gamma\left(a_k^i - a_k^j\right)\rho_k \xi_k \mathcal{B}_k^{(1)}(u) du \\ &+ e^{-\kappa_k \tau}\int_0^\tau e^{\kappa_k u}\gamma\frac{\xi_k^2}{2}\left(\mathcal{B}_k^{(0)}(u)\right)^2 du \\ &= \omega^2\gamma \mathcal{B}_k^{(3)}(\tau) + \gamma^2 \mathcal{B}_k^{(2)}(\tau). \end{aligned} \tag{8.30}$$

Then we can write the function $B_k(\tau)$ as follows:

$$B_k(\tau) = \gamma \mathcal{B}_k^{(0)}(\tau) + \alpha\omega\gamma \mathcal{B}_k^{(1)}(\tau) + \alpha^2\left(\omega^2\gamma \mathcal{B}_k^{(3)}(\tau) + \gamma^2 \mathcal{B}_k^{(2)}(\tau)\right). \tag{8.31}$$

A direct substitution of (8.31) into (8.8) allows us to express the function $A(\tau)$:

$$\begin{aligned} A(\tau) &= \omega\left(r^i - r^j\right)\tau + \sum_{k=1}^d \kappa_k\theta_k\int_0^\tau B_k(u)du \\ &= \omega\left(r^i - r^j\right)\tau + \gamma\sum_{k=1}^d \underbrace{\kappa_k\theta_k\int_0^\tau \mathcal{B}_k^{(0)}(u)du}_{:=\mathcal{A}_k^{(0)}(\tau)} + \omega\gamma\alpha\sum_{k=1}^d \underbrace{\kappa_k\theta_k\int_0^\tau \mathcal{B}_k^{(1)}(u)du}_{:=\mathcal{A}_k^{(1)}(\tau)} \\ &+ \omega^2\gamma\alpha^2\sum_{k=1}^d \underbrace{\kappa_k\theta_k\int_0^\tau \mathcal{B}_k^{(3)}(u)du}_{:=\mathcal{A}_k^{(3)}(\tau)} + \alpha^2\gamma^2\sum_{k=1}^d \underbrace{\kappa_k\theta_k\int_0^\tau \mathcal{B}_k^{(2)}(u)du}_{:=\mathcal{A}_k^{(2)}(\tau)}. \end{aligned} \tag{8.32}$$

We consider then the price in terms of Fourier transform as in (4.2) by replacing the argument $\omega = \mathtt{i}\lambda$. A Taylor-McLaurin expansion w.r.t. $\alpha$ gives the following:

$$\begin{aligned}
C(S(t),K,\tau) \approx & \frac{e^{-r^i\tau}}{2\pi}\int_{\mathcal{Z}} e^{\mathtt{i}\lambda\left(r^i-r^j\right)\tau+\mathtt{i}\lambda x+\gamma\sum_{k=1}^{d}\left(\mathcal{A}_k^{(0)}(\tau)+\mathcal{B}_k^{(0)}(\tau)V_k\right)}\Phi(\lambda)d\lambda \\
& + \alpha\sum_{k=1}^{d}\left(\mathcal{A}_k^{(1)}(\tau)+\mathcal{B}_k^{(1)}(\tau)V_k\right) \\
& \times \frac{e^{-r^i\tau}}{2\pi}\int_{\mathcal{Z}} \gamma\mathtt{i}\lambda e^{\mathtt{i}\lambda\left(r^i-r^j\right)\tau+\mathtt{i}\lambda x+\gamma\sum_{k=1}^{d}\left(\mathcal{A}_k^{(0)}(\tau)+\mathcal{B}_k^{(0)}(\tau)V_k\right)}\Phi(\lambda)d\lambda \\
& + \alpha^2\sum_{k=1}^{d}\left(\mathcal{A}_k^{(2)}(\tau)+\mathcal{B}_k^{(2)}(\tau)V_k\right) \\
& \times \frac{e^{-r^i\tau}}{2\pi}\int_{\mathcal{Z}} \gamma^2 e^{\mathtt{i}\lambda\left(r^i-r^j\right)\tau+\mathtt{i}\lambda x+\gamma\sum_{k=1}^{d}\left(\mathcal{A}_k^{(0)}(\tau)+\mathcal{B}_k^{(0)}(\tau)V_k\right)}\Phi(\lambda)d\lambda \\
& + \alpha^2\sum_{k=1}^{d}\left(\mathcal{A}_k^{(3)}(\tau)+\mathcal{B}_k^{(3)}(\tau)V_k\right) \\
& \times \frac{e^{-r^i\tau}}{2\pi}\int_{\mathcal{Z}} \gamma\mathtt{i}\lambda^2 e^{\mathtt{i}\lambda\left(r^i-r^j\right)\tau+\mathtt{i}\lambda x+\gamma\sum_{k=1}^{d}\left(\mathcal{A}_k^{(0)}(\tau)+\mathcal{B}_k^{(0)}(\tau)V_k\right)}\Phi(\lambda)d\lambda \\
& + \frac{\alpha^2}{2}\left[\sum_{k=1}^{d}\left(\mathcal{A}_k^{(1)}(\tau)+\mathcal{B}_k^{(1)}(\tau)V_k\right)\right]^2 \\
& \times \frac{e^{-r^i\tau}}{2\pi}\int_{\mathcal{Z}} \gamma^2\mathtt{i}\lambda^2 e^{\mathtt{i}\lambda\left(r^i-r^j\right)\tau+\mathtt{i}\lambda x+\gamma\sum_{k=1}^{d}\left(\mathcal{A}_k^{(0)}(\tau)+\mathcal{B}_k^{(0)}(\tau)V_k\right)}\Phi(\lambda)d\lambda.
\end{aligned} \tag{8.33}$$

Recall now from (5.9) the definition of the integrated Black-Scholes variance. In the previous formula, in the first term on the right hand side, we recognise the Black-Scholes price in terms of the characteristic function when the integrated variance is $v = \sigma^2\tau$:

$$C_{\text{BS}}\left(S(t),K,\sigma,\tau\right) = \frac{e^{-r^i\tau}}{2\pi}\int_{\mathcal{Z}} e^{\mathtt{i}\lambda\left(r^i-r^j\right)\tau+\mathtt{i}\lambda x+\frac{(\mathtt{i}\lambda)^2-\mathtt{i}\lambda}{2}v}\Phi(\lambda)d\lambda, \tag{8.34}$$

so that the price expansion is of the form:

$$\begin{aligned}
C(S(t),K,\tau) \approx & C_{\text{BS}}\left(S(t),K,\sigma,\tau\right) \\
& + \alpha\sum_{k=1}^{d}\left(\mathcal{A}_k^{(1)}(\tau)+\mathcal{B}_k^{(1)}(\tau)V_k\right)\partial^2_{xv}C_{\text{BS}}\left(S(t),K,\sigma,\tau\right) \\
& + \alpha^2\sum_{k=1}^{d}\left(\mathcal{A}_k^{(2)}(\tau)+\mathcal{B}_k^{(2)}(\tau)V_k\right)\partial^2_{vv}C_{\text{BS}}\left(S(t),K,\sigma,\tau\right) \\
& + \alpha^2\sum_{k=1}^{d}\left(\mathcal{A}_k^{(3)}(\tau)+\mathcal{B}_k^{(3)}(\tau)V_k\right)\partial^3_{xxv}C_{\text{BS}}\left(S(t),K,\sigma,\tau\right) \\
& + \frac{\alpha^2}{2}\left[\sum_{k=1}^{d}\left(\mathcal{A}_k^{(1)}(\tau)+\mathcal{B}_k^{(1)}(\tau)V_k\right)\right]^2\partial^4_{xxvv}C_{\text{BS}}\left(S(t),K,\sigma,\tau\right).
\end{aligned} \tag{8.35}$$

From the previous expression we can deduce the relation defining the deterministic functions $\mathcal{B}_k^{(h)}, \mathcal{A}_k^{(h)},\ h = 0, ..., 3$.

$$\mathcal{B}_k^{(0)}(\tau) = \left(a_k^i - a_k^j\right)^2 \frac{1 - e^{-\kappa_k \tau}}{\kappa_k}; \tag{8.36}$$

$$\mathcal{B}_k^{(1)}(\tau) = \left(a_k^i - a_k^j\right) \rho_k \xi_k e^{-\kappa_k \tau} \int_0^\tau e^{\kappa_k u} \mathcal{B}_k^{(0)}(u) du; \tag{8.37}$$

$$\mathcal{B}_k^{(2)}(\tau) = \frac{\xi_k^2}{2\kappa_k} e^{-\kappa_k \tau} \int_0^\tau e^{\kappa_k u} \left(\mathcal{B}_k^{(0)}(u)\right)^2 du; \tag{8.38}$$

$$\mathcal{B}_k^{(3)}(\tau) = \left(a_k^i - a_k^j\right) \rho_k \xi_k e^{-\kappa_k \tau} \int_0^\tau e^{\kappa_k u} \mathcal{B}_k^{(1)}(u) du \tag{8.39}$$

and

$$\mathcal{A}_k^{(0)}(\tau) = \kappa_k \theta_k \int_0^\tau \mathcal{B}_k^{(0)}(u) du; \tag{8.40}$$

$$\mathcal{A}_k^{(1)}(\tau) = \kappa_k \theta_k \int_0^\tau \mathcal{B}_k^{(1)}(u) du; \tag{8.41}$$

$$\mathcal{A}_k^{(2)}(\tau) = \kappa_k \theta_k \int_0^\tau \mathcal{B}_k^{(2)}(u) du; \tag{8.42}$$

$$\mathcal{A}_k^{(3)}(\tau) = \kappa_k \theta_k \int_0^\tau \mathcal{B}_k^{(3)}(u) du. \tag{8.43}$$

Computing the trivial integrals completes the proof.

**8.4. Proof of Proposition** (4.5)**.** We follow the procedure in Da Fonseca and Grasselli (2011). We suppose an expansion for the integrated implied variance of the form $v = \sigma_{\mathrm{imp}}^2 \tau = \zeta_0 + \alpha \zeta_1 + \alpha^2 \zeta_2$ and we consider the Black-Scholes formula as a function of the integrated implied variance and the log exchange rate $x = \log S$: $C_{\mathrm{BS}}\left(S(t), K, \sigma, \tau\right) = C_{\mathrm{BS}}\left(x(t), K, \sigma_{imp}^2 \tau, \tau\right)$. A Taylor-McLaurin expansion gives us the following:

$$\begin{aligned} C_{\mathrm{BS}}\left(x(t), K, \sigma_{imp}^2 \tau, \tau\right) &= C_{\mathrm{BS}}\left(x(t), K, \zeta_0, \tau\right) + \alpha \zeta_1 \partial_v C_{\mathrm{BS}}\left(x(t), K, \zeta_0, \tau\right) \\ &+ \frac{\alpha^2}{2}\left(2\zeta_2 \partial_v C_{\mathrm{BS}}\left(x(t), K, \zeta_0, \tau\right) + \zeta_1^2 \partial_{v^2}^2 C_{\mathrm{BS}}\left(x(t), K, \zeta_0, \tau\right)\right). \end{aligned} \tag{8.44}$$

By comparing this with the price expansion (8.35) we deduce that the coefficients must be of the form:

(8.45)

$$\zeta_0 = v_0;$$

(8.46)

$$\zeta_1 = \frac{\sum_{k=1}^d \left(\mathcal{A}_k^{(1)}(\tau) + \mathcal{B}_k^{(1)}(\tau) V_k\right) \partial_{xv}^2 C_{\mathrm{BS}}}{\partial_v C_{\mathrm{BS}}};$$

$$\begin{aligned} \zeta_2 &= \frac{-\zeta_1^2 \partial_{vv}^2 C_{\mathrm{BS}} + 2\sum_{k=1}^d \left(\mathcal{A}_k^{(2)}(\tau) + \mathcal{B}_k^{(2)}(\tau) V_k\right) \partial_{vv}^2 C_{\mathrm{BS}}}{2\partial_v C_{\mathrm{BS}}} \\ &+ \frac{2\sum_{k=1}^d \left(\mathcal{A}_k^{(3)}(\tau) + \mathcal{B}_k^{(3)}(\tau) V_k\right) \partial_{xxv}^3 C_{\mathrm{BS}} + \left[\sum_{k=1}^d \left(\mathcal{A}_k^{(1)}(\tau) + \mathcal{B}_k^{(1)}(\tau) V_k\right)\right]^2 \partial_{xxvv}^4 C_{\mathrm{BS}}}{2\partial_v C_{\mathrm{BS}}}, \end{aligned} \tag{8.47}$$

where the Black-Scholes formula $C_{\text{BS}}\left(x(t), K, \sigma^2_{imp}\tau, \tau\right)$ is evaluated at the point $(x, K, v_0, \tau)$. In order to find the values of $\zeta_1, \zeta_2$, we differentiate (5.1)-(5.4) thus obtaining the following ODE's:

$$\begin{aligned}
\frac{\partial \mathcal{B}^{(0)}_k}{\partial \tau} &= -\kappa_k \mathcal{B}^{(0)}_k(\tau) + \left(a^i_k - a^j_k\right)^2;\\
\frac{\partial \mathcal{B}^{(1)}_k}{\partial \tau} &= -\kappa_k \mathcal{B}^{(1)}_k(\tau) + \left(a^i_k - a^j_k\right)\rho_k \xi_k \mathcal{B}^{(0)}_k(\tau);\\
\frac{\partial \mathcal{B}^{(2)}_k}{\partial \tau} &= -\kappa_k \mathcal{B}^{(2)}_k(\tau) + \frac{1}{2}\xi^2_k \mathcal{B}^{(0)}_k(\tau)^2;\\
\frac{\partial \mathcal{B}^{(3)}_k}{\partial \tau} &= -\kappa_k \mathcal{B}^{(3)}_k(\tau) + \left(a^i_k - a^j_k\right)\rho_k \xi_k \mathcal{B}^{(1)}_k(\tau).
\end{aligned}$$

We consider a Taylor-McLaurin expansion in terms of $\tau$:

$$\mathcal{B}^{(0)}_k(\tau) = \left(a^i_k - a^j_k\right)^2 \tau - \frac{\tau^2}{2}\kappa_k \left(a^i_k - a^j_k\right)^2; \tag{8.48}$$

$$\mathcal{B}^{(1)}_k(\tau) = \frac{\tau^2}{2}\left(a^i_k - a^j_k\right)^3 \rho_k \xi_k - \frac{2}{3}\tau^3 \kappa_k \left(a^i_k - a^j_k\right)^3 \rho^k \xi^k; \tag{8.49}$$

$$\mathcal{B}^{(2)}_k(\tau) = \frac{\tau^3}{6}\xi^2_k \left(a^i_k - a^j_k\right)^4; \tag{8.50}$$

$$\mathcal{B}^{(3)}_k(\tau) = \frac{\tau^3}{6}\left(a^i_k - a^j_k\right)^4 \rho^2_k \xi^2_k. \tag{8.51}$$

Noting from (5.5)-(5.8) that $\mathcal{A}^{(i)}_k$ are one order in $\tau$ higher than the corresponding $\mathcal{B}^{(i)}_k$, the following approximations hold:

$$\sum_{k=1}^{d}\left(\mathcal{A}^{(0)}_k(\tau) + \mathcal{B}^{(0)}_k(\tau)V_k\right) = \sum_{k=1}^{d}\left(a^i_k - a^j_k\right)V_k \tau + o(\tau); \tag{8.52}$$

$$\sum_{k=1}^{d}\left(\mathcal{A}^{(1)}_k(\tau) + \mathcal{B}^{(1)}_k(\tau)V_k\right) = \sum_{k=1}^{d}\rho_k \xi_k \left(a^i_k - a^j_k\right)^3 V_k \frac{\tau^2}{2} + o(\tau^2); \tag{8.53}$$

$$\sum_{k=1}^{d}\left(\mathcal{A}^{(2)}_k(\tau) + \mathcal{B}^{(2)}_k(\tau)V_k\right) = \sum_{k=1}^{d}\xi^2_k \left(a^i_k - a^j_k\right)^4 V_k \frac{\tau^3}{6} + o(\tau^3); \tag{8.54}$$

$$\sum_{k=1}^{d}\left(\mathcal{A}^{(3)}_k(\tau) + \mathcal{B}^{(3)}_k(\tau)V_k\right) = \sum_{k=1}^{d}\rho^2_k \xi^2_k \left(a^i_k - a^j_k\right)^4 V_k \frac{\tau^3}{6} + o(\tau^3). \tag{8.55}$$

We introduce two new variables: the log- forward moneyness $m_f := \log\left(\frac{Se^{(r^i - r^j)\tau}}{K}\right)$ and also $V := \left(\mathbf{a}^i - \mathbf{a}^j\right)^\top \text{Diag}(\mathbf{V})\left(\mathbf{a}^i - \mathbf{a}^j\right)\tau$. Then, from Lewis (2000), we consider the following ratios among the derivatives of the Black-Scholes formula:

$$\frac{\partial^2_{xv} C_{\text{BS}}(x, K, V, \tau)}{\partial_v C_{\text{BS}}(x, K, V, \tau)} = \frac{1}{2} + \frac{m_f}{V}; \tag{8.56}$$

$$\frac{\partial^2_{vv} C_{\text{BS}}(x, K, V, \tau)}{\partial_v C_{\text{BS}}(x, K, V, \tau)} = \frac{m^2_f}{2V^2} - \frac{1}{2V} - \frac{1}{8}; \tag{8.57}$$

$$\frac{\partial^3_{xxv} C_{\text{BS}}(x, K, V, \tau)}{\partial_v C_{\text{BS}}(x, K, V, \tau)} = \frac{1}{4} + \frac{m_f - 1}{V} + \frac{m^2_f}{V^2}; \tag{8.58}$$

$$\frac{\partial^4_{xxvv} C_{\text{BS}}(x, K, V, \tau)}{\partial_v C_{\text{BS}}(x, K, V, \tau)} = \frac{m^4_f}{2V^4} + \frac{m^2_f(m_f - 1)}{2V^3}. \tag{8.59}$$

Upon substitution of (8.52)-(8.59) into (8.45)-(8.47), we obtain the values for $\zeta_i, i = 0, 1, 2$ allowing us to express the expansion of the implied volatility.

$$\zeta_0 = \left(\mathbf{a}^i - \mathbf{a}^j\right)^\top \mathrm{Diag}(\mathbf{V}) \left(\mathbf{a}^i - \mathbf{a}^j\right) \tau; \tag{8.60}$$

$$\zeta_1 = \left(\sum_{k=1}^{d} \frac{\rho_k \xi_k}{2} \left(a_k^i - a_k^j\right)^4 V_k\right) \frac{m_f}{\left(\mathbf{a}^i - \mathbf{a}^j\right)^\top \mathrm{Diag}(\mathbf{V}) \left(\mathbf{a}^i - \mathbf{a}^j\right)} \tau; \tag{8.61}$$

$$\zeta_2 = \frac{m_f^2}{\left(\left(\mathbf{a}^i - \mathbf{a}^j\right)^\top \mathrm{Diag}(\mathbf{V}) \left(\mathbf{a}^i - \mathbf{a}^j\right)\right)^2} \tau \left[\frac{1}{12} \left(\sum_{k=1}^{d} \xi_k^2 \left(a_k^i - a_k^j\right)^4 V_k\right)\right.$$

$$\left. + \frac{1}{6} \left(\sum_{k=1}^{d} \rho_k^2 \xi_k^2 \left(a_k^i - a_k^j\right)^4 V_k\right) - \frac{5}{16} \frac{\left(\sum_{k=1}^{d} \rho_k \xi_k \left(a_k^i - a_k^j\right)^3 V_k\right)^2}{\left(\mathbf{a}^i - \mathbf{a}^j\right)^\top \mathrm{Diag}(\mathbf{V}) \left(\mathbf{a}^i - \mathbf{a}^j\right)}\right]. \tag{8.62}$$

This completes the proof.

## 9. Images and Tables

| | 6 | 5 | 4 | 3 | 2 |
|---|---|---|---|---|---|
| $V_1$ | 0.0137 | 0.0137 | 0.0136 | 0.0137 | 0.0135 |
| $V_2$ | 0.0391 | 0.0365 | 0.0278 | 0.0293 | 0.0273 |
| $a_1^{\mathrm{USD}}$ | 0.6650 | 0.6713 | 0.6165 | 0.6371 | 0.6518 |
| $a_2^{\mathrm{USD}}$ | 1.0985 | 1.0531 | 0.9700 | 0.9795 | 0.9514 |
| $a_1^{\mathrm{EUR}}$ | 1.6177 | 1.6222 | 1.5648 | 1.5804 | 1.6061 |
| $a_2^{\mathrm{EUR}}$ | 1.3588 | 1.3208 | 1.2746 | 1.2797 | 1.2737 |
| $a_1^{\mathrm{JPY}}$ | 0.2995 | 0.3151 | 0.2732 | 0.3035 | 0.3116 |
| $a_2^{\mathrm{JPY}}$ | 1.6214 | 1.5922 | 1.5882 | 1.5858 | 1.5816 |
| $\kappa_1$ | 0.9418 | 1.1432 | 1.5138 | 1.7349 | 1.8685 |
| $\kappa_2$ | 1.7909 | 1.9998 | 1.9014 | 0.7142 | 0.7210 |
| $\theta_1$ | 0.0370 | 0.0349 | 0.0329 | 0.0329 | 0.0297 |
| $\theta_2$ | 0.0909 | 0.0839 | 0.0670 | 0.1236 | 0.1091 |
| $\xi_1$ | 0.4912 | 0.5138 | 0.5542 | 0.5847 | 0.5962 |
| $\xi_2$ | 1.0000 | 0.9997 | 0.8736 | 0.8318 | 0.8568 |
| $\rho_1$ | 0.5231 | 0.5118 | 0.4916 | 0.4727 | 0.4567 |
| $\rho_2$ | -0.3980 | -0.3956 | -0.3943 | -0.3902 | -0.3728 |
| Res. norm. | 4.6996e-004 | 3.4244e-004 | 1.8618e-004 | 1.1145e-004 | 5.2514e-005 |

TABLE 1. This table reports the results of the calibration of the model. We concentrate on the two factor case. For each column, a different number of expiries, ranging from 6 to 2, is chosen. More specifically, 6 means that the following expiries are considered: 1, 2, 3, 6, 9 months and 1 year, whereas 5 means that the longest maturity, i.e. 1 year is excluded from the sample. We proceed analogously in the subsequent columns by excluding the longest expiry date up to the point where we perform the calibration on the 2-sample, where we fit the smile at 1 and 2 months. We consider market data as of July 23rd 2010. The reference exchange rates are EURJPY=112.29, EURUSD=1.2921 and USDJPY=86.90. Res. norm. is the residual of the objective function for the given set of parameters.

| | 5 | 4 | 3 | 2 |
|---|---|---|---|---|
| $V_1$ | 0.1244% | -0.2866% | 0.0960% | -1.1068% |
| $V_2$ | -6.5645% | -28.9269% | -25.0960% | -30.0900% |
| $a_1^{\mathrm{USD}}$ | 0.9368% | -7.3035% | -4.1928% | -1.9883% |
| $a_2^{\mathrm{USD}}$ | -4.1309% | -11.6957% | -10.8309% | -13.3918% |
| $a_1^{\mathrm{EUR}}$ | 0.2745% | -3.2714% | -2.3082% | -0.7190% |
| $a_2^{\mathrm{EUR}}$ | -2.7989% | -6.1962% | -5.8255% | -6.2652% |
| $a_1^{\mathrm{JPY}}$ | 5.1809% | -8.8010% | 1.3206% | 4.0245% |
| $a_2^{\mathrm{JPY}}$ | -1.7985% | -2.0460% | -2.1910% | -2.4522% |
| $\kappa_1$ | 21.3845% | 60.7349% | 84.2055% | 98.3943% |
| $\kappa_2$ | 11.6649% | 6.1715% | -60.1213% | -59.7402% |
| $\theta_1$ | -5.6226% | -11.1784% | -11.0318% | -19.7810% |
| $\theta_2$ | -7.7145% | -26.3082% | 36.0430% | 20.0453% |
| $\xi_1$ | 4.6020% | 12.8359% | 19.0424% | 21.3784% |
| $\xi_2$ | -0.0244% | -12.6344% | -16.8201% | -14.3193% |
| $\rho_1$ | -2.1702% | -6.0305% | -9.6495% | -12.7003% |
| $\rho_2$ | -0.6031% | -0.9375% | -1.9522% | -6.3351% |

TABLE 2. In this table we consider the calibration on the largest sample as a basic case. We report the percentage difference between the model parameters resulting from the subsamples

| | 6 | 5 | 4 | 3 | 2 |
|---|---|---|---|---|---|
| $V_1$ | 0.0438 | 0.0430 | 0.0405 | 0.0421 | 0.0412 |
| $V_2$ | 0.0465 | 0.0450 | 0.0408 | 0.0370 | 0.0335 |
| $a_1^{\mathrm{USD}}$ | 0.7201 | 0.7165 | 0.7086 | 0.7099 | 0.7082 |
| $a_2^{\mathrm{USD}}$ | 1.0211 | 1.0182 | 1.0095 | 0.9915 | 0.9685 |
| $a_1^{\mathrm{EUR}}$ | 1.2517 | 1.2534 | 1.2603 | 1.2477 | 1.2538 |
| $a_2^{\mathrm{EUR}}$ | 1.2624 | 1.2616 | 1.2619 | 1.2575 | 1.2589 |
| $a_1^{\mathrm{JPY}}$ | 0.5159 | 0.5155 | 0.5093 | 0.5206 | 0.5142 |
| $a_2^{\mathrm{JPY}}$ | 1.5053 | 1.5083 | 1.5223 | 1.5307 | 1.5372 |
| $\theta_1$ | 0.1154 | 0.1169 | 0.1203 | 0.1391 | 0.1300 |
| $\theta_2$ | 0.1344 | 0.1377 | 0.1350 | 0.1253 | 0.1081 |
| $\xi_1$ | 0.8892 | 0.8898 | 0.8992 | 0.9700 | 0.9925 |
| $\xi_2$ | 0.9338 | 0.9450 | 0.9458 | 0.9616 | 0.9659 |
| $\rho_1$ | 0.5226 | 0.5132 | 0.4950 | 0.4756 | 0.4591 |
| $\rho_2$ | -0.4042 | -0.4030 | -0.4004 | -0.3887 | -0.3721 |

TABLE 3. This table reports the results of the calibration of the model. In this case we are assuming $\kappa_k = 1, k = 1, 2$. For each column, a different number of expiries, ranging from 6 to 2, is chosen. Res. norm. is the residual of the objective function for the given set of parameters.

| | 5 | 4 | 3 | 2 |
|---|---|---|---|---|
| $V_1$ | -1.8915% | -7.6930% | -3.8971% | -5.9284% |
| $V_2$ | -3.1003% | -12.1687% | -20.3324% | -28.0033% |
| $a_1^{\mathrm{USD}}$ | -0.5056% | -1.6003% | -1.4171% | -1.6497% |
| $a_2^{\mathrm{USD}}$ | -0.2832% | -1.1348% | -2.9033% | -5.1535% |
| $a_1^{\mathrm{EUR}}$ | 0.1322% | 0.6825% | -0.3164% | 0.1703% |
| $a_2^{\mathrm{EUR}}$ | -0.0691% | -0.0438% | -0.3903% | -0.2826% |
| $a_1^{\mathrm{JPY}}$ | -0.0857% | -1.2718% | 0.9087% | -0.3396% |
| $a_2^{\mathrm{JPY}}$ | 0.1994% | 1.1235% | 1.6872% | 2.1131% |
| $\theta_1$ | 1.3740% | 4.2649% | 20.5911% | 12.6966% |
| $\theta_2$ | 2.4412% | 0.4181% | -6.8073% | -19.5908% |
| $\xi_1$ | 0.0622% | 1.1246% | 9.0864% | 11.6096% |
| $\xi_2$ | 1.1985% | 1.2881% | 2.9747% | 3.4434% |
| $\rho_1$ | -1.7994% | -5.2886% | -8.9844% | -12.1501% |
| $\rho_2$ | -0.2922% | -0.9392% | -3.8297% | -7.9304% |

TABLE 4. In this table we consider the calibration on the largest sample as a basic case, when $\kappa_k = 1, k = 1, 2$. We report the percentage difference between the model parameters resulting from the subsamples

| | 6 | 5 | 4 | 3 | 2 |
|---|---|---|---|---|---|
| $k = 1$ | -0.1715 | -0.1841 | -0.2076 | -0.2276 | -0.2445 |
| $k = 2$ | -0.6745 | -0.6640 | -0.5086 | -0.5153 | -0.5768 |

TABLE 5. For all $k = 1, 2$ and for each sample we report the quantity $2\kappa_k\theta_k - \xi_k^2$. In all cases the quantity is negative and its absolute value is a measure of the violation of the Feller condition.

| | USD/EUR | USD/JPY | EUR/JPY |
|---|---|---|---|
| 10DC | -0.0006 | -0.0002 | -0.0027 |
| 15DC | | 0.0003 | -0.0017 |
| 25DC | -0.0012 | 0.0005 | -0.0005 |
| 0 | -0.0022 | 0.0009 | 0.0021 |
| 25DP | -0.0008 | 0.0012 | 0.0042 |
| 15DP | | 0.0004 | 0.0031 |
| 10DP | 0.0009 | -0.0001 | 0.0011 |

TABLE 6. This table reports the raw difference between the market implied volatility and the implied volatility for 1 year, when we calibrate the model to the previous 5 expiries. Blanks on the first column reflect missing market data for 15DC and 15DP.

| | USD/EUR 9m | USD/EUR 1y | USD/JPY 9m | USD/JPY 1y | EUR/JPY 9m | EUR/JPY 1y |
|---|---|---|---|---|---|---|
| 10DC | -0.0031 | 0.0003 | -0.0013 | -0.0001 | -0.0006 | -0.0019 |
| 15DC | | | -0.0007 | 0.0002 | 0.0008 | -0.0012 |
| 25DC | -0.0023 | -0.0007 | 0.0004 | 0.0004 | 0.0021 | -0.0006 |
| 0 | -0.0021 | -0.0021 | 0.0020 | 0.0006 | 0.0036 | 0.0012 |
| 25DP | -0.0010 | -0.0007 | 0.0011 | 0.0010 | 0.0028 | 0.0033 |
| 15DP | | | -0.0008 | 0.0003 | 0.0004 | 0.0025 |
| 10DP | -0.0006 | 0.0015 | -0.0014 | -0.0000 | -0.0026 | 0.0007 |

TABLE 7. This table reports the raw difference between the market implied volatility and the implied volatility for 1 year and 9 months, when we calibrate the model to the previous 4 expiries. Blanks on the first two columns reflect missing market data for 15DC and 15DP.

| USD/EUR | 6m | 9m | 1y |
|---|---|---|---|
| 10DC | -0.0060 | -0.0023 | 0.0015 |
| 25DC | -0.0034 | -0.0019 | 0.0002 |
| 0 | -0.0018 | -0.0020 | -0.0017 |
| 25DP | -0.0006 | -0.0010 | -0.0005 |
| 10DP | 0.0002 | -0.0002 | 0.0019 |

| USD/JPY | 6m | 9m | 1y |
|---|---|---|---|
| 10DC | -0.0058 | -0.0011 | 0.0000 |
| 15DC | -0.0042 | -0.0005 | 0.0002 |
| 25DC | -0.0009 | 0.0005 | -0.0000 |
| 0 | 0.0016 | 0.0020 | 0.0000 |
| 25DP | 0.0004 | 0.0009 | 0.0007 |
| 15DP | -0.0019 | -0.0012 | 0.0002 |
| 10DP | -0.0037 | -0.0020 | -0.0002 |

| EUR/JPY | 6m | 9m | 1y |
|---|---|---|---|
| 10DC | -0.0008 | -0.0003 | -0.0009 |
| 15DC | 0.0011 | 0.0007 | -0.0006 |
| 25DC | 0.0031 | 0.0015 | -0.0005 |
| 0 | 0.0041 | 0.0025 | 0.0006 |
| 25DP | 0.0014 | 0.0018 | 0.0026 |
| 15DP | -0.0022 | -0.0004 | 0.0019 |
| 10DP | -0.0053 | -0.0032 | 0.0003 |

TABLE 8. This table reports the raw difference between the market implied volatility and the implied volatility for 1 year, 9 and 6 months, when we calibrate the model to the previous 3 expiries.

USD/EUR

| | 3m | 6m | 9m | 1y |
|---|---|---|---|---|
| 10DC | -0.0062 | -0.0043 | -0.0006 | 0.0031 |
| 25DC | -0.0027 | -0.0022 | -0.0007 | 0.0010 |
| 0 | -0.0013 | -0.0009 | -0.0014 | -0.0015 |
| 25DP | -0.0006 | 0.0001 | -0.0005 | -0.0005 |
| 10DP | 0.0008 | 0.0009 | 0.0003 | 0.0020 |

USD/JPY

| | 3m | 6m | 9m | 1y |
|---|---|---|---|---|
| 10DC | -0.0090 | -0.0052 | -0.0004 | 0.0007 |
| 15DC | -0.0062 | -0.0038 | -0.0002 | 0.0003 |
| 25DC | -0.0037 | -0.0009 | 0.0003 | -0.0006 |
| 0 | 0.0010 | 0.0014 | 0.0015 | -0.0011 |
| 25DP | -0.0004 | 0.0006 | 0.0009 | 0.0002 |
| 15DP | -0.0022 | -0.0014 | -0.0008 | 0.0002 |
| 10DP | -0.0039 | -0.0029 | -0.0013 | 0.0003 |

EUR/JPY

| | 3m | 6m | 9m | 1y |
|---|---|---|---|---|
| 10DC | -0.0045 | -0.0003 | 0.0002 | -0.0006 |
| 15DC | -0.0021 | 0.0014 | 0.0009 | -0.0007 |
| 25DC | 0.0010 | 0.0030 | 0.0012 | -0.0013 |
| 0 | 0.0030 | 0.0038 | 0.0019 | -0.0008 |
| 25DP | 0.0009 | 0.0014 | 0.0014 | 0.0016 |
| 15DP | -0.0026 | -0.0020 | -0.0005 | 0.0013 |
| 10DP | -0.0057 | -0.0050 | -0.0030 | -0.0000 |

TABLE 9. This table reports the raw difference between the market implied volatility and the implied volatility for 1 year, 9, 6 and 3 months, when we calibrate the model to the previous 2 expiries.

| | USD/EUR | USD/JPY | EUR/JPY |
|---|---|---|---|
| 10DC | 0.0058 | 0.0030 | 0.0046 |
| 15DC | | 0.0022 | 0.0035 |
| 25DC | 0.0058 | 0.0009 | 0.0041 |
| 0 | 0.0035 | 0.0007 | 0.0019 |
| 25DP | 0.0029 | 0.0024 | 0.0016 |
| 15DP | | 0.0027 | -0.0012 |
| 10DP | 0.0042 | 0.0029 | -0.0044 |

TABLE 10. This table reports the raw difference between the market implied volatility and the implied volatility for 1 year, when we calibrate the model to the previous 5 expiries and $\kappa_k = 1, k = 1, 2$. Blanks on the first column reflect missing market data for 15DC and 15DP.

| | USD/EUR 9m | USD/EUR 1y | USD/JPY 9m | USD/JPY 1y | EUR/JPY 9m | EUR/JPY 1y |
|---|---|---|---|---|---|---|
| 10DC | 0.0083 | 0.0092 | 0.0034 | 0.0052 | 0.0073 | 0.0090 |
| 15DC | | | 0.0021 | 0.0042 | 0.0061 | 0.0078 |
| 25DC | 0.0060 | 0.0092 | 0.0005 | 0.0029 | 0.0053 | 0.0079 |
| 0 | 0.0029 | 0.0066 | -0.0000 | 0.0025 | 0.0024 | 0.0050 |
| 25DP | 0.0033 | 0.0056 | 0.0026 | 0.0044 | 0.0025 | 0.0041 |
| 15DP | | | 0.0034 | 0.0049 | 0.0006 | 0.0010 |
| 10DP | 0.0062 | 0.0067 | 0.0041 | 0.0054 | -0.0027 | -0.0024 |

TABLE 11. This table reports the raw difference between the market implied volatility and the implied volatility for 1 year and 9 months, when we calibrate the model to the previous 4 expiries and $\kappa_k = 1, k = 1, 2$. Blanks on the first two columns reflect missing market data for 15DC and 15DP.

| USD/EUR | 6m | 9m | 1y |
|---|---|---|---|
| 10DC | 0.0073 | 0.0116 | 0.0131 |
| 25DC | 0.0042 | 0.0091 | 0.0128 |
| 0 | 0.0010 | 0.0056 | 0.0100 |
| 25DP | 0.0020 | 0.0058 | 0.0087 |
| 10DP | 0.0056 | 0.0088 | 0.0099 |

| USD/JPY | 6m | 9m | 1y |
|---|---|---|---|
| 10DC | 0.0025 | 0.0048 | 0.0067 |
| 15DC | 0.0009 | 0.0030 | 0.0053 |
| 25DC | -0.0013 | 0.0008 | 0.0034 |
| 0 | -0.0028 | -0.0004 | 0.0023 |
| 25DP | 0.0002 | 0.0027 | 0.0048 |
| 15DP | 0.0013 | 0.0041 | 0.0057 |
| 10DP | 0.0022 | 0.0053 | 0.0067 |

| EUR/JPY | 6m | 9m | 1y |
|---|---|---|---|
| 10DC | 0.0051 | 0.0114 | 0.0138 |
| 15DC | 0.0036 | 0.0099 | 0.0123 |
| 25DC | 0.0022 | 0.0085 | 0.0120 |
| 0 | -0.0001 | 0.0048 | 0.0083 |
| 25DP | 0.0017 | 0.0047 | 0.0071 |
| 15DP | 0.0010 | 0.0028 | 0.0040 |
| 10DP | -0.0016 | -0.0003 | 0.0006 |

TABLE 12. This table reports the raw difference between the market implied volatility and the implied volatility for 1 year, 9 and 6 months, when we calibrate the model to the previous 3 expiries and $\kappa_k = 1, k = 1, 2$.

| USD/EUR | | | | |
|---|---|---|---|---|
| | 3m | 6m | 9m | 1y |
| 10DC | 0.0057 | 0.0085 | 0.0125 | 0.0136 |
| 25DC | 0.0023 | 0.0042 | 0.0086 | 0.0121 |
| 0 | -0.0018 | -0.0000 | 0.0041 | 0.0082 |
| 25DP | -0.0008 | 0.0010 | 0.0043 | 0.0068 |
| 10DP | 0.0025 | 0.0050 | 0.0077 | 0.0086 |
| USD/JPY | | | | |
| | 3m | 6m | 9m | 1y |
| 10DC | 0.0013 | 0.0026 | 0.0046 | 0.0063 |
| 15DC | 0.0004 | 0.0003 | 0.0020 | 0.0040 |
| 25DC | -0.0015 | -0.0028 | -0.0012 | 0.0010 |
| 0 | -0.0033 | -0.0051 | -0.0033 | -0.0010 |
| 25DP | -0.0010 | -0.0015 | 0.0005 | 0.0022 |
| 15DP | -0.0004 | 0.0002 | 0.0025 | 0.0038 |
| 10DP | -0.0002 | 0.0017 | 0.0043 | 0.0053 |
| EUR/JPY | | | | |
| | 3m | 6m | 9m | 1y |
| 10DC | 0.0002 | 0.0048 | 0.0106 | 0.0127 |
| 15DC | -0.0009 | 0.0025 | 0.0082 | 0.0101 |
| 25DC | -0.0026 | -0.0001 | 0.0056 | 0.0086 |
| 0 | -0.0038 | -0.0035 | 0.0008 | 0.0037 |
| 25DP | -0.0008 | -0.0007 | 0.0016 | 0.0034 |
| 15DP | -0.0010 | -0.0007 | 0.0005 | 0.0012 |
| 10DP | -0.0021 | -0.0026 | -0.0020 | -0.0015 |

TABLE 13. This table reports the raw difference between the market implied volatility and the implied volatility for 1 year, 9, 6 and 3 months, when we calibrate the model to the previous 3 expiries and $\kappa_k = 1, k = 1, 2$.

| Order | $S^{\mathrm{USD,EUR}}$ | $S^{\mathrm{JPY,USD}}$ | $S^{\mathrm{JPY,EUR}}$ |
|---|---|---|---|
| 1 | $+\infty$ | $+\infty$ | $+\infty$ |
| 2 | $+\infty$ | 12.1962 | 5.1612 |
| 3 | $+\infty$ | 2.9537 | 2.0580 |
| 4 | 3.3968 | 1.7990 | 1.3763 |
| 5 | 2.0070 | 2.0819 | 1.0614 |
| 6 | 1.5057 | 1.9550 | 0.8736 |
| 7 | 1.2265 | 1.7030 | 0.7463 |
| 8 | 1.0427 | 1.5143 | 0.6534 |
| 9 | 0.9104 | 1.3715 | 0.5822 |
| 10 | 0.8099 | 1.2591 | 0.5256 |
| 11 | 0.7303 | 1.1676 | 0.4794 |
| 12 | 0.6656 | 1.0912 | 0.4409 |
| 13 | 0.6119 | 1.0260 | 0.4083 |
| 14 | 0.5664 | 0.9695 | 0.3803 |
| 15 | 0.5274 | 0.9199 | 0.3559 |
| 16 | 0.4935 | 0.8758 | 0.3346 |
| 17 | 0.4638 | 0.8363 | 0.3157 |
| 18 | 0.4376 | 0.8006 | 0.2988 |
| 19 | 0.4142 | 0.7681 | 0.2837 |
| 20 | 0.3932 | 0.7385 | 0.2701 |

TABLE 14. Times of moment explosions for moments up to order 5 for the three currency pairs.

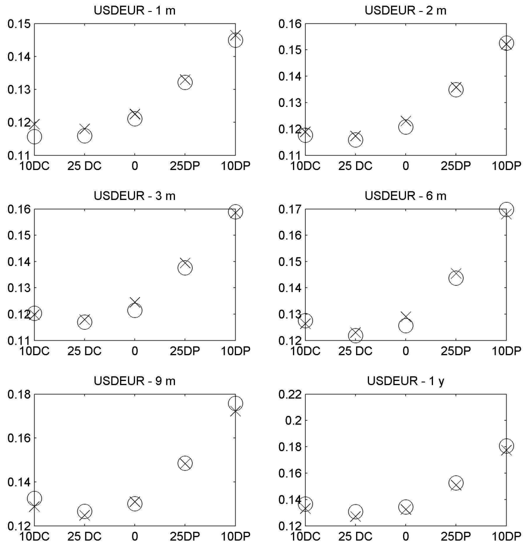



FIGURE 1. Calibration of USD/EUR implied volatility surface.

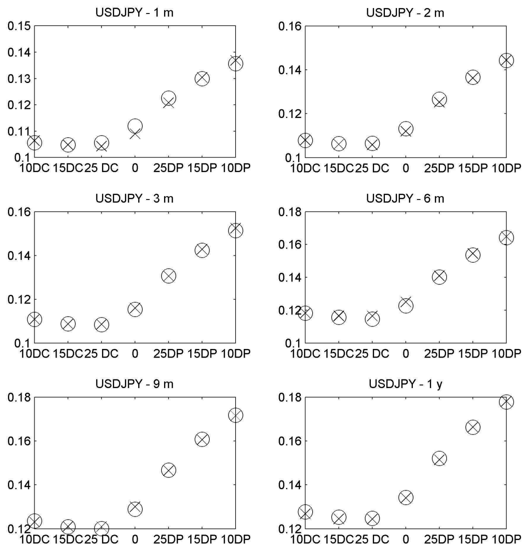



FIGURE 2. Calibration of USD/JPY implied volatility surface.

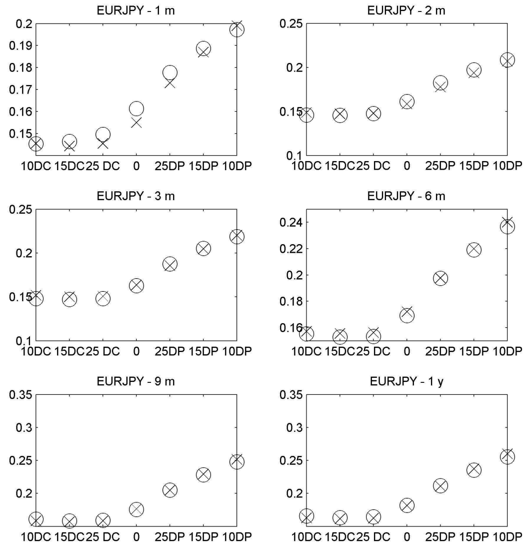



FIGURE 3. Calibration of EUR/JPY implied volatility surface.

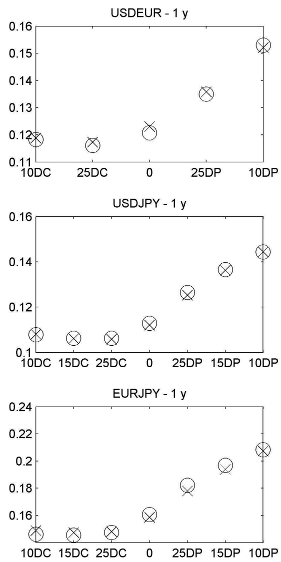



FIGURE 4. 1-year out-of-sample performance, using parameters calibrated from the 5 expiries sub-sample

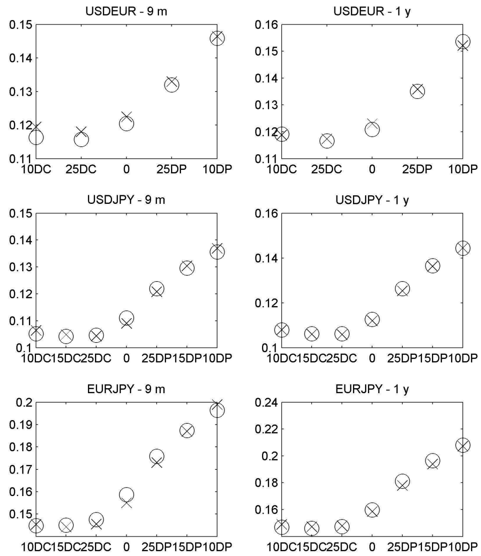



FIGURE 5. 9-month and 1-year out-of-sample performance, using parameters calibrated from the 4 expiries sub-sample

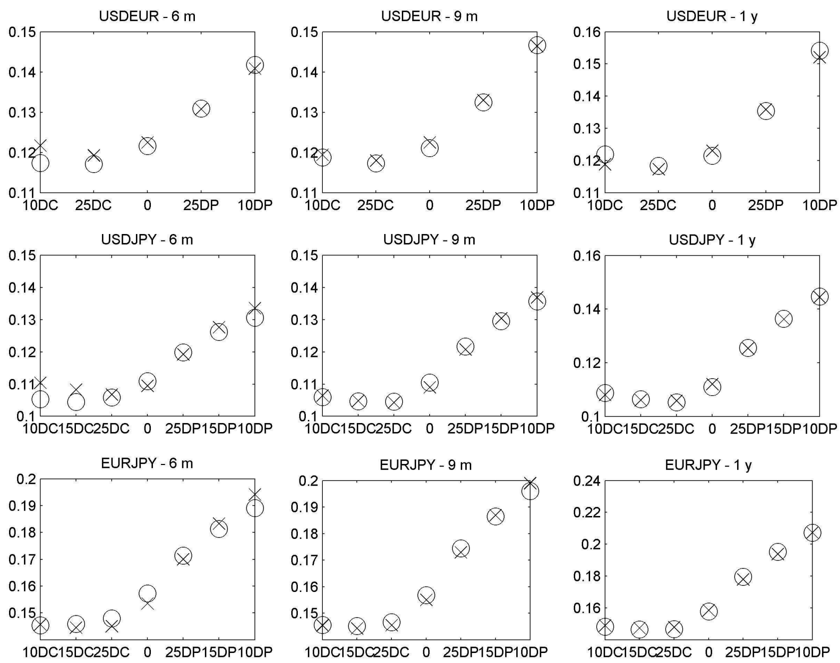



FIGURE 6. 6-month, 9-month and 1-year out-of-sample performance, using parameters calibrated from the 4 expiries sub-sample

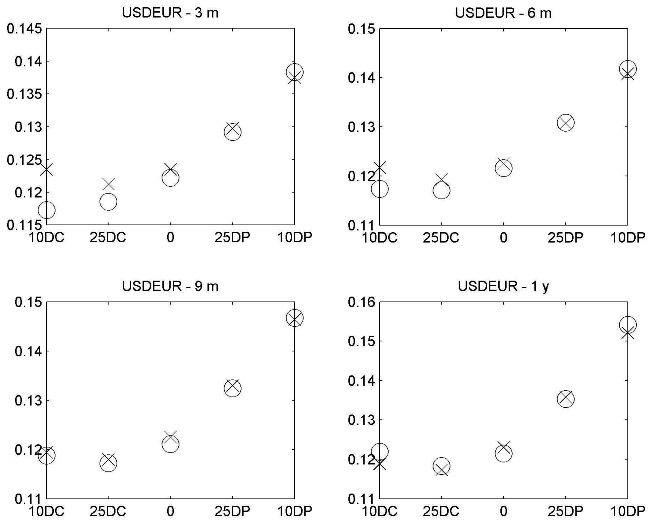



FIGURE 7. 6-month, 9-month and 1-year out-of-sample performance, using parameters calibrated from the 4 expiries sub-sample

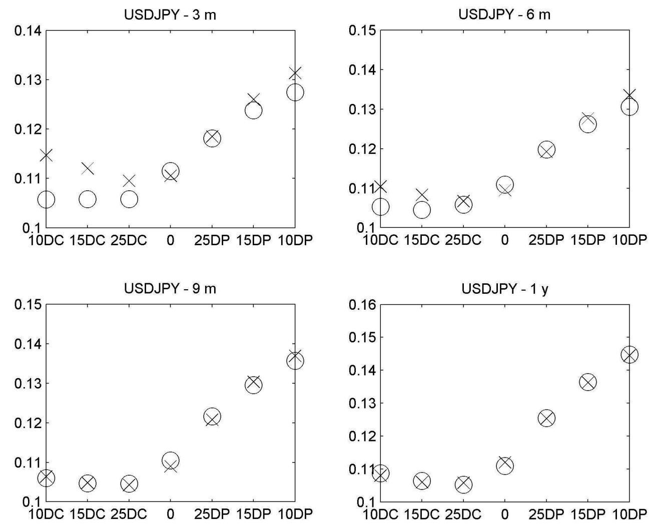



FIGURE 8. 6-month, 9-month and 1-year out-of-sample performance, using parameters calibrated from the 4 expiries sub-sample

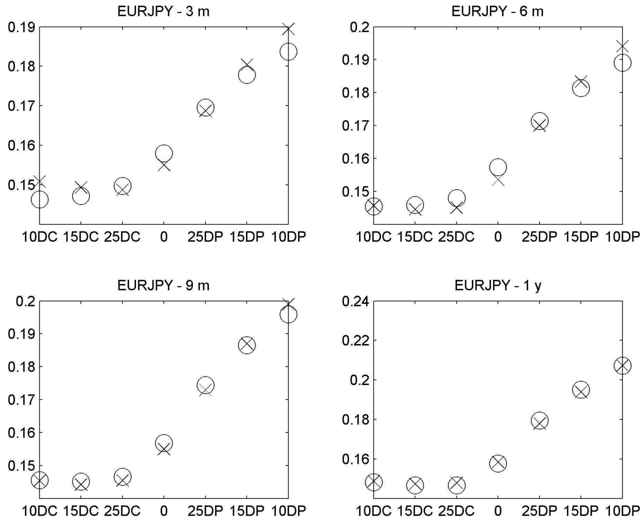



FIGURE 9. 6-month, 9-month and 1-year out-of-sample performance, using parameters calibrated from the 4 expiries sub-sample

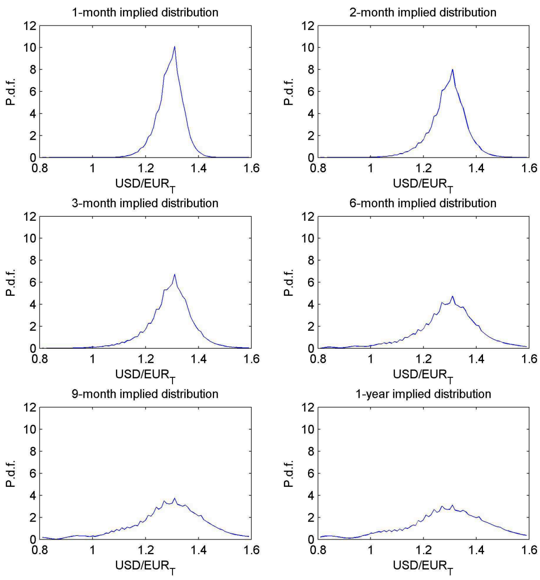



FIGURE 10. Implied risk-neutral conditional distributions for the USD/EUR exchange rate inferred from option prices.

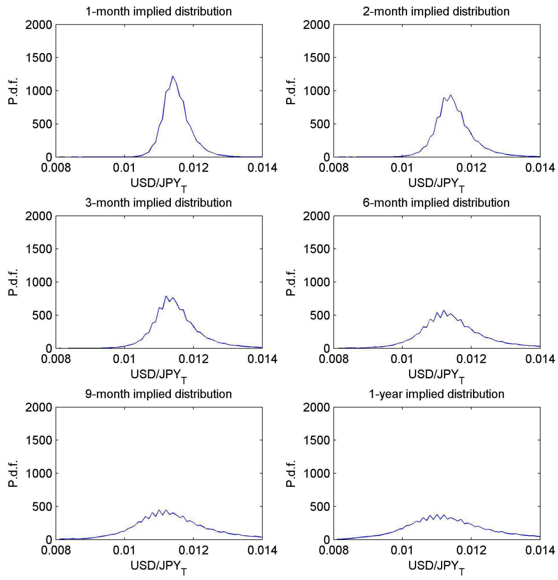



FIGURE 11. Implied risk-neutral conditional distributions for the USD/JPY exchange rate inferred from option prices.

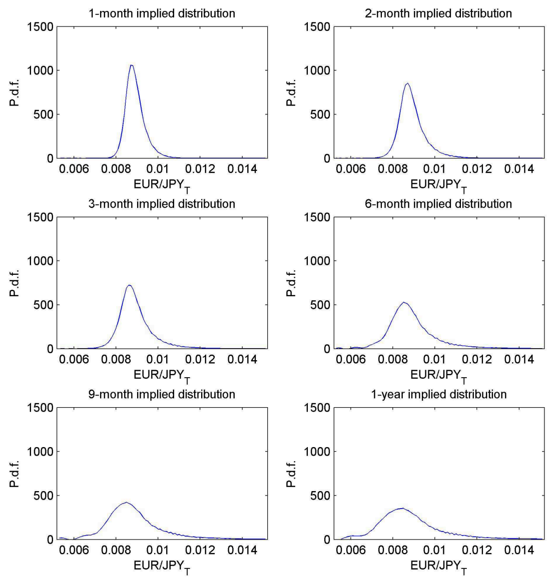



FIGURE 12. Implied risk-neutral conditional distributions for the EUR/JPY exchange rate inferred from option prices.

CHAPTER 5

# The Wishart case

### 1. Basic definitions and assumptions

In this chapter we propose an extension of the previous approach to the setting of matrix stochastic processes on the state space $S_d^+$. We will see that all the good analytical properties of the previous approach are preserved, while we allow for a more general dynamics. This increased freedom may be of interest in the context of the evaluation of complex multi-currency derivatives.

### 2. Wishart-based exchange model

We present another model in which again $N$ currencies are traded. As before, we start by considering the value of each of these currencies in units of a universal numéraire. For applications and calibration purposes we will focus on the case $N = 3$. We work under the risk neutral measure defined by the universal numéraire and call $S^{0,i}(t)$ the value at time $t$ of one unit of the currency $i$ in terms of our universal numéraire (note that $S^{0,i}(t)$ can itself be thought as an exchange rate, between the universal numéraire and the currency $i$). We model each of the $S^{0,i}(t)$ via a multifactor Wishart stochastic volatility model of dimension $d$ and a matrix Brownian motion, $Z_t \in M_d$. The dimension $d$ can be chosen according to the specific problem and may reflect a PCA-type analysis. We further assume that these stochastic volatility components are *common* between the different $S^{0,i}(t)$. Formally, we write

$$\frac{dS^{0,i}(t)}{S^{0,i}(t)} = (r^0 - r^i)dt - Tr\left[A_i\sqrt{\Sigma(t)}dZ(t)\right], \qquad i = 1, \dots, N; \tag{2.1}$$

$$d\Sigma(t) = (\Omega\Omega^\top + M\Sigma(t) + \Sigma(t)M^\top)dt + \sqrt{\Sigma(t)}dW(t)Q + Q^\top dW(t)^\top. \tag{2.2}$$

We assume $\Omega, M, Q \in M_d$, $W = (W_t)_{t\geq 0} \in M_d$ is a matrix Brownian motion. In order to ensure the typical mean reverting behavior of the process we assume that $M$ is negative semi-definite, moreover the matrix $\Omega$ satisfies the condition $\Omega\Omega^\top = \beta Q^\top Q$, for $\beta \geq d - 1$. This last condition ensures that the boundary of the cone $S_d^+$ is not reached by the process. The diffusion term exhibits a structure which is completely analogous to the one introduced in the previous chapter: in the present case we have that the dynamics of the exchange rate is driven by a linear projection of the variance factor $\sqrt{\Sigma(t)}$ along a direction parametrized by the symmetric matrix $A_i$. In consequence the total instantaneous variance is $Tr\left[A_i\Sigma(t)A_i\right]dt$. We assume the existence of $N$ basic traded asset (one for each monetary area) which are called money-market accounts, the values of which are driven by deterministic ODE's of the type:

$$dB^j(t) = r^j B^j(t)dt. \tag{2.3}$$

Finally, we assume a correlation structure between the two matrix Brownian motions $Z(t)$ and $W(t)$, by means of a matrix $R$ according to the following relationship:

$$W(t) = Z(t)R^\top + B(t)\sqrt{I_d - RR^\top}, \tag{2.4}$$

where $B(t)$ is a matrix Brownian motion independent of $Z(t)$. This is the basic setup of the model. We denote by $S^{i,j}(t)$, $i,j = 1,..,n$ the exchange rate between currency $i$ and $j$. More precisely, we have that $S^{i,j}(t) = S^{0,j}(t)/S^{0,i}(t)$ has the following dynamics:

$$\frac{dS^{i,j}(t)}{S^{i,j}(t)} = (r^i - r^j)dt + Tr[(A_i - A_j)\Sigma(t)A_i]dt + Tr[(A_i - A_j)\sqrt{\Sigma(t)}dZ(t)], \tag{2.5}$$

Note that at this stage we are still working in the risk neutral measure defined by the universal numéraire. The additional drift term in (2.5) can be understood as a quanto adjustment between the currencies $i$ and $j$.

PROPOSITION 5.1. *The dynamics of the exchange rate (2.5) satisfies the triangular relation, namely:*

$$dS^{i,j}(t) = d\left(S^{i,l}(t)S^{l,j}(t)\right). \tag{2.6}$$

PROOF. See Section 9 □

### 3. Risk neutral probability measures

Up to now we have worked in the risk neutral measure defined by our (rather unspecified) universal numéraire. In practical pricing applications, it is more convenient to change the numéraire to any of the currencies included in our FX multi-dimensional system. Without loss of generality, let us consider the risk neutral measure defined by the $i$-th money market account $B^i(t)$ and derive the dynamical equations for the standard FX rate $S^{i,j}(t)$, its inverse $S^{j,i}(t)$, and a generic cross $S^{j,l}(t)$.
The Girsanov change of measure that transfers to the $\mathbb{Q}^i$ risk neutral measure (i.e. the risk neutral measure in the $i$-th country) is simply determined by assuming that under $\mathbb{Q}^i$ the drift of the exchange rate $S^{i,j}(t)$ is given by $r^i - r^j$ (or equivalently by the fact that the money market account $B^f(t)$ is a $\mathbb{Q}^i$-martingale once discounted by the interest rate $r^i$). The associated Radon-Nikodym derivative is:

$$\left.\frac{d\mathbb{Q}^i}{d\mathbb{Q}^0}\right|_{\mathcal{F}_t} = \exp\left(-\int_0^t Tr[A_i\sqrt{\Sigma(s)}dZ(s)] - \frac{1}{2}\int_0^t Tr[A_i\Sigma(s)A_i]ds\right). \tag{3.1}$$

In Mayerhofer (2012), conditions ensuring that the stochastic exponential above is a true martingale are provided. In the following, we proceed along the lines of the previous chapter. The possibility of buying the foreign currency and investing it at the foreign short rate of interest, is equivalent to the possibility of investing in a domestic asset with price process $\tilde{B}(t)$. Assume $i$ is the domestic economy and $j$ is the foreign economy then:

$$\begin{aligned}
d\tilde{B}^i_j(t) &= d\left(B^j(t)S^{i,j}(t)\right)\\
&= B^j(t)S^{i,j}(t)\left((r^i - r^j)dt + Tr[(A_i - A_j)\Sigma(t)A_i]dt + Tr[(A_i - A_j)\sqrt{\Sigma(t)}dZ(t)]\right)\\
&\quad + B^jS^{i,j}(t)r^jdt\\
&= \tilde{B}^i_j(t)\left(r^idt + Tr[(A_i - A_j)\Sigma(t)A_i]dt + Tr[(A_i - A_j)\sqrt{\Sigma(t)}dZ(t)]\right)\\
&= \tilde{B}^i_j(t)\left(r^idt + Tr[(A_i - A_j)\sqrt{\Sigma(t)}dZ^{\mathbb{Q}^i}(t)]\right),
\end{aligned} \tag{3.2}$$

where the matrix brownian motion under $\mathbb{Q}^i$ is given by.

$$dZ^{\mathbb{Q}^i} = dZ + \sqrt{\Sigma(t)}A_idt, \tag{3.3}$$

then the $\mathbb{Q}^i$-risk neutral dynamics of the exchange rate is of the form

$$dS^{i,j}(t) = d\left(\frac{\tilde{B}^i_j(t)}{B_j(t)}\right)$$
$$= S^{i,j}(t)\left((r^i - r^j)dt + Tr\left[(A_i - A_j)\sqrt{\Sigma(t)}dZ^{\mathbb{Q}^i}(t)\right]\right). \tag{3.4}$$

The measure change has however also an impact on the variance processes, via the correlation matrix $R$. Recall that we have, from (2.4),

$$dW(t) = dZ(t)R^\top + dB(t)\sqrt{I_d - RR^\top}. \tag{3.5}$$

The component of $dB(t)$ which is orthogonal to the spot driver $dZ(t)$ is not affected by the measure change; this is a natural choice that is consistent with the foreign-domestic symmetry (see the discussion in Chapter 4). We are now able to derive the risk neutral dynamics of the factor process $\Sigma(t)$ governing the volatility of the exchange rates under $\mathbb{Q}^i$, which is given as

$$dW^{\mathbb{Q}^i}(t) = \left(dZ(t) + \sqrt{\Sigma}A_i dt\right)R^\top + dB(t)\sqrt{I_d - RR^\top}. \tag{3.6}$$

From (2.2) and (3.5) we derive the $\mathbb{Q}^i$-risk neutral dynamics of $\Sigma$ as follows:

$$\begin{aligned}
&d\Sigma(t) = \\
&(\Omega\Omega^\top + M\Sigma(t) + \Sigma(t)M^\top)dt \\
&+ \sqrt{\Sigma(t)}\left(\left(dZ(t) + \sqrt{\Sigma(t)}A_i dt\right)R^\top + dB(t)\sqrt{I_d - RR^\top}\right)Q \\
&+ Q^\top\left(R\left(dZ_t^\top + A_i\sqrt{\Sigma(t)}dt\right) + \sqrt{I_d - RR^\top}^\top dB^\top\right)\sqrt{\Sigma(t)} \\
&- \Sigma(t)A_i R^\top Q dt \\
&- Q^\top R A_i \Sigma(t)dt.
\end{aligned}$$

Now define:

$$M^{\mathbb{Q}^i} := M - Q^\top R A_i, \tag{3.7}$$

so that using (3.6) we can finally write

$$\begin{aligned}
d\Sigma(t) &= (\Omega\Omega^\top + M^{\mathbb{Q}^i}\Sigma(t) + \Sigma(t)M^{\mathbb{Q}^i,\top})dt \\
&\sqrt{\Sigma(t)}dW^{\mathbb{Q}^i}(t)Q + Q^\top dW^{\mathbb{Q}^i,\top}(t)\sqrt{\Sigma(t)}
\end{aligned} \tag{3.8}$$

and so the relations between the parameters are:

$$R^{\mathbb{Q}^i} = R, \tag{3.9}$$

$$Q^{\mathbb{Q}^i} = Q, \tag{3.10}$$

$$M^{\mathbb{Q}^i} = M - Q^\top R A_i. \tag{3.11}$$

We observe that, like in the multi-Heston case, the functional form of the model is invariant under the measure change between the 0 and the $i$th-risk neutral measure. The inverse FX rate under the $\mathbb{Q}^i$-risk neutral measure follows from Ito calculus, recalling that $S^{j,i} = \left(S^{i,j}\right)^{-1}$:

$$\frac{dS^{j,i}(t)}{S^{j,i}(t)} = S^{i,j}(t)d\left(\frac{1}{S^{i,j}(t)}\right) \tag{3.12}$$
$$= \left(r^j - r^i + Tr\left[(A_j - A_i)\Sigma(t)(A_j - A_i)\right]\right)dt + Tr\left[(A_j - A_i)\sqrt{\Sigma(t)}dZ^{\mathbb{Q}^i}(t)\right],$$

which includes the self-quanto adjustment. Similarly, the SDE of a generic cross FX rate becomes

$$\begin{aligned}\frac{dS^{j,l}(t)}{S^{j,l}(t)} &= \frac{S^{i,j}(t)}{S^{i,l}(t)} d\left(\frac{S^{i,l}(t)}{S^{i,j}(t)}\right) \\ &= \left(r^j - r^l + Tr\left[(A_j - A_l)\Sigma(t)(A_j - A_i)\right]\right) dt + Tr\left[(A_j - A_l)\sqrt{\Sigma(t)}dZ^{\mathbb{Q}^i}(t)\right].\end{aligned} \tag{3.13}$$

The additional drift term is the quanto adjustment as described by the current model choice. By applying Girsanov's theorem again, this time switching to the $\mathbb{Q}^j$ risk neutral measure, the term is removed while the Wishart parameters become

$$R^{\mathbb{Q}^j} = R^{\mathbb{Q}^i}, \tag{3.14}$$

$$Q^{\mathbb{Q}^j} = Q^{\mathbb{Q}^i}, \tag{3.15}$$

$$M^{\mathbb{Q}^j} = M^{\mathbb{Q}^i} - Q^{\mathbb{Q}^i,\top} R^{\mathbb{Q}^i}(A_j - A_i). \tag{3.16}$$

These are the fundamental transformation rules for the model parameters.
Previously, we mentioned that we assume that the component of $dW(t)$, which is orthogonal to the spot driver $dZ(t)$, is not affected by the measure change. This assumption is in line with the procedure that has been introduced in Chapter 4 and implies that the model is consistent with the foreign-domestic parity as in Del Baño Rollin (2008).

## 4. Features of the model

**4.1. Stochastic Skew.** In the present framework we have a nice feature: the skew is stochastic. To be more precise we define the skew as:

$$Corr_t(Noise(d\log S^{i,j}(t)), Noise(Vol(d\log S^{i,j}(t)))). \tag{4.1}$$

By computing the volatility of $S^{i,j}$

$$d\left\langle S^{i,j}(t), S^{i,j}(t)\right\rangle_t = Tr[(A_i - A_j)\Sigma(t)(A_i - A_j)]dt, \tag{4.2}$$

we obtain that the skew is then proportional to the quantity

$$d\left\langle S^{i,j}(t), Vol(S^{i,j}(t))\right\rangle_t \propto Tr\left[(A_i - A_j)\Sigma(t)(A_i - A_j)^2 Q^\top R\right]. \tag{4.3}$$

This proportionality may be appreciated once we compute the skew, which can be derived by proceeding along the following steps. We assume that the matrices $A$ are all symmetric. In Section 3 we have seen that the $\mathbb{Q}^i$-risk neutral dynamics of $S^{i,j}$ is given as in the next relation. This allows us to express the dynamics by means of a scalar Brownian motion $B_1$.

$$\begin{aligned}\frac{dS^{i,j}(t)}{S^{i,j}(t)} &= (r^i - r^j)dt + Tr\left[(A_i - A_j)\sqrt{\Sigma(t)}dZ^{\mathbb{Q}^i}(t)\right] \\ &= (r^i - r^j)dt + \sqrt{Tr\left[(A_i - A_j)\Sigma(t)(A_i - A_j)\right]}\underbrace{\frac{Tr\left[(A_i - A_j)\sqrt{\Sigma(t)}dZ^{\mathbb{Q}^i}(t)\right]}{\sqrt{Tr\left[(A_i - A_j)\Sigma(t)(A_i - A_j)\right]}}}_{=:dB_1(t)}.\end{aligned}$$

The dynamics of the variance is given by:

$$\begin{aligned} & dTr\left[(A_i - A_j)\Sigma(t)(A_i - A_j)\right] \\ & = \left(Tr\left[(A_i - A_j)\Omega\Omega^\top(A_i - A_j)\right] + 2Tr\left[(A_i - A_j)M\Sigma(t)(A_i - A_j)\right]\right)dt \\ & + 2Tr\left[(A_i - A_j)\sqrt{\Sigma}dW^{\mathbb{Q}^i}(t)Q(A_i - A_j)\right]. \end{aligned} \tag{4.4}$$

In order to derive the shape of the scalar brownian motion driving this dynamics we compute the following covariation:

$$\begin{aligned} & d\left\langle Tr\left[(A_i - A_j)\Sigma(t)(A_i - A_j)\right]\right\rangle \\ & = 4\left\langle Tr\left[(A_i - A_j)\sqrt{\Sigma(t)}dW^{\mathbb{Q}^i}(t)Q(A_i - A_j)\right], Tr\left[(A_i - A_j)\sqrt{\Sigma(t)}dW^{\mathbb{Q}^i}(t)Q(A_i - A_j)\right]\right\rangle \\ & = 4\left\langle \sum_{a,b,c,d,e=1}^{d}(A_i - A_j)_{ab}\sqrt{\Sigma(t)}_{bc}dW^{\mathbb{Q}^i}_{cd}(t)Q_{de}(A_i - A_j)_{ea}, \right. \\ & \left. \sum_{p,q,r,s,t=1}^{d}(A_i - A_j)_{pq}\sqrt{\Sigma(t)}_{qr}dW^{\mathbb{Q}^i}_{rs}(t)Q_{st}(A_i - A_j)_{tp}\right\rangle \\ & = \sum^{d}(A_i - A_j)_{ab}\sqrt{\Sigma(t)}_{br}Q_{se}(A_i - A_j)_{ea}(A_i - A_j)_{pq}\sqrt{\Sigma(t)}_{qr}Q_{st}(A_i - A_j)_{tp}dt \\ & = 4Tr\left[(A_i - A_j)^2\Sigma(t)(A_i - A_j)^2Q^\top Q\right]dt. \end{aligned} \tag{4.5}$$

Then we can express the dynamics of the variance as follows:

$$\begin{aligned} & dTr\left[(A_i - A_j)\Sigma(t)(A_i - A_j)\right] \\ & = (\ldots)\,dt + 2\sqrt{Tr\left[(A_i - A_j)^2\Sigma(t)(A_i - A_j)^2Q^\top Q\right]}\underbrace{\frac{Tr\left[(A_i - A_j)\sqrt{\Sigma}dW^{\mathbb{Q}^i}(t)Q(A_i - A_j)\right]}{\sqrt{Tr\left[(A_i - A_j)^2\Sigma(t)(A_i - A_j)^2Q^\top Q\right]}}}_{=:dB_2(t)}, \end{aligned}$$

so that we can finally compute the covariation between the two noises. Notice that in the calculation below we are assuming the invariance of the correlation with respect to the change of measure which was clarified in Section 3. The skew is computed as

$$
\begin{aligned}
&Corr_t(Noise(d\log S^{i,j}(t)), Noise(Vol(d\log S^{i,j}(t))))\\
&= d\left\langle B_1, B_2\right\rangle\\
&= \left\langle \frac{Tr\left[(A_i-A_j)\sqrt{\Sigma(t)}dZ_t^{\mathbb{Q}^i}\right]}{\sqrt{Tr\left[(A_i-A_j)\Sigma(t)(A_i-A_j)\right]}}, \frac{Tr\left[(A_i-A_j)\sqrt{\Sigma(t)}dW^{\mathbb{Q}^i}(t)Q(A_i-A_j)\right]}{\sqrt{Tr\left[(A_i-A_j)^2\Sigma(t)(A_i-A_j)^2Q^\top Q\right]}}\right\rangle\\
&= \left\langle \frac{Tr\left[(A_i-A_j)\sqrt{\Sigma(t)}dZ^{\mathbb{Q}^i}(t)\right]}{\sqrt{Tr\left[(A_i-A_j)\Sigma(t)(A_i-A_j)\right]}}, \frac{Tr\left[(A_i-A_j)\sqrt{\Sigma(t)}dZ^{\mathbb{Q}^i}(t)R^\top Q(A_i-A_j)\right]}{\sqrt{Tr\left[(A_i-A_j)^2\Sigma(t)(A_i-A_j)^2Q^\top Q\right]}}\right\rangle\\
&= \frac{\sum^d (A_i-A_j)_{pq}\sqrt{\Sigma(t)}_{qr}dZ_{rp}^{\mathbb{Q}^i}(t)(A_i-A_j)_{ab}\sqrt{\Sigma(t)}_{bc}dZ_{cd}^{\mathbb{Q}^i}(t)R_{de}^\top Q_{ef}(A_i-A_j)_{fa}}{\sqrt{Tr\left[(A_i-A_j)\Sigma(t)(A_i-A_j)\right]}\sqrt{Tr\left[(A_i-A_j)^2\Sigma(t)(A_i-A_j)^2Q^\top Q\right]}}\\
&= \frac{\sum^d (A_i-A_j)_{dq}\sqrt{\Sigma(t)}_{qc}\sqrt{\Sigma(t)}_{cb}(A_i-A_j)_{ba}(A_i-A_j)_{af}Q_{fe}^\top R_{ed}}{\sqrt{Tr\left[(A_i-A_j)\Sigma(t)(A_i-A_j)\right]}\sqrt{Tr\left[(A_i-A_j)^2\Sigma(t)(A_i-A_j)^2Q^\top Q\right]}}dt\\
&= \frac{Tr\left[(A_i-A_j)\Sigma(t)(A_i-A_j)^2Q^\top R\right]}{\sqrt{Tr\left[(A_i-A_j)\Sigma(t)(A_i-A_j)\right]}\sqrt{Tr\left[(A_i-A_j)^2\Sigma(t)(A_i-A_j)^2Q^\top Q\right]}}dt. \qquad (4.6)
\end{aligned}
$$

This quantity represents the skew in Equation (4.1). By looking at the numerator we realize that the proportionality relation (4.3) is satisfied.

**4.2. A stochastic variance-covariance matrix.** We would like to discuss the positive definiteness of the variance-covariance matrix. For simplicity, we consider the case of three currencies, meaning that we will have a $2\times 2$ candidate covariance matrix:

$$
\begin{pmatrix} \left\langle \frac{dS^{i,j}(t)}{S^{i,j}(t)}\right\rangle & \left\langle \frac{dS^{i,j}(t)}{S^{i,j}(t)}, \frac{dS^{i,l}(t)}{S^{i,l}(t)}\right\rangle \\ \left\langle \frac{dS^{i,j}(t)}{S^{i,j}(t)}, \frac{dS^{i,l}(t)}{S^{i,l}(t)}\right\rangle & \left\langle \frac{dS^{i,l}(t)}{S^{i,l}(t)}\right\rangle \end{pmatrix}. \qquad (4.7)
$$

We know that:

$$
\left\langle \frac{dS^{i,j}(t)}{S^{i,j}(t)}\right\rangle = Tr\left[(A_i-A_j)\,\Sigma(t)\,(A_i-A_j)\right]dt, \qquad (4.8)
$$

$$
\left\langle \frac{dS^{i,j}(t)}{S^{i,j}(t)}, \frac{dS^{i,l}(t)}{S^{i,l}(t)}\right\rangle = Tr\left[(A_i-A_j)\,\Sigma(t)\,(A_i-A_l)\right]dt. \qquad (4.9)
$$

We first look at (4.8). We recall that we assumed $A_i, A_j, A_l \in S_d$. Without loss of generality (otherwise put $V=-V'$ for $V'\in S_d^+$), let $(A_i-A_j)\in S_d^+$. Recall that the cone $S_d^+$ is self dual, meaning that:

$$
S_d^+ = \left\{u\in S_d \quad | \quad Tr\left[uv\right]\geq 0, \forall v\in S_d^+\right\}. \qquad (4.10)
$$

Let $O$ be an orthogonal matrix, then we may write: $(A_i-A_j)=O\Lambda O^\top$, where $\Lambda$ is a Diagonal matrix containing the eigenvalues of $(A_i-A_j)$ on the main Diagonal. Then we have:

$$
\begin{aligned}
Tr\left[(A_i-A_j)\,\Sigma(t)\,(A_i-A_j)\right] &= Tr\left[O\Lambda O^\top\Sigma(t)O\Lambda O^\top\right]\\
&= Tr\left[\Sigma(t)O\Lambda^2O^\top\right]\geq 0, \qquad (4.11)
\end{aligned}
$$

by self-duality. This shows that variances are positive. Now we would like to check that the variance-covariance matrix is in $S_d^+$. Now let

$$
\mathcal{M}(t) = (A_i-A_j)\sqrt{\Sigma(t)}, \qquad (4.12)
$$

$$
\mathcal{N}(t) = (A_i-A_l)\sqrt{\Sigma(t)}, \qquad (4.13)
$$

then, using the Cauchy-Schwarz inequality for matrices we have

$$Tr\left[(A_i - A_j)\,\Sigma(t)\,(A_i - A_j)\right] Tr\left[(A_i - A_l)\,\Sigma(t)\,(A_i - A_l)\right] \tag{4.14}$$

$$= Tr\left[\mathcal{M}(t)\mathcal{M}^\top(t)\right] Tr\left[\mathcal{N}(t)\mathcal{N}^\top(t)\right] \tag{4.15}$$

$$\geq Tr\left[\mathcal{M}(t)\mathcal{N}^\top(t)\right]^2 \tag{4.16}$$

$$= Tr\left[(A_i - A_j)\,\Sigma(t)\,(A_i - A_l)\right]^2. \tag{4.17}$$

This implies that the determinant of the instantaneous variance-covariance matrix is positive, so we conclude that the variance-covariance matrix is well defined, and, as a side effect, we have the usual bound for the correlations, i.e. all correlations are bounded by one (in absolute value).

## 5. Option pricing

As in the multi-Heston case, in this section we provide the calculation of the Laplace transform and the characteristic function of $x^{i,j}(t) := \log S^{i,j}(t)$, which will be useful for option pricing purposes. Let us consider a call option $C(S^{i,j}(t), K^{i,j}, \tau), i, j = 1, .., N, i \neq j$, on a generic FX rate $S^{i,j}(t) = \exp(x^{i,j}(t))$ with strike $K^{i,j}$, maturity $T$ ($\tau = T - t$ is the time to maturity) and face equal to one unit of the foreign currency. For ease of notation set: $R^{\mathbb{Q}^i} = R$ and $Q^{\mathbb{Q}^i} = Q$ and the shorthand $M^{\mathbb{Q}^i} = \tilde{M}$. We proceed to prove the following: Being an affine model, the characteristic function conditioned on the initial values

$$\phi^{i,j}(\omega, t, \tau, x, \Sigma) = \mathbb{E}_t^{\mathbb{Q}^i}[e^{\mathtt{i}\omega x^{i,j}(T)} | x^{i,j}(t) = x, \Sigma(t) = \Sigma] \tag{5.1}$$

can be derived analytically (here $\mathtt{i} = \sqrt{-1}$). Standard numerical integration methods can then be used to invert the Fourier transform to obtain the probability density at $T$ or the vanilla price via integration against the payoff, with overall little computational effort. In fact, from the usual risk-neutral argument, the initial price of the call option can be written as (domestic) risk neutral expected value:

$$C(S^{i,j}(t), K^{i,j}, \tau) = e^{-r^i\tau}\mathbb{E}_t^{\mathbb{Q}^i}\left[\left(e^{x^{i,j}(T)} - K^{i,j}\right)^+\right],$$

and by applying standard arguments (see e.g. Carr and Madan (1999), Bakshi and Madan (2000), Duffie et al. (2000) and Sepp (2003)) it can be expressed in terms of the integral of the product of the Fourier transform of the payoff and the characteristic function of the log-asset price:

$$C(S^{i,j}(t), K^{i,j}, \tau) = e^{-r^i\tau}\frac{1}{2\pi}\int_{\mathcal{Z}} \phi^{i,j}(-\mathtt{i}\lambda, t, \tau, x, \Sigma)\Phi(\lambda)d\lambda, \tag{5.2}$$

where

$$\Phi(\lambda) = \int_{\mathcal{Z}} e^{\mathtt{i}\lambda x}\left(e^x - K^{i,j}\right)^+ dx$$

is the Fourier transform of the payoff function and $\mathcal{Z}$ denotes the strip of regularity of the payoff, that is the admissible domain where the integral in (5.2) is well defined. In other words, the pricing problem is essentially solved once the (conditional) characteristic function of the log-exchange rate is known. We recall the relationship between the characteristic function and the moment generating function. If we denote via $G^{i,j}(\omega, t, \tau, x, \Sigma)$ the moment generating function, given by

$$G^{i,j}(\omega, t, \tau, x, \Sigma) = \mathbb{E}_t^{\mathbb{Q}^i}[e^{\omega x^{i,j}(T)} | x^{i,j}(t) = x, \Sigma(t) = \Sigma], \tag{5.3}$$

we simply have $\phi^{i,j}(\omega, t, \tau, x, \mathbf{V}) = G^{i,j}(\mathtt{i}\omega, t, \tau, x, \Sigma)$. Consequence, in our affine model, it is sufficient to derive the Laplace transform, which is explicitly given by the following proposition:

PROPOSITION 5.2. *Assume that the family of matrices $A$ is symmetric then, in the Wishart model, the Laplace Transform of $x^{i,j}(t) := \log S^{i,j}(t)$ is given by:*

$$G^{i,j}(\omega, t, T, x, \Sigma) = exp\left[\omega x + \mathcal{A}(\tau) + Tr\left[\mathcal{B}(\tau)\Sigma\right]\right], \tag{5.4}$$

*where:*

$$\mathcal{A} = \omega\left(r^i - r^j\right)\tau - \frac{\beta}{2}Tr\left[\log \mathcal{F}(\tau) + \left(\tilde{M}^\top + \omega\left(A_i - A_j\right)R^\top Q\right)\tau\right], \tag{5.5}$$

$$\mathcal{B}(\tau) = \mathcal{B}_{22}(\tau)^{-1}\mathcal{B}_{21}(\tau) \tag{5.6}$$

*and $\mathcal{B}_{22}(\tau), \mathcal{B}_{21}(\tau)$ are submatrices in:*

(5.7)
$$\begin{pmatrix} \mathcal{B}_{11}(\tau) & \mathcal{B}_{12}(\tau) \\ \mathcal{B}_{21}(\tau) & \mathcal{B}_{22}(\tau) \end{pmatrix} = \exp\tau\begin{bmatrix} \tilde{M} + \omega Q^\top R\left(A_i - A_j\right) & -2Q^\top Q \\ \frac{\omega^2 - \omega}{2}\left(A_i - A_j\right)^2 & -\left(\tilde{M}^\top + \omega\left(A_i - A_j\right)R^\top Q\right) \end{bmatrix}.$$

PROOF. See Section 9. □

As we stated above, for $\omega = i\lambda$, we obtain the characteristic function of the log exchange rate, hence we can compute option prices as e.g. in Carr and Madan (1999) via:

$$\frac{exp\left\{-\alpha c\right\}}{2\pi}\int_{-\infty}^{+\infty} e^{-\mathtt{i}vc}\varphi(v)dv, \tag{5.8}$$

where:

$$\varphi(v) = e^{-r^i(T-t)}\frac{G\left(\mathtt{i}\left(v - (\alpha+1)\,\mathtt{i}\right), t, T, x, V\right)}{\left(\alpha + \mathtt{i}v\right)\left(1 + \alpha + \mathtt{i}v\right)}. \tag{5.9}$$

## 6. Expansions

As in the previous chapter, we will perform in the sequel the calibration of our model, which can be obtained by relying on a standard non-linear least squares procedure. This will be employed to minimize the distance between model and market implied volatilities. The model implied volatilities are extracted from the prices produced by the FFT routine. This procedure for the Wishart model is more demanding from a numerical point of view than the analogous one for the multi-Heston case. An alternative approach is to fit implied volatilities via a simpler function. A possibility is to find a relationship between the prices produced by the model, and the standard Black-Scholes formula. The next result states that it is possible to approximate the prices of options under the Wishart model, via a suitable expansion of the standard Black-Scholes formula and its derivatives, analogously to what has been done in the previous chapter. The proof, which is reported in the appendix, relies on arguments which may be found in Lewis (2000) and Da Fonseca and Grasselli (2011) (we drop all currency indices, it is intended that we are considering the $(i, j)$ FX pair). Define $\tau := T - t$ and let us define the real deterministic functions $\tilde{\mathcal{B}}^0, \tilde{\mathcal{B}}^1, \tilde{\mathcal{B}}^{20}, \tilde{\mathcal{B}}^{21}$ as follows:

$$\tilde{\mathcal{B}}^0 = \int_0^\tau e^{(\tau-u)\tilde{M}^\top}\left(A_i - A_j\right)e^{(\tau-u)\tilde{M}}du, \tag{6.1}$$

$$\tilde{\mathcal{B}}^1 = \int_0^\tau e^{(\tau-u)\tilde{M}^\top}\left(\tilde{\mathcal{B}}^0(u)Q^\top R\left(A_i - A_j\right) + \left(A_i - A_j\right)R^\top Q\tilde{\mathcal{B}}^0(u)\right)e^{(\tau-u)\tilde{M}}du, \tag{6.2}$$

$$\tilde{\mathcal{B}}^{20} = \int_0^\tau e^{(\tau-u)\tilde{M}^\top}2\tilde{\mathcal{B}}^0(u)Q^\top Q\tilde{\mathcal{B}}^0(u)e^{(\tau-u)\tilde{M}}du, \tag{6.3}$$

$$\tilde{\mathcal{B}}^{21} = \int_0^\tau e^{(\tau-u)\tilde{M}^\top}\left(\tilde{\mathcal{B}}^1(u)Q^\top R\left(A_i - A_j\right) + \left(A_i - A_j\right)R^\top Q\tilde{\mathcal{B}}^1(u)\right)e^{(\tau-u)\tilde{M}}du. \tag{6.4}$$

Moreover, the real deterministic scalar functions $\tilde{\mathcal{A}}^0(\tau), \tilde{\mathcal{A}}^1(\tau), \tilde{\mathcal{A}}^{20}(\tau), \tilde{\mathcal{A}}^{21}(\tau)$ are given by:

$$\tilde{\mathcal{A}}^0(\tau) = Tr\left[\Omega\Omega^\top \int_0^\tau \tilde{\mathcal{B}}^0(u)du\right], \tag{6.5}$$

$$\tilde{\mathcal{A}}^1(\tau) = Tr\left[\Omega\Omega^\top \int_0^\tau \tilde{\mathcal{B}}^1(u)du\right], \tag{6.6}$$

$$\tilde{\mathcal{A}}^{20}(\tau) = Tr\left[\Omega\Omega^\top \int_0^\tau \tilde{\mathcal{B}}^{20}(u)du\right], \tag{6.7}$$

$$\tilde{\mathcal{A}}^{21}(\tau) = Tr\left[\Omega\Omega^\top \int_0^\tau \tilde{\mathcal{B}}^{21}(u)du\right]. \tag{6.8}$$

Finally, let

$$v = \sigma^2\tau = \tilde{\mathcal{A}}^0(\tau) + Tr\left[\tilde{\mathcal{B}}^0(\tau)\Sigma\right] \tag{6.9}$$

be the integrated variance.

PROPOSITION 5.3. *Assume that the vol of vol matrix $Q$ has been scaled by the factor $\alpha > 0$. Then the call price $C(S(t), K, \tau)$ in the Wishart-based exchange model can be approximated in terms of the vol of vol scale factor $\alpha$ by differentiating the Black Scholes formula $C_{B\&S}(S(t), K, \sigma, \tau)$ with respect to the log exchange rate $x(t) = \ln S(t)$ and the integrated variance $v = \sigma^2\tau$:*

$$\begin{aligned} C(S(t), K, \tau) \approx\ & C_{B\&S}(S(t), K, \sigma, \tau) \\ & + \alpha\left(\tilde{\mathcal{A}}^1(\tau) + Tr\left[\tilde{\mathcal{B}}^1(\tau)\Sigma(t)\right]\right)\partial^2_{xv} C_{B\&S}(S(t), K, \sigma, \tau) \\ & + \alpha^2\left(\tilde{\mathcal{A}}^{20}(\tau) + Tr\left[\tilde{\mathcal{B}}^{20}(\tau)\Sigma(t)\right]\right)\partial^2_{v^2} C_{B\&S}(S(t), K, \sigma, \tau) \\ & + \alpha^2\left(\tilde{\mathcal{A}}^{21}(\tau) + Tr\left[\tilde{\mathcal{B}}^{21}(\tau)\Sigma(t)\right]\right)\partial^3_{x^2v} C_{B\&S}(S(t), K, \sigma, \tau) \\ & + \frac{\alpha^2}{2}\left(\tilde{\mathcal{A}}^1(\tau) + Tr\left[\tilde{\mathcal{B}}^1(\tau)\Sigma(t)\right]\right)^2\partial^4_{x^2v^2} C_{B\&S}(S(t), K, \sigma, \tau), \end{aligned} \tag{6.10}$$

PROOF. See Section 9 □

Finally, as in the multi-Heston case, we can state another formula, which does not involve the computation of option prices and constitutes an approximation of the implied volatility surface for a short time to maturity. This formula may constitute a useful alternative in order to get a quicker calibration for short maturities. The proof is again provided in detail in the appendix.

PROPOSITION 5.4. *For a short time to maturity the implied volatility expansion in terms of the vol-of-vol scale factor $\alpha$ in the Wishart-based exchange model is given by:*

$$\begin{aligned} \sigma^2_{imp} \approx\ & Tr\left[(A_i - A_j)\Sigma(A_i - A_j)\right] + \alpha\frac{Tr\left[(A_i - A_j)Q^\top R(A_i - A_j)\Sigma(t)\right]m_f}{Tr\left[(A_i - A_j)\Sigma(t)(A_i - A_j)\right]} \\ & + \alpha^2\frac{m_f^2}{Tr\left[(A_i - A_j)\Sigma(t)(A_i - A_j)\right]^2}\left[\frac{1}{3}Tr\left[(A_i - A_j)Q^\top Q(A_i - A_j)\Sigma(t)\right]\right. \\ & + \frac{1}{3}Tr\Big[\left[(A_i - A_j)Q^\top R(A_i - A_j) + (A_i - A_j)R^\top Q(A_i - A_j)\right] \\ & \left.\times Q^\top R(A_i - A_j)\Sigma\Big] - \frac{5}{4}\frac{Tr\left[(A_i - A_j)Q^\top R(A_i - A_j)\Sigma\right]^2}{Tr\left[(A_i - A_j)\Sigma(t)(A_i - A_j)\right]}\right]. \end{aligned} \tag{6.11}$$

*where, as in Chapter 4, $m_f = \log\left(\frac{S^{i,j}(t)e^{(r_i - r_j)\tau}}{K}\right)$ denotes the log-moneyness.*

PROOF. See Section 9 □

## 7. Calibration to market data

We perform the calibration along the same lines as in Chapter 4. This means that we will be minimizing the squared distance between market and model implied volatilities. We perform the calibration experiment on the same data set as in Chapter 4. To keep the exposition of the present chapter self-contained, we report the details.

**7.1. Description of the data.** We consider implied volatility surfaces for USD/EUR, USD/JPY and EUR/JPY. With these currencies we are able to construct a triangular relation between rates. We consider market data as of 23rd July 2010. Our sample includes expiry dates ranging from 3 days to 5 years. It is important to stress that in the forex market implied volatilities surfaces are expressed in terms of maturity and delta: the market practice is to quote volatilities for strangles and risk reversals which can then be employed to reconstruct a whole surface of implied volatilities via an interpolation method.

**7.2. Calibration results.** For the calibration of the model we use a non-linear least-squares optimizer to minimize the following function:

$$\sum_i \left(\sigma^{imp}_{i,mkt} - \sigma^{imp}_{i,model}\right)^2 \tag{7.1}$$

The choice of this norm constitutes the market practice. The use of a norm in price should be avoided as the numerical range for option prices may be large, thus introducing a bias in the optimization. To be more precise, the construction of the objective function is performed along the following steps. First we consider a function implementing the Fourier/Laplace transform, i.e. formulas (5.4), (5.5), (5.6) and (5.7). The Fourier/Laplace transform is then invoked by an FFT pricing routine which implements the Carr-Madan methodology, thus returning a surface of prices for different moneyness and maturities, finally we construct the implied volatility surface produced by the model by using this surface of prices as input for a standard Black-Scholes implied volatility solver. For a more detailed discussion, see Da Fonseca and Grasselli (2011).

We would like to provide some practical suggestions for the implementation of the calibration procedure, which is more delicate in this case with respect to the multi-Heston model. In Algorithm 7.1, we provide an excerpt of the procedure we employ to price options, which implements the Carr and Madan (1999) methodology. A direct inspection of the vector $ncf$, defined in Algorithm 7.1, reveals that, for larger maturities and for large values of $v(j)$ (the point where the characteristic function is computed), this quantity decays quickly. We can employ this fact to speed up the computation of the prices as follows: we fix a threshold $\epsilon = 10^{-20}$, and stop the computation of the characteristic function as soon as the absolute value of $ncf(j)$ satisfies $|ncf(j)| \leq \epsilon$. This procedure has been adopted in Da Fonseca and Grasselli (2011). Another issue that should be taken into consideration is the presence of some numerical instabilities. According to Da Fonseca (2011), these numerical instabilities are observed when the value of the characteristic function is small but the individual blocks of the matrix exponential are large. A possible solution, adopted in Da Fonseca and Grasselli (2011), is to calculate the characteristic function at the point where the instability is observed by relying on the Runge-Kutta method, (see e.g. Chapter 1). The alternative approach that we employ is trivial but faster: we simply approximate $v(j)$ with $v(j-1)$, which corresponds to a simple but effective interpolation approach.

**Algorithm 7.1:** FFTPRICER($params$)

$\epsilon = 1e - 20;$
$N = 4096;$
$\eta = 0.18;$
$\lambda = (2\pi)/(N\eta);$
$v = 0 : \eta : ((N-1)\eta);$
$b = (N\lambda)/2;$
$ncf = zeros(1, length(v));$

**for** $j \leftarrow 1$ **to** $length(v)$
  $ncf(j) = e^{-r(T-t)} G\left(i\left(v(j) - (\alpha+1)i\right), t, T, x, V\right);$(v(j))
  **if** ABS(ncf(j)) $\leq \epsilon$
    **then** $ncf(j+1 : end) = 0;$ $break;$
  **if** isNaN(ncf(j)) $== TRUE$
    **then** $ncf(j) = ncf(j-1);$

We report in Table 1 the parameters that we obtained by performing the calibration first on the whole sample ranging from 1 month up to 1 year, and then on various subsamples obtained by excluding more and more longer maturities. In all cases we are able to obtain a very good fit of market implied volatilities. The result of the fit for the whole sample that we consider can be appreciated by looking at Figures 1, 2 and 3. This is a result that we expected since the present model is a generalization of the framework that we introduced in chapter 4. In Table 2 we report the relative change in the values of the parameters when we consider the various sub-samples. Unfortunately, for some parameters, large variations are observed. This suggests that the flexibility of the Wishart model is not justified if we are only interested in calibrating vanilla prices, hence we strongly believe that the added value of our approach should be expected in the context of the evaluation of more complex derivatives. Clearly, it is possible to perform a calibration on a reduced form of the model, as we showed in Chapter 4. Anyhow it is interesting to note that we are in front of a very general model, which is able to fit again three FX implied volatility surfaces, while retaining a large amount of flexibility which may be employed to capture other stylized facts concerning the market or some particular products.

## 8. Conclusions

In this chapter we showed that the multi-Heston model admits a generalization giving rise to the Wishart FX model. The Wishart FX model retains the same level of analytical tractability of the multi-Heston approach. In particular, we are still considering a model which is theoretically consistent with triangular relationship among FX rates. The model is highly analytically tractable and is consistent with respect to spot inversion. Finally, and more importantly, the model features clear a clear relationship between the parameters under the various risk neutral measure coexisting on the market. Thanks to this relationship we are able to perform again a simultaneous calibration of the model to three FX implied volatility surfaces. The fit is again very satisfactory.

## 9. Proofs

**9.1. Proof of Proposition 5.1.** Application of the Ito formula to the product $S^{i,l}(t)S^{l,j}(t)$ (using the property $Tr[dWA]Tr[dWB] = Tr[AB]$) leads to (2.5). In formulas:

$$\frac{dS^{i,l}(t)}{S^{i,l}(t)} = (r^i - r^l)dt + Tr[(A_i - A_l)\Sigma(t)A_i]dt + Tr[(A_i - A_l)\sqrt{\Sigma(t)}dZ(t)],$$
$$\frac{dS^{l,j}(t)}{S^{l,j}(t)} = (r^l - r^j)dt + Tr[(A_l - A_j)\Sigma(t)A_l]dt + Tr[(A_l - A_j)\sqrt{\Sigma(t)}dZ(t)],$$

$$\begin{aligned} dS^{i,j}(t) &= dS^{i,l}(t)_t S^{l,j}_t + S^{i,l}(t)_t dS^{l,j}_t + d\left\langle S^{i,l}(t)_t, S^{l,j}_t\right\rangle \\ &= S^{l,j}_t S^{i,l}(t)\left((r^i - r^l)dt + Tr[(A_i - A_l)\Sigma(t)A_i]dt + Tr[(A_i - A_l)\sqrt{\Sigma(t)}dZ(t)]\right) \\ &+ S^{l,j}_t S^{i,l}(t)\left((r^l - r^j)dt + Tr[(A_l - A_j)\Sigma(t)A_l]dt + Tr[(A_l - A_j)\sqrt{\Sigma(t)}dZ(t)]\right) \\ &+ \left\langle Tr[(A_i - A_l)\sqrt{\Sigma(t)}dZ(t)], Tr[(A_l - A_j)\sqrt{\Sigma(t)}dZ(t)].\right\rangle \end{aligned} \tag{9.1}$$

We concentrate on the covariation term:

$$\begin{aligned} &\left\langle Tr[(A_i - A_l)\sqrt{\Sigma(t)}dZ(t)], Tr[(A_l - A_j)\sqrt{\Sigma(t)}dZ(t)]\right\rangle \\ &= \left\langle \sum_{p,q,r=1}^{d} (A_i - A_l)_{pq}\sqrt{\Sigma(t)}_{qr}dZ(t)_{rp}, \sum_{s,t,u=1}^{d} (A_l - A_j)_{st}\sqrt{\Sigma(t)}_{tu}dZ(t)_{us}\right\rangle \\ &= \sum_{p,q,r,s,t,u=1}^{d} (A_i - A_l)_{pq}\sqrt{\Sigma(t)}_{qr}dZ(t)_{rp}(A_l - A_j)_{st}\sqrt{\Sigma(t)}_{tu}dZ(t)_{us}\delta_{r=u}\delta_{p=s}dt \\ &= \sum_{s,q,u,t=1}^{d} (A_i - A_l)_{sq}\sqrt{\Sigma(t)}_{qu}(A_l - A_j)_{st}\sqrt{\Sigma(t)}_{tu}dt \\ &= \sum_{s,q,u,t=1}^{d} (A_i - A_l)_{sq}\sqrt{\Sigma(t)}_{qu}\sqrt{\Sigma(t)}_{ut}(A_l^\top - A_j^\top)_{ts}dt. \end{aligned} \tag{9.2}$$

By assuming that the matrices $A$ are symmetric we get that:

$$\left\langle Tr[(A_i - A_l)\sqrt{\Sigma(t)}dZ(t)], Tr[(A_l - A_j)\sqrt{\Sigma(t)}dZ(t)]\right\rangle = Tr\left[(A_i - A_l)\Sigma(t)(A_l - A_j)\right]. \tag{9.3}$$

Finally, using the fact that terms in the trace commute we obtain:

$$\frac{dS^{i,j}(t)}{S^{i,j}(t)} = (r^i - r^j)dt + Tr[(A_i - A_j)\Sigma(t)A_i]dt + Tr[(A_i - A_j)\sqrt{\Sigma(t)}dZ(t)] \tag{9.4}$$

as desired.

**9.2. Proof of Proposition 5.2.** Recall that $\phi^{i,j}(\omega, t, \tau, x, \Sigma) = G^{i,j}(i\omega, t, \tau, x, \Sigma)$ where $G$ was defines as $G^{i,j}(\omega, t, \tau, x, \Sigma) = \mathbb{E}_t^{\mathbb{Q}^i}[e^{\omega x_T}]$, $x^{i,j}(t) := \log S^{i,j}(t)$. These functions represent the characteristic and the moment generating functions of the log-exchange rate. Following Da Fonseca et al. (2008), in order to determine these quantities, we first write the PDE satisfied by $G$. First of all we compute the dynamics of $x(t) = x^{i,j}(t)$ under the measure $\mathbb{Q}^i$:

$$d\log S^{i,j}(t) = \left(\left(r^i - r^j\right) - \frac{1}{2}Tr\left[\left(A_i - A_j\right)\Sigma(t)\left(A_i - A_j\right)\right]\right)dt + Tr\left[\left(A_i - A_j\right)\sqrt{\Sigma(t)}dZ^{\mathbb{Q}^i}(t)\right], \tag{9.5}$$

As before, the Laplace Transform solves the following PDE of backward Kolmogorov type in terms of $\tau = T - t$:

$$\frac{\partial}{\partial \tau}G^{i,j} = \mathcal{A}_{x,\Sigma}G^{i,j}. \tag{9.6}$$

$$G^{i,j}(\omega, T, 0, x, \Sigma) = e^{\omega x} \tag{9.7}$$

To solve this PDE we first determine the infinitesimal generator $\mathcal{A}_{x,\Sigma}$. This will feature the contribution of three terms: the process $x$, the Wishart process $\Sigma$ and the mixed term which corresponds to the coefficient of the term $\frac{\partial^2}{\partial x \partial \Sigma_{pt}}$ and arises from the correlation structure. The first is trivial, the second is known to us thanks to Bru (1991) and is of the form:

$$Tr\left[\left(\Omega\Omega^\top + \tilde{M}\Sigma + \Sigma\tilde{M}^\top\right)D + 2\Sigma D Q^\top Q D\right], \tag{9.8}$$

where $D$ is the differential operator:

$$D_{pt} = \frac{\partial}{\partial \Sigma_{pt}}. \tag{9.9}$$

We calculate the mixed term. To this end we notice that under the measure $\mathbb{Q}^i$:

$$\begin{aligned}
&d\left\langle \log S^{i,j}(t), d\Sigma(t)_{pt}\right\rangle \\
&= 2\left\langle Tr\left[\left(A_i - A_j\right)\sqrt{\Sigma(t)}dZ^{\mathbb{Q}^i}(t)\right], \sum_{q,r,s=1}^{d}\sqrt{\Sigma(t)}_{pq}dZ^{\mathbb{Q}^i}_{qr}(t)R^\top_{rs}Q_{st}\right\rangle \\
&= 2\left\langle \sum_{a,b,c,=1}^{d}\left(A_i - A_j\right)_{ab}\sqrt{\Sigma(t)}_{bc}dZ^{\mathbb{Q}^i}_{ca}, \sum_{q,r,s=1}^{d}\sqrt{\Sigma}_{pq}dZ^{\mathbb{Q}^i}_{qr}R^\top_{rs}Q_{st}\right\rangle \delta_{c=q}\delta_{a=r}dt \\
&= 2\sum_{b,q,r,s=1}^{d}\left(A_i - A_j\right)_{rb}\sqrt{\Sigma(t)}_{bq}\sqrt{\Sigma(t)}_{pq}R^\top_{rs}Q_{st}dt \\
&= 2\sum_{b,q,r,s=1}^{d}\Sigma(t)_{pb}\left(A_i - A_j\right)_{br}R^\top_{rs}Q_{st}dt,
\end{aligned}$$

where we have used the fact that

$$2Tr\left[\Sigma\left(A_i - A_j\right)R^\top Q D\right]\frac{\partial}{\partial x} = 2\sum_{p,b,r,s=1}^{d}D_{tp}\Sigma_{pb}\left(A_i - A_j\right)_{br}R^\top_{rs}Q_{st}\frac{\partial}{\partial x} \tag{9.10}$$

and that $D$ is symmetric. Now we can state the PDE satisfied by $G$:

$$\begin{aligned}
\frac{\partial G}{\partial \tau} &= \left(\left(r^i - r^j\right) - \frac{1}{2}Tr\left[\left(A_i - A_j\right)\Sigma\left(A_i - A_j\right)\right]\right)\frac{\partial G}{\partial x} \\
&+ \frac{1}{2}Tr\left[\left(A_i - A_j\right)\Sigma\left(A_i - A_j\right)\right]\frac{\partial^2 G}{\partial x^2} \\
&+ Tr\left[\left(\Omega\Omega^\top + \tilde{M}\Sigma + \Sigma\tilde{M}^\top\right)DG + 2\left(\Sigma D Q^\top Q D\right)G\right] \\
&+ 2Tr\left[\Sigma\left(A_i - A_j\right)R^\top Q D\right]\frac{\partial G}{\partial x}.
\end{aligned} \tag{9.11}$$

The Wishart process $\Sigma$ belongs to the class of Affine Processes defined on the cone of positive semi-definite matrices, meaning that we can make a guess of the following kind:

$$G^{i,j}(\omega,t,\tau,x,\Sigma) = exp\left[\mathcal{C}(\tau)x + \mathcal{A}(\tau) + Tr\left[\mathcal{B}(\tau)\Sigma\right]\right], \tag{9.12}$$

with $\mathcal{C},\mathcal{A} \in \mathbb{R}$ and $\mathcal{B} \in \mathcal{S}_d$ s.t. the transform is well defined, moreover these functions satisfy the following terminal conditions:

$$\mathcal{A}(0) = 0 \in \mathbb{R}, \tag{9.13}$$

$$\mathcal{C}(0) = \omega \in \mathbb{R}, \tag{9.14}$$

$$\mathcal{B}(0) = 0 \in \mathcal{S}_d. \tag{9.15}$$

We substitute the candidate (9.12) into (9.11) and then:

$$\begin{aligned}
&\frac{\partial}{\partial\tau}\mathcal{C}(\tau)x + \frac{\partial}{\partial\tau}\mathcal{A}(\tau) + Tr\left[\frac{\partial}{\partial\tau}\mathcal{B}(\tau)\Sigma\right] = \\
&\mathcal{C}(\tau)\left(\left(r^i - r^j\right) - \frac{1}{2}Tr\left[\left(A_i - A_j\right)\Sigma\left(A_i - A_j\right)\right]\right) \\
&+ \frac{1}{2}Tr\left[\left(A_i - A_j\right)\Sigma\left(A_i - A_j\right)\right]\mathcal{C}^2(\tau) \\
&+ Tr\left[\left(\Omega\Omega^\top + \tilde{M}\Sigma + \Sigma\tilde{M}^\top\right)\mathcal{B}(\tau) + 2\Sigma\mathcal{B}(\tau)Q^\top Q\mathcal{B}(\tau)\right] \\
&+ Tr\left[\Sigma\left(A_i - A_j\right)R^\top Q\mathcal{B}(\tau)\mathcal{C}(\tau)\right] + Tr\left[\mathcal{C}(\tau)\mathcal{B}(\tau)Q^\top R\Sigma\left(A_i - A_j\right)\right].
\end{aligned} \tag{9.16}$$

We identify terms and deduce that:

$$\frac{\partial}{\partial\tau}\mathcal{C}(\tau) = 0, \tag{9.17}$$

hence: $\mathcal{C}(\tau) = \omega\ \forall\tau$. We also have the following (matrix) Riccati ODE:

$$\begin{aligned}
\frac{\partial}{\partial\tau}\mathcal{B} = {} & \mathcal{B}(\tau)\left(\tilde{M} + \omega Q^\top R\left(A_i - A_j\right)\right) + \left(\tilde{M}^\top + \omega\left(A_i - A_j\right)R^\top Q\right)\mathcal{B}(\tau) + 2\mathcal{B}(\tau)Q^\top Q\mathcal{B}(\tau) \\
& + \frac{\omega^2 - \omega}{2}\left(A_i - A_j\right)^2
\end{aligned} \tag{9.18}$$

and the final ODE which may then be solved upon direct integration of:

$$\frac{\partial}{\partial\tau}\mathcal{A} = \omega\left(r^i - r^j\right) + Tr\left[\Omega\Omega^\top\mathcal{B}(\tau)\right]. \tag{9.19}$$

It is possible to linearize (9.18). In fact, following Grasselli and Tebaldi (2008) and writing:

$$\mathcal{B}(\tau) = \mathcal{F}^{-1}(\tau)\mathcal{G}(\tau), \tag{9.20}$$

for $\mathcal{F}(\tau) \in GL(d)$ and $\mathcal{G}(\tau) \in M_d$, then we have

$$\begin{aligned}
\frac{\partial}{\partial\tau}\left[\mathcal{F}(\tau)\mathcal{B}(\tau)\right] &= \frac{\partial}{\partial\tau}\left[\mathcal{F}(\tau)\right]\mathcal{B}(\tau) + \mathcal{F}(\tau)\frac{\partial}{\partial\tau}\mathcal{B}(\tau) \\
&= \frac{\partial}{\partial\tau}\left[\mathcal{F}(\tau)\right]\mathcal{B}(\tau) + \\
&+ \mathcal{F}(\tau)\Big(\mathcal{B}(\tau)\left(\tilde{M} + \omega Q^\top R\left(A_i - A_j\right)\right) + \left(\tilde{M}^\top + \omega\left(A_i - A_j\right)R^\top Q\right)\mathcal{B}(\tau) \\
&+ 2\mathcal{B}(\tau)Q^\top Q\mathcal{B}(\tau)\Big),
\end{aligned} \tag{9.21}$$

which gives rise to the following system of ODE's

$$\frac{\partial}{\partial \tau}\mathcal{F} = -\mathcal{F}(\tau)\left(\tilde{M}^\top + \omega\left(A_i - A_j\right)R^\top Q\right) - 2\mathcal{G}(\tau)Q^\top Q \tag{9.22}$$

$$\frac{\partial}{\partial \tau}\mathcal{G} = \mathcal{G}(\tau)\left(\tilde{M} + \omega Q^\top R\left(A_i - A_j\right)\right) + \mathcal{F}(\tau)\frac{\omega^2 - \omega}{2}\left(A_i - A_j\right)^2, \tag{9.23}$$

with $\mathcal{F}(0) = I_d$ and $\mathcal{G}(0) = \mathcal{B}(0)$. The solution of the above system is

$$\begin{aligned} &(\mathcal{B}(0), I_d)\begin{pmatrix} \mathcal{B}_{11}(\tau) & \mathcal{B}_{12}(\tau) \\ \mathcal{B}_{21}(\tau) & \mathcal{B}_{22}(\tau) \end{pmatrix} \\ &= (\mathcal{B}(0), I_d)\exp\tau\begin{bmatrix} \tilde{M} + \omega Q^\top R\left(A_i - A_j\right) & -2Q^\top Q \\ \frac{\omega^2-\omega}{2}\left(A_i - A_j\right)^2 & -\left(\tilde{M}^\top + \omega\left(A_i - A_j\right)R^\top Q\right) \end{bmatrix} \end{aligned} \tag{9.24}$$

and so the solution for $\mathcal{B}(\tau)$ is

$$\mathcal{B}(\tau) = \left(\mathcal{B}(0)\mathcal{B}_{12}(\tau) + \mathcal{B}_{22}(\tau)\right)^{-1}\left(\mathcal{B}(0)\mathcal{B}_{11}(\tau) + \mathcal{B}_{21}(\tau)\right) \tag{9.25}$$

and since $\mathcal{B}(0) = 0$ we finally have: $\mathcal{B}(\tau) = \mathcal{B}_{22}(\tau)^{-1}\mathcal{B}_{21}(\tau)$. What we need to do now is to compute the solution of the last ODE. Instead of performing a direct integration we follow Da Fonseca et al. (2008) and proceed as follows: we start from (9.22) and write:

$$-\frac{1}{2}\left(\frac{\partial}{\partial \tau}\mathcal{F} + \mathcal{F}(\tau)\left(\tilde{M}^\top + \omega\left(A_i - A_j\right)R^\top Q\right)\right)\left(Q^\top Q\right)^{-1} = \mathcal{G}(\tau). \tag{9.26}$$

We plug this into (9.20), then we insert the resulting formula for $\mathcal{B}(\tau)$ into (9.19) and so we have:

$$\frac{\partial}{\partial \tau}\mathcal{A} = \omega\left(r^i - r^j\right) + Tr\left[-\frac{\beta}{2}\left(\mathcal{F}^{-1}(\tau)\frac{\partial}{\partial \tau}\mathcal{F} + \left(\tilde{M}^\top + \omega\left(A_i - A_j\right)R^\top Q\right)\right)\right]$$

hence the solution is:

$$\mathcal{A} = \omega\left(r^i - r^j\right)\tau - \frac{\beta}{2}Tr\left[\log\mathcal{F}(\tau) + \left(\tilde{M}^\top + \omega\left(A_i - A_j\right)R^\top Q\right)\tau\right]. \tag{9.27}$$

**9.3. Proof of Proposition 5.3.** We proceed as in Chapter 4, Proposition 4.4. The matrix Riccati ODE (9.18) may be rewritten as follows after replacing $Q$ by a small perturbation $\alpha Q, \alpha \in \mathbb{R}$:

$$\begin{aligned} \frac{\partial}{\partial \tau}\mathcal{B} &= \mathcal{B}(\tau)\left(\tilde{M} + \alpha\omega Q^\top R\left(A_i - A_j\right)\right) + \left(\tilde{M}^\top + \omega\alpha\left(A_i - A_j\right)R^\top Q\right)\mathcal{B}(\tau) \\ &\quad + 2\alpha^2\mathcal{B}(\tau)Q^\top Q\mathcal{B}(\tau) + \frac{\omega^2 - \omega}{2}\left(A_i - A_j\right)^2 \end{aligned} \tag{9.28}$$

$$\mathcal{B}(0) = 0. \tag{9.29}$$

We consider now an expansion in terms of $\alpha$ of the form $\mathcal{B} = \mathcal{B}^0 + \alpha\mathcal{B}^1 + \alpha^2\mathcal{B}^2$. We substitute this expansion and identify terms by powers of $\alpha$. We obtain the following ODE's.

$$\frac{\partial}{\partial \tau}\mathcal{B}^0 = \mathcal{B}^0(\tau)\tilde{M} + \tilde{M}^\top\mathcal{B}^0(\tau) + \frac{\omega^2 - \omega}{2}\left(A_i - A_j\right)^2 \tag{9.30}$$

$$\begin{aligned} \frac{\partial}{\partial \tau}\mathcal{B}^1 &= \mathcal{B}^1(\tau)\tilde{M} + \tilde{M}^\top\mathcal{B}^1(\tau) + \mathcal{B}^0(\tau)Q^\top R\left(A_i - A_j\right)\omega \\ &\quad + \omega\left(A_i - A_j\right)R^\top Q\mathcal{B}^0(\tau) \end{aligned} \tag{9.31}$$

$$\begin{aligned} \frac{\partial}{\partial \tau}\mathcal{B}^2 &= \mathcal{B}^2(\tau)\tilde{M} + \tilde{M}^\top\mathcal{B}^2(\tau) + \mathcal{B}^1(\tau)Q^\top R\left(A_i - A_j\right)\omega \\ &\quad + \omega\left(A_i - A_j\right)R^\top Q\mathcal{B}^1(\tau) + 2\mathcal{B}^0(\tau)Q^\top Q\mathcal{B}^0(\tau). \end{aligned} \tag{9.32}$$

Let $\gamma := \frac{\omega^2-\omega}{2}$ then these equations admit the following solutions:

$$\begin{aligned}\mathcal{B}^0(\tau) &= \frac{\omega^2-\omega}{2}\int_0^\tau e^{(\tau-u)\tilde{M}^\top}\left(A_i-A_j\right)^2 e^{(\tau-u)\tilde{M}}du\\ &:= \gamma\tilde{\mathcal{B}}^0(\tau), \quad (9.33)\\ \mathcal{B}^1(\tau) &= \gamma\omega\int_0^\tau e^{(\tau-u)\tilde{M}^\top}\left(\tilde{\mathcal{B}}^0(u)Q^\top R\left(A_i-A_j\right)+\left(A_i-A_j\right)R^\top Q\tilde{\mathcal{B}}^0(u)\right)e^{(\tau-u)\tilde{M}}du\\ &:= \gamma\omega\tilde{\mathcal{B}}^1(\tau), \quad (9.34)\\ \mathcal{B}^2(\tau) &= \gamma\omega^2\int_0^\tau e^{(\tau-u)\tilde{M}^\top}\left(\tilde{\mathcal{B}}^1(u)Q^\top R\left(A_i-A_j\right)+\left(A_i-A_j\right)R^\top Q\tilde{\mathcal{B}}^1(u)\right)e^{(\tau-u)\tilde{M}}du\\ &\quad+\gamma^2\int_0^\tau e^{(\tau-u)\tilde{M}^\top}2\tilde{\mathcal{B}}^0(u)Q^\top Q\tilde{\mathcal{B}}^0(u)e^{(\tau-u)\tilde{M}}du\\ &:= \gamma^2\tilde{\mathcal{B}}^{20}(\tau)+\gamma\omega^2\tilde{\mathcal{B}}^{21}(\tau). \quad (9.35)\end{aligned}$$

whereby we implicitly defined the matrices $\tilde{\mathcal{B}}^0(\tau), \tilde{\mathcal{B}}^1(\tau), \tilde{\mathcal{B}}^{20}(\tau), \tilde{\mathcal{B}}^{21}(\tau)$. We can now write the function $\mathcal{B}(\tau)$ as follows:

$$\mathcal{B}(\tau) = \gamma\tilde{\mathcal{B}}^0(\tau)+\alpha\gamma\omega\tilde{\mathcal{B}}^1(\tau)+\alpha^2\gamma^2\tilde{\mathcal{B}}^{20}(\tau)+\alpha^2\gamma\omega^2\tilde{\mathcal{B}}^{21}(\tau). \quad (9.36)$$

A direct substitution of (9.36) into (9.19) allows us to express the function $\mathcal{A}(\tau)$ as:

$$\begin{aligned}\mathcal{A}(\tau) &= \omega\left(r_i-r_j\right)\tau+\gamma Tr\left[\Omega\Omega^\top\int_0^\tau\tilde{\mathcal{B}}^0(u)du\right]+\alpha\gamma\omega Tr\left[\Omega\Omega^\top\int_0^\tau\tilde{\mathcal{B}}^1(u)du\right]\\ &\quad+\alpha^2\gamma^2 Tr\left[\Omega\Omega^\top\int_0^\tau\tilde{\mathcal{B}}^{20}(u)du\right]+\alpha^2\gamma\omega^2 Tr\left[\Omega\Omega^\top\int_0^\tau\tilde{\mathcal{B}}^{21}(u)du\right]\\ &= \omega\left(r_i-r_j\right)\tau+\gamma\tilde{\mathcal{A}}^0(\tau)+\alpha\gamma\omega\tilde{\mathcal{A}}^1(\tau)+\alpha^2\gamma^2\tilde{\mathcal{A}}^{20}(\tau)+\alpha^2\gamma\omega^2\tilde{\mathcal{A}}^{21}(\tau), \quad (9.37)\end{aligned}$$

having again implicitly defined the functions $\tilde{\mathcal{A}}^0(\tau), \tilde{\mathcal{A}}^1(\tau), \tilde{\mathcal{A}}^{20}(\tau), \tilde{\mathcal{A}}^{21}(\tau)$. We consider now the pricing in terms of the Fourier transform, i.e. $\omega = i\lambda$, as in (5.2). Let $\mathcal{Z}$ denote the strip of regularity of the payoff as in Chapter 4. A Taylor-McLaurin expansion w.r.t. $\alpha$ gives the following:

$$\begin{aligned}C(S(t),K,\tau) \approx\; &\frac{e^{-r_i\tau}}{2\pi}\int_{\mathcal{Z}} e^{\mathtt{i}\lambda(r_i-r_j)\tau+\mathtt{i}\lambda x+\gamma\left(\tilde{\mathcal{A}}^0(\tau)+Tr\left[\tilde{\mathcal{B}}^0(\tau)\Sigma\right]\right)}\Phi(\lambda)d\lambda\\ &+\alpha\left(\tilde{\mathcal{A}}^1(\tau)+Tr\left[\tilde{\mathcal{B}}^1(\tau)\Sigma\right]\right)\\ &\frac{e^{-r_i\tau}}{2\pi}\int_{\mathcal{Z}}\gamma\mathtt{i}\lambda e^{\mathtt{i}\lambda(r_i-r_j)\tau+\mathtt{i}\lambda x+\gamma\left(\tilde{\mathcal{A}}^0(\tau)+Tr\left[\tilde{\mathcal{B}}^0(\tau)\Sigma\right]\right)}\hat{F}(\lambda)d\lambda\\ &+\alpha^2\left(\tilde{\mathcal{A}}^{20}(\tau)+Tr\left[\tilde{\mathcal{B}}^{20}(\tau)\Sigma\right]\right)\\ &\times\frac{e^{-r_i\tau}}{2\pi}\int_{\mathcal{Z}}\gamma^2 e^{\mathtt{i}\lambda(r_i-r_j)\tau+\mathtt{i}\lambda x+\gamma\left(\tilde{\mathcal{A}}^0(\tau)+Tr\left[\tilde{\mathcal{B}}^0(\tau)\Sigma\right]\right)}\hat{F}(\lambda)d\lambda\\ &+\alpha^2\left(\tilde{\mathcal{A}}^{21}(\tau)+Tr\left[\tilde{\mathcal{B}}^{21}(\tau)\Sigma\right]\right)\\ &\times\frac{e^{-r_i\tau}}{2\pi}\int_{\mathcal{Z}}\gamma\mathtt{i}\lambda^2 e^{\mathtt{i}\lambda(r_i-r_j)\tau+\mathtt{i}\lambda x+\gamma\left(\tilde{\mathcal{A}}^0(\tau)+Tr\left[\tilde{\mathcal{B}}^0(\tau)\Sigma\right]\right)}\hat{F}(\lambda)d\lambda\\ &+\frac{\alpha^2}{2}\left(\tilde{\mathcal{A}}^1(\tau)+Tr\left[\tilde{\mathcal{B}}^1(\tau)\Sigma\right]\right)^2\\ &\times\frac{e^{-r_i\tau}}{2\pi}\int_{\mathcal{Z}}\gamma^2\mathtt{i}\lambda^2 e^{\mathtt{i}\lambda(r_i-r_j)\tau+\mathtt{i}\lambda x+\gamma\left(\tilde{\mathcal{A}}^0(\tau)+Tr\left[\tilde{\mathcal{B}}^0(\tau)\Sigma\right]\right)}\hat{F}(\lambda)d\lambda.\end{aligned}$$

Recall now from (6.9) the definition of the integrated Black-Scholes variance. In the previous formula in the first term we recognise the Black Scholes price in terms of the characteristic function when the integrated variance is $v = \sigma^2\tau$:

$$C_{B\&S}\left(S(t), K, \sigma, \tau\right) = \frac{e^{-r_i\tau}}{2\pi}\int_{\mathcal{Z}} e^{\mathrm{i}\lambda(r_i-r_j)\tau + i\lambda x + \frac{(i\lambda)^2 - i\lambda}{2}v}\Phi(\lambda)d\lambda, \tag{9.38}$$

so that the price expansion is of the form:

$$\begin{aligned} C(S(t), K, \tau) \approx\; & C_{B\&S}\left(S(t), K, \sigma, \tau\right) + \alpha\left(\tilde{\mathcal{A}}^1(\tau) + Tr\left[\tilde{\mathcal{B}}^1(\tau)\Sigma\right]\right)\partial^2_{xv}C_{B\&S}\left(S(t), K, \sigma, \tau\right) \\ & + \alpha^2\left(\tilde{\mathcal{A}}^{20}(\tau) + Tr\left[\tilde{\mathcal{B}}^{20}(\tau)\Sigma\right]\right)\partial^2_{v^2}C_{B\&S}\left(S(t), K, \sigma, \tau\right) \\ & + \alpha^2\left(\tilde{\mathcal{A}}^{21}(\tau) + Tr\left[\tilde{\mathcal{B}}^{21}(\tau)\Sigma\right]\right)\partial^3_{x^2v}C_{B\&S}\left(S(t), K, \sigma, \tau\right) \\ & + \frac{\alpha^2}{2}\left(\tilde{\mathcal{A}}^1(\tau) + Tr\left[\tilde{\mathcal{B}}^1(\tau)\Sigma\right]\right)^2\partial^4_{x^2v^2}C_{B\&S}\left(S(t), K, \sigma, \tau\right), \end{aligned} \tag{9.39}$$

which completes the proof.

**9.4. Proof of Proposition 5.4.** We follow the procedure in Da Fonseca and Grasselli (2011). We suppose an expansion for the integrated implied variance of the form $v = \sigma^2_{imp}\tau = \zeta_0 + \alpha\zeta_1 + \alpha^2\zeta_2$ and we consider the Black Scholes formula as a function of the integrated implied variance and the log exchange rate $x = \log S$: $C_{B\&S}(S(t), K, \sigma, \tau) = C_{B\&S}(x(t), K, \sigma^2_{imp}\tau, \tau)$. A Taylor-McLaurin expansion gives us the following:

$$\begin{aligned} C_{B\&S}(x(t), K, \sigma^2_{imp}\tau, \tau) =\; & C_{B\&S}(x(t), K, \zeta_0, \tau) + \alpha\zeta_1\partial_v C_{B\&S}(x(t), K, \zeta_0, \tau) \\ & + \frac{\alpha^2}{2}\left(2\zeta_2\partial_v C_{B\&S}(x(t), K, \zeta_0, \tau) + \zeta_1^2\partial^2_{v^2}C_{B\&S}(x(t), K, \zeta_0, \tau)\right). \end{aligned} \tag{9.40}$$

By comparing this with the price expansion (9.39) we deduce that the coefficients must be of the form:

$$\zeta_0 = v_0 \tag{9.41}$$

$$\zeta_1 = \frac{\left(\tilde{\mathcal{A}}^1(\tau) + Tr\left[\tilde{\mathcal{B}}^1(\tau)\Sigma\right]\right)\partial^2_{xv}C_{B\&S}}{\partial_v C_{B\&S}} \tag{9.42}$$

$$\begin{aligned} \zeta_2 =\; & \frac{-\zeta_1^2\partial^2_{\xi^2}C_{B\&S} + 2\left(\tilde{\mathcal{A}}^{20}(\tau) + Tr\left[\tilde{\mathcal{B}}^{20}(\tau)\Sigma\right]\right)\partial^2_{v^2}C_{B\&S}}{2\partial_v C_{B\&S}} \\ & + \frac{2\left(\tilde{\mathcal{A}}^{21}(\tau) + Tr\left[\tilde{\mathcal{B}}^{21}(\tau)\Sigma\right]\right)\partial^3_{x^2v}C_{B\&S} + \left(\tilde{\mathcal{A}}^1(\tau) + Tr\left[\tilde{\mathcal{B}}^1(\tau)\Sigma\right]\right)^2\partial^4_{x^2v^2}C_{B\&S}}{2\partial_v C_{B\&S}}, \end{aligned} \tag{9.43}$$

where the Black Scholes formula $C_{B\&S}(x(t), K, \sigma^2_{imp}\tau, \tau)$ is evaluated at the point $(x, K, v_0, \tau)$. In order to find the values of $\zeta_1, \zeta_2$, we differentiate (6.1)-(6.4) thus obtaining the following ODE's:

$$\frac{\partial}{\partial\tau}\tilde{\mathcal{B}}^0 = \tilde{\mathcal{B}}^0(\tau)\tilde{M} + \tilde{M}^\top\tilde{\mathcal{B}}^0(\tau) + (A_i - A_j)^2, \tag{9.44}$$

$$\frac{\partial}{\partial\tau}\tilde{\mathcal{B}}^1 = \tilde{\mathcal{B}}^1(\tau)\tilde{M} + \tilde{M}^\top\tilde{\mathcal{B}}^1(\tau) + \tilde{\mathcal{B}}^0(\tau)Q^\top R\left(A_i - A_j\right) + \left(A_i - A_j\right)R^\top Q\tilde{\mathcal{B}}^0(\tau), \tag{9.45}$$

$$\frac{\partial}{\partial\tau}\tilde{\mathcal{B}}^{20} = \tilde{\mathcal{B}}^{20}(\tau)\tilde{M} + \tilde{M}^\top\tilde{\mathcal{B}}^{20}(\tau) + 2\tilde{\mathcal{B}}^0(\tau)Q^\top Q\tilde{\mathcal{B}}^0(\tau), \tag{9.46}$$

$$\frac{\partial}{\partial\tau}\tilde{\mathcal{B}}^{21} = \tilde{\mathcal{B}}^{21}(\tau)\tilde{M} + \tilde{M}^\top\tilde{\mathcal{B}}^{21}(\tau+)\tilde{\mathcal{B}}^1(\tau)Q^\top R\left(A_i - A_j\right) + \left(A_i - A_j\right)R^\top Q\tilde{\mathcal{B}}^1(\tau). \tag{9.47}$$

We consider a Taylor-McLaurin expansion in terms of $\tau$

$$\tilde{\mathcal{B}}^0(\tau) \approx (A_i - A_j)^2 \tau + \frac{\tau^2}{2} \left( (A_i - A_j)^2 \tilde{M} + \tilde{M}^\top (A_i - A_j)^2 \right) \tag{9.48}$$

$$\begin{aligned} \tilde{\mathcal{B}}^1(\tau) \approx & \frac{\tau^2}{2} \left[ (A_i - A_j)^2 Q^\top R (A_i - A_j) + (A_i - A_j) R^\top Q (A_i - A_j)^2 \right] \\ & + \frac{\tau^3}{6} \Big( \left[ (A_i - A_j)^2 Q^\top R (A_i - A_j) + (A_i - A_j) R^\top Q (A_i - A_j)^2 \right] \tilde{M} \\ & + \tilde{M}^\top \left[ (A_i - A_j)^2 Q^\top R (A_i - A_j) + (A_i - A_j) R^\top Q (A_i - A_j)^2 \right] \Big) \\ & + \frac{\tau^3}{6} \Big( \left[ (A_i - A_j)^2 \tilde{M} + \tilde{M}^\top (A_i - A_j)^2 \right] Q^\top R (A_i - A_j) \\ & + (A_i - A_j) R^\top Q \left[ (A_i - A_j)^2 \tilde{M} + \tilde{M}^\top (A_i - A_j)^2 \right] \Big) \end{aligned} \tag{9.49}$$

$$\tilde{\mathcal{B}}^{20}(\tau) \approx \frac{\tau^3}{6} 4 (A_i - A_j)^2 Q^\top Q (A_i - A_j)^2 \tag{9.50}$$

$$\begin{aligned} \tilde{\mathcal{B}}^{21}(\tau) \approx & \frac{\tau^3}{6} \Big( \left[ (A_i - A_j)^2 Q^\top R (A_i - A_j) + (A_i - A_j) R^\top Q (A_i - A_j)^2 \right] \\ & \times Q^\top R (A_i - A_j) + (A_i - A_j) R^\top Q \times \\ & \left[ (A_i - A_j)^2 Q^\top R (A_i - A_j) + (A_i - A_j) R^\top Q (A_i - A_j)^2 \right] \Big). \end{aligned} \tag{9.51}$$

Noticing from (6.5)-(6.8) that $\tilde{\mathcal{A}}^i(\tau)$ are one order in $\tau$ higher than the corresponding $\tilde{\mathcal{B}}^i(\tau)$, the following approximations hold:

$$\tilde{\mathcal{A}}^0(\tau) + Tr \left[ \tilde{\mathcal{B}}^0(\tau) \Sigma \right] = Tr \left[ (A_i - A_j)^2 \Sigma \right] \tau + o(\tau) \tag{9.52}$$

$$\tilde{\mathcal{A}}^1(\tau) + Tr \left[ \tilde{\mathcal{B}}^1(\tau) \Sigma \right] = Tr \left[ (A_i - A_j)^2 Q^\top R (A_i - A_j) \Sigma \right] \tau^2 + o(\tau^2) \tag{9.53}$$

$$\tilde{\mathcal{A}}^{20}(\tau) + Tr \left[ \tilde{\mathcal{B}}^{20}(\tau) \Sigma \right] = \frac{2}{3} Tr \left[ (A_i - A_j)^2 Q^\top Q (A_i - A_j)^2 \right] \tau^3 + o(\tau^3) \tag{9.54}$$

$$\begin{aligned} \tilde{\mathcal{A}}^{21}(\tau) + Tr \left[ \tilde{\mathcal{B}}^{21}(\tau) \Sigma \right] = & Tr \Big[ \Big( (A_i - A_j)^2 Q^\top R (A_i - A_j) \\ & + (A_i - A_j) R^\top Q (A_i - A_j)^2 \Big) Q^\top R (A_i - A_j) \Sigma \Big] \frac{\tau^3}{3} + o(\tau^3). \end{aligned} \tag{9.55}$$

We introduce two variables: the log-moneyness $m_f = \log \left( \frac{S^{i,j}(t) e^{(r_i - r_j)\tau}}{K} \right)$ and the variance $V = Tr \left[ (A_i - A_j) \Sigma (A_i - A_j) \right] \tau$ Then, from Lewis (2000), we consider the following ratios among the derivatives of the Black-Scholes formula:

$$\frac{\partial^2_{x,v} C_{B\&S}(x, K, V, \tau)}{\partial_v C_{B\&S}(x, K, V, \tau)} = \frac{1}{2} + \frac{m_f}{V}; \tag{9.56}$$

$$\frac{\partial^2_{v^2} C_{B\&S}(x, K, V, \tau)}{\partial_v C_{B\&S}(x, K, V, \tau)} = \frac{m_f^2}{2V^2} - \frac{1}{2V} - \frac{1}{8}; \tag{9.57}$$

$$\frac{\partial^3_{x^2,v} C_{B\&S}(x, K, V, \tau)}{\partial_v C_{B\&S}(x, K, V, \tau)} = \frac{1}{4} + \frac{m_f - 1}{V} + \frac{m_f^2}{V^2}; \tag{9.58}$$

$$\frac{\partial^4_{x^2,v^2} C_{B\&S}(x, K, V, \tau)}{\partial_v C_{B\&S}(x, K, V, \tau)} = \frac{m_f^4}{2V^4} + \frac{m_f^2 (m_f - 1)}{2V^3}. \tag{9.59}$$

Upon substitution of (9.52)-(9.59) into (9.42)-(9.43) we obtain the values for $\zeta_i, i = 0, 1, 2$ allowing us to express the expansion of the implied volatility.

$$\zeta_0 = Tr\left[(A_i - A_j)\,\Sigma\,(A_i - A_j)\right]\tau, \tag{9.60}$$

$$\zeta_1 = \frac{Tr\left[(A_i - A_j)^2\,Q^\top R\,(A_i - A_j)\,\Sigma\right] m_f \tau}{Tr\left[(A_i - A_j)\,\Sigma\,(A_i - A_j)\right]}, \tag{9.61}$$

$$\begin{aligned}\zeta_3 = {} & \frac{m_f^2}{Tr\left[(A_i - A_j)\,\Sigma\,(A_i - A_j)\right]^2}\tau\left[\frac{1}{3}Tr\left[(A_i - A_j)^2\,Q^\top Q\,(A_i - A_j)^2\,\Sigma\right]\right.\\ & + \frac{1}{3}Tr\Big[\left[(A_i - A_j)^2\,Q^\top R\,(A_i - A_j) + (A_i - A_j)\,R^\top Q\,(A_i - A_j)^2\right]\\ & \left.\times\, Q^\top R\,(A_i - A_j)\,\Sigma\Big] - \frac{5}{4}\frac{Tr\left[(A_i - A_j)^2\,Q^\top R\,(A_i - A_j)\,\Sigma\right]^2}{Tr\left[(A_i - A_j)\,\Sigma\,(A_i - A_j)\right]}\right].\end{aligned} \tag{9.62}$$

By plugging these expressions we obtain the result.

## 10. Images and Tables

| | 6 | 5 | 4 | 3 | 2 |
|---|---|---|---|---|---|
| $\Sigma(t)(1,1)$ | 0.0452 | 0.0422 | 0.0349 | 0.0394 | 0.0370 |
| $\Sigma(t)(1,2)$ | 0.0407 | 0.0400 | 0.0346 | 0.0343 | 0.0295 |
| $\Sigma(t)(2,2)$ | 0.0470 | 0.0480 | 0.0502 | 0.0387 | 0.0343 |
| $A_{us}(1,1)$ | 1.0106 | 0.7901 | 0.9145 | 0.8058 | 0.8760 |
| $A_{us}(1,2)$ | 0.1463 | 0.2991 | 0.3164 | 0.3896 | 0.3320 |
| $A_{us}(2,2)$ | 0.8406 | 0.7763 | 0.8187 | 0.6600 | 0.6796 |
| $A_{eur}(1,1)$ | 1.7852 | 1.7156 | 1.8698 | 1.7785 | 1.7894 |
| $A_{eur}(1,2)$ | -0.1293 | -0.1821 | -0.1576 | -0.1856 | -0.2456 |
| $A_{eur}(2,2)$ | 1.2319 | 1.3205 | 1.2576 | 1.1938 | 1.1811 |
| $M(1,1)$ | -0.3502 | -0.3334 | -0.2514 | -0.3337 | -0.3177 |
| $M(1,2)$ | -0.2927 | -0.4266 | -0.4628 | -0.5200 | -0.4575 |
| $M(2,1)$ | -0.0211 | -0.0303 | 0.0638 | 0.0984 | -0.0656 |
| $M(2,2)$ | -0.3692 | -0.6085 | -0.5257 | -0.8810 | -0.5317 |
| $\beta$ | 1.0145 | 1.0003 | 1.0232 | 1.0141 | 1.0205 |
| $Q(1,1)$ | 0.2893 | 0.3404 | 0.3212 | 0.3573 | 0.3461 |
| $Q(1,2)$ | 0.2169 | 0.2795 | 0.2603 | 0.2682 | 0.2442 |
| $Q(2,1)$ | 0.2410 | 0.2075 | 0.1849 | 0.2308 | 0.2355 |
| $Q(2,2)$ | 0.2752 | 0.3126 | 0.3276 | 0.3466 | 0.3439 |
| $R(1,1)$ | 0.5720 | 0.5939 | 0.5210 | 0.5113 | 0.5116 |
| $R(1,2)$ | -0.1237 | -0.0218 | -0.0266 | 0.0520 | 0.0244 |
| $R(2,1)$ | 0.0059 | -0.1274 | -0.0125 | 0.0550 | 0.0151 |
| $R(2,2)$ | -0.3473 | -0.4203 | -0.4222 | -0.3897 | -0.3775 |
| $A_{jpy}(1,1)$ | 0.9653 | 0.8162 | 0.9316 | 0.9842 | 0.9585 |
| $A_{jpy}(1,2)$ | 0.0565 | 0.0334 | 0.0134 | -0.0475 | -0.0261 |
| $A_{jpy}(2,2)$ | 1.4183 | 1.4213 | 1.3948 | 1.4562 | 1.4369 |

TABLE 1. This table reports the results of the calibration of the Wishart model. For each column, a different number of expiries, ranging from 6 to 2, is chosen.

| | 5 | 4 | 3 | 2 |
|---|---|---|---|---|
| $\Sigma(t)(1,1)$ | -6.7531% | -22.729% | -12.877% | -18.22% |
| $\Sigma(t)(1,2)$ | -1.7384% | -14.944% | -15.598% | -27.397% |
| $\Sigma(t)(2,2)$ | 2.0343% | 6.6436% | -17.714% | -27.153% |
| $A_{us}(1,1)$ | -21.815% | -9.5083% | -20.268% | -13.313% |
| $A_{us}(1,2)$ | 104.52% | 116.31% | 166.39% | 127.03% |
| $A_{us}(2,2)$ | -7.6501% | -2.6065% | -21.49% | -19.152% |
| $A_{eur}(1,1)$ | -3.8997% | 4.7399% | -0.37673% | 0.2308% |
| $A_{eur}(1,2)$ | 40.846% | 21.945% | 43.54% | 89.987% |
| $A_{eur}(2,2)$ | 7.2001% | 2.0892% | -3.0864% | -4.1233% |
| $M(1,1)$ | -4.8057% | -28.198% | -4.7105% | -9.2728% |
| $M(1,2)$ | 45.739% | 58.123% | 77.647% | 56.299% |
| $M(2,1)$ | 43.447% | -402.23% | -566.09% | 210.63% |
| $M(2,2)$ | 64.832% | 42.397% | 138.64% | 44.036% |
| $\beta$ | -1.3909% | 0.8582% | -0.038535% | 0.59754% |
| $Q(1,1)$ | 17.648% | 11.023% | 23.491% | 19.617% |
| $Q(1,2)$ | 28.846% | 20.012% | 23.649% | 12.576% |
| $Q(2,1)$ | -13.922% | -23.302% | -4.2575% | -2.2839% |
| $Q(2,2)$ | 13.59% | 19.051% | 25.951% | 24.965% |
| $R(1,1)$ | 3.8164% | -8.9239% | -10.618% | -10.569% |
| $R(1,2)$ | -82.359% | -78.485% | -142.08% | -119.73% |
| $R(2,1)$ | -2243.8% | -309.96% | 824.57% | 153.99% |
| $R(2,2)$ | 21.031% | 21.558% | 12.204% | 8.6949% |
| $A_{jpy}(1,1)$ | -15.446% | -3.4868% | 1.9618% | -0.69799% |
| $A_{jpy}(1,2)$ | -40.995% | -76.352% | -183.97% | -146.13% |
| $A_{jpy}(2,2)$ | 0.21467% | -1.6532% | 2.673% | 1.3154% |

TABLE 2. In this table we consider the calibration on the largest sample as a basic case. We report the percentage difference between the model parameters resulting from the subsamples

| USD/EUR | | | |
|---|---|---|---|
| | 6m | 9m | 1y |
| 10DC | -0.0005 | 0.0027 | 0.0033 |
| 25DC | 0.0007 | 0.0040 | 0.0065 |
| 0 | 0.0007 | 0.0038 | 0.0069 |
| 25DP | 0.0005 | 0.0029 | 0.0046 |
| 10DP | -0.0001 | 0.0021 | 0.0020 |
| USD/JPY | | | |
| | 6m | 9m | 1y |
| 10DC | 0.0015 | 0.0039 | 0.0060 |
| 15DC | 0.0010 | 0.0038 | 0.0064 |
| 25DC | 0.0003 | 0.0033 | 0.0061 |
| 0 | 0.0000 | 0.0034 | 0.0064 |
| 25DP | 0.0018 | 0.0049 | 0.0070 |
| 15DP | 0.0015 | 0.0047 | 0.0064 |
| 10DP | 0.0013 | 0.0046 | 0.0057 |
| EUR/JPY | | | |
| | 6m | 9m | 1y |
| 10DC | 0.0064 | 0.0132 | 0.0155 |
| 15DC | 0.0068 | 0.0139 | 0.0163 |
| 25DC | 0.0073 | 0.0146 | 0.0182 |
| 0 | 0.0061 | 0.0120 | 0.0154 |
| 25DP | 0.0055 | 0.0087 | 0.0107 |
| 15DP | 0.0026 | 0.0045 | 0.0050 |
| 10DP | -0.0017 | -0.0009 | -0.0009 |

TABLE 3. This table reports the raw difference between the market implied volatility and the implied volatility for 1 year, 9 and 6 months, when we calibrate the model to the previous 3 expiries.

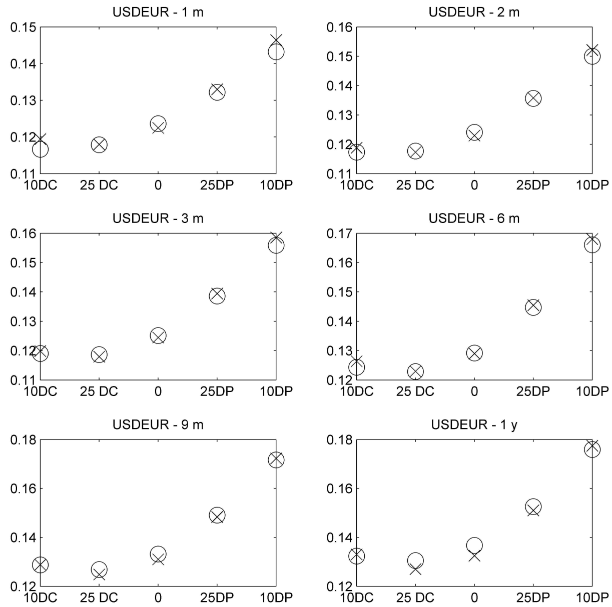



FIGURE 1. Calibration of USD/EUR implied volatility surface.

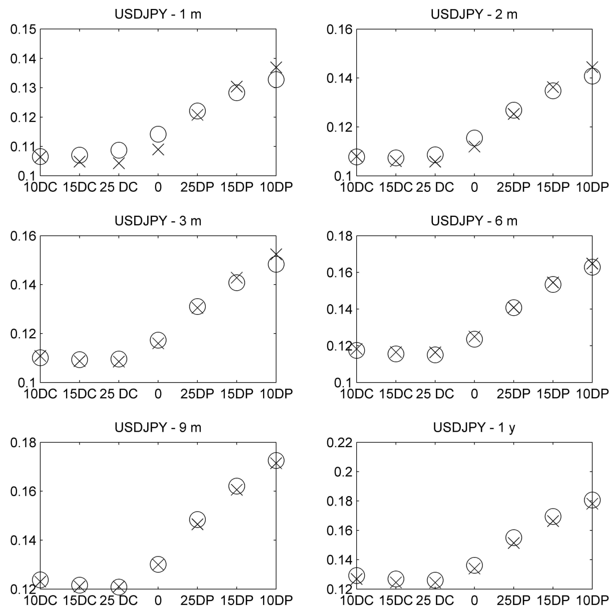



FIGURE 2. Calibration of USD/JPY implied volatility surface.

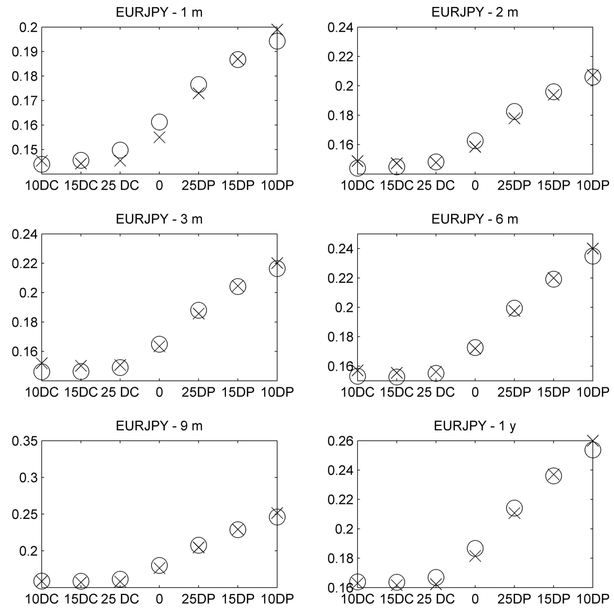



FIGURE 3. Calibration of EUR/JPY implied volatility surface.

| | USD/EUR 9m | USD/EUR 1y | USD/JPY 9m | USD/JPY 1y | EUR/JPY 9m | EUR/JPY 1y |
|---|---|---|---|---|---|---|
| 10DC | 0.0044 | 0.0064 | 0.0030 | 0.0059 | 0.0076 | 0.0091 |
| 15DC | | | 0.0031 | 0.0064 | 0.0085 | 0.0101 |
| 25DC | 0.0062 | 0.0096 | 0.0031 | 0.0064 | 0.0096 | 0.0123 |
| 0 | 0.0053 | 0.0088 | 0.0040 | 0.0076 | 0.0082 | 0.0107 |
| 25DP | 0.0029 | 0.0049 | 0.0059 | 0.0090 | 0.0059 | 0.0071 |
| 15DP | | | 0.0057 | 0.0087 | 0.0019 | 0.0018 |
| 10DP | 0.0007 | 0.0008 | 0.0055 | 0.0083 | -0.0034 | -0.0039 |

TABLE 4. This table reports the raw difference between the market implied volatility and the implied volatility for 1 year and 9 months, when we calibrate the model to the previous 4 expiries.

# Bibliography


A. Ahdida and A. Alfonsi. Exact and high order discretization schemes for Wishart processes and their affine extensions. *ArXiv e-prints*, 2010.

J. Andersen and L. Andreasen. Volatile volatilities. *Risk*, 12:163–168, 2002.

L. B. Andersen and R. Brotherton-Ratcliffe. Extended Libor Market Models with Stochastic Volatility. *Working Paper Gen Re Securities*, 2001. doi: 10.2139/ssrn.294853.

L. B. G. Andersen and V. V. Piterbarg. Moment explosions in stochastic volatility models. *Finance and Stochastics*, 11:29–50, 2007.

B.D.O. Anderson and J.B. Moore. *Linear optimal control*. Prentice-Hall, first edition, 1971.

G.S. Bakshi and D. Madan. Spanning and Derivative-Security Valuation. *Journal of Financial Economics*, 55:205–238, 2000.

O. E. Barndorff-Nielsen and R. Stelzer. Multivariate supOU processes. *ArXiv e-prints*, 2010.

O. E. Barndorff-Nielsen and R Stelzer. Positive-definite matrix processes of finite variation. *to appear in Probab. Math. Statist.*, 27(1), 2007.

R. Beneder and M. Elkenbracht-Huizing. Foreign exchange options and the volatility smile. *Medium Econometrische Toepassingen*, 2:30–36, 2003.

T. Björk. *Arbitrage theory in continuous time*. Oxford university press, Oxford, second edition, 2004.

F. Black. The pricing of commodity contracts. *Journal of Financial Economics*, 3(1-2):167–179, 1976. URL `http://ideas.repec.org/a/eee/jfinec/v3y1976i1-2p167-179.html`.

F. Black and M. Scholes. The Pricing of Option and Corporate Liabilities. *Journal of Political Economy*, (81):637–654, May 1973.

A. Brace, D. Gatarek, and M. Musiela. The market model of interest rate dynamics. *Mathematical Finance*, 7(2):127–155, 1997. URL `http://econpapers.repec.org/RePEc:bla:mathfi:v:7:y:1997:i:2:p:127-155`.

D.T. Breeden and R.H. Litzenberger. Prices of State Contingent Claims Implicit in Options Prices. *Journal of Business*, (51(4)):621–651, May 1978.

D. Brigo and F. Mercurio. Analytical Pricing of the Smile in a Forward LIBOR Market Model. *Quantitative Finance*, 3(1):15–27, 2003.

D. Brigo and F. Mercurio. Interest Rate Models: Theory and Practice. *Springer Finance, Heidelberg, 2nd edition*, 2006.

R. Brockett. *Finite Dimensional Linear Systems*. Wiley, New York, 1970.

M. F. Bru. Wishart processes. *Journal of Theoretical Probability*, 4:725–751, 1991. ISSN 0894-9840. URL `http://dx.doi.org/10.1007/BF01259552`.

A. Buraschi, A. Cieslak, and F. Trojani. Correlation risk and the term structure of interest rates. 2008.

A. Buraschi, P. Porchia, and F. Trojani. Correlation risk and optimal portfolio choice. *The Journal of Finance*, 65(1):393–420, 2010. ISSN 1540-6261. doi: 10.1111/j.1540-6261.2009.01533.x. URL `http://dx.doi.org/10.1111/j.1540-6261.2009.01533.x`.

P. Carr and D. B. Madan. Option valuation using the fast fourier transform. *Journal of Computational Finance*, 2:61–73, 1999.

A. Castagna. *FX Options and Smile Risk*. Wiley, 2010.

C. Chiarella, C. Hsiao, and T. To. Risk premia and wishart term structure models. *Working Paper*, 2010.

P. Christoffersen, S. L. Heston, and K. Jacobs. The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work so Well. *Management Science*, 72:1914–1932, 2009.

I. Clark. *Foreign Exchange Option Pricing: A Practitioner's Guide*. Wiley, 2011.

P. Collin-Dufresne and R. Goldstein. Pricing swaptions within an affine framework. *Journal of Derivatives*, 3(1-2):167–179, 2002. URL `http://ideas.repec.org/a/eee/jfinec/v3y1976i1-2p167-179.html`.

G. M. Constantinides. A theory of the nominal term structure of interest rates. *Review of Financial Studies*, 5(4):531–552, 1992.

John Cox, Jonathan Ingersoll, and Stephen A. Ross. A Theory of the Term Structure of Interest Rates. *Econometrica*, 53:385–407, 1985.

C. Cuchiero, D. Filipović, E. Mayerhofer, and J. Teichmann. Affine Processes on Positive Semidefinite Matrices. *Annals of Applied Probability, to appear*, 2009.

J. Da Fonseca. personal communication, 2011.

J. Da Fonseca and M. Grasselli. Riding on the smiles. *to appear in Quantitative Finance*, 2011.

J. Da Fonseca, M. Grasselli, and F. Ielpo. Estimating the Wishart Affine Stochastic Correlation Model Using the Empirical Characteristic Function. *SSRN eLibrary*, 2007a.

J. Da Fonseca, M. Grasselli, and C. Tebaldi. Option pricing when correlations are stochastic: an analytical framework. *Review of Derivatives Research*, 10(2):151–180, 2007b. URL `http://econpapers.repec.org/RePEc:kap:revdev:v:10:y:2007:i:2:p:151-180`.

J. Da Fonseca, M. Grasselli, and C. Tebaldi. A multifactor volatility heston model. *Quantitative Finance*, 8(6):591–604, 2008. URL `http://econpapers.repec.org/RePEc:taf:quantf:v:8:y:2008:i:6:p:591-604`.

J. Da Fonseca, M. Grasselli, and F. Ielpo. Hedging (Co)Variance Risk with Variance Swaps. *International Journal of Theoretical and Applied Finance, to appear*, 2009.

T. Damm. Applied linear algebra. unpublished lecture notes, Padova 2009, 2009.

S. Del Baño Rollin. Spot inversion in the heston model. *Unpublished manuscript, available at http://www.crm.es/Publications/08/Pr837.pdf*, 2008.

C. Donati-Martin, Y. Doumerc, H. Matsumoto, and M. Yor. Some properties of Wishart process and a matrix extension of the Hartman–Watson law. *Publ. Res. Inst. Math. Sci.*, (40):1385–1412, 2004.

D. Duffie and R. Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6(4):379–406, 1996. URL `http://ideas.repec.org/a/bla/mathfi/v6y1996i4p379-406.html`.

D. Duffie, J. Pan, and K. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68:1343–1376, 2000.

D. Duffie, D. Filipović, and W. Schachermayer. Affine processes and applications in finance. *Annals of Applied Probability*, 13:984–1053, 2003.

E. Eberlein and F. Özkan. The lévy libor model. *Finance and Stochastics*, 9(3):327–348, 07 2005. URL `http://ideas.repec.org/a/spr/finsto/v9y2005i3p327-348.html`.

P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling extremal events for insurance and finance*. Applications of mathematics. Springer, 1997. ISBN 9783540609315.

M. L. Esquível, C. Veiga, and E. Wystup. Unifying exotic option closed formulas. *Review of Derivatives Research, forthcoming*, 2010.

J. Faraut and A. Koràny. *Analysis on Symmetric Cones*. Oxford Mathematical Monographs. Oxford University Press, first edition, 1994.

W. Feller. Two Singular Diffusion Problems. *Annals of Mathematics*, 54:173–182, 1951.

M. B. Garman and S. W. Kohlhagen. Foreign currency option values. *Journal of International Money and Finance*, 2(3):231 – 237, 1983. ISSN 0261-5606. doi: DOI:10.1016/S0261-5606(83)80001-1. URL `http://www.sciencedirect.com/science/article/pii/S0261560683800011`.

H. Geman, Elkaroui N., and Rochet J. Changes of numeraire, changes of probability measure and option pricing. *Journal of Applied Probability*, 32:443–458, 1995.

K. Glasserman and K. Kim. Moment explosions and stationary distributions in affine diffusion models. *forthcoming in Mathematical Finance*, 2011.

P. Glasserman and S. Kou. The term structure of simple forward rates with jump risk. *Mathematics Finance*, 13:383–410, 2003.

C. Gourieroux. Continuous time wishart process for stochastic risk. *Econometric Reviews*, 25:177–217, 2006. ISSN 0747-4938.

C. Gourieroux and R. Sufana. Wishart Quadratic Term Structure Models. *SSRN eLibrary*, 2003.

C. Gourieroux and R. Sufana. Derivative Pricing With Wishart Multivariate Stochastic Volatility. *Journal of Business and Economic Statistics*, 2004(October):1–44, 2005.

C. Gourieroux, A. Monfort, and R. Sufana. International Money and Stock Market Contingent Claims. Working Papers 2005-41, Centre de Recherche en Economie et Statistique, 2005. URL `http://econpapers.repec.org/RePEc:crs:wpaper:2005-41`.

C. Gouriéroux and R. Sufana. Discrete time Wishart term structure models. *Journal of Economic Dynamics and Control*, 35:815–824, 2011.

M. Grasselli and C. Tebaldi. Solvable Affine Term Structure Models. *Mathematical Finance*, 18:135–153, 2008.

A. Gupta and D. Nagar. *Matrix Variate Distributions*, volume 104 of *Monographs and surveys in pure and applied mathematics*. Chapman and Hall CRC, Boca Raton, first edition, 2000.

D. Heath and E. Platen. Currency derivatives under a minimal market model. *The ICFAI Journal of Derivatives Marktes*, 3:68–86, 2006a.

D. Heath and E. Platen. *A Benchmark Approach to Quantitative Finance*. Springer-Verlag, 2006b.

L. S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:327–343, 1993.

F. Jamshidian. Pricing of Contingent Claims in the One Factor Term Structure Model. 1987.

F. Jamshidian. An exact bond option formula. *Journal of Finance*, 44(1):205–209, 1989.

F. Jamshidian. Libor and swap market models and measures. *Finance and Stochastics*, 1(4):293–330, 1997. URL `http://ideas.repec.org/a/spr/finsto/v1y1997i4p293-330.html`.

J. Jong and X. Zhou. *Stochastic controls: Hamiltonian systems and HJB equations*. Springer, New York, 1st edition, 1999.

M. Joshi and R. Rebonato. A stochastic volatility displaced-diffusion extension of the libor market model. *Quantitative Finance*, 3(6):458–469, December 2003.

I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, Berlin, 2nd edition, 1991.

M. Keller-Ressel. Moment explosions and long-term behavior of affine stochastic volatility models. *Mathematical Finance*, 21:73–98, 2011.

M. Keller-Ressel. *Affine processes - theory and applications in finance*. PhD thesis, Vienna Institute of Technology, 2008.

M. Keller-Ressel and T. Steiner. Yield curve shapes and the asymptotic short rate distribution in affine one-factor models. *Finance and Stochastics*, 12(2):149–172, 2008. URL `http://econpapers.repec.org/RePEc:spr:finsto:v:12:y:2008:i:2:p:149-172`.

M. Keller-Ressel, A. Papapantoleon, and J. Teichmann. The affine LIBOR models. *ArXiv e-prints*, April 2009.

M. Leippold and F. Trojani. Asset Pricing with Matrix Jump Diffusions. *SSRN eLibrary*, 2010.

J. J. Levin. On the matrix Riccati equation. *Proc. Amer. Math. Soc.*, 10:519–524, 1959.

A.L. Lewis. *Option valuation under stochastic volatility: with Mathematica code*. Finance Press, 2000. ISBN 9780967637204.

E. Mayerhofer. Wishart processes and wishart distributions: An affine processes point of view. *Arxiv*, 2012.

E. Mayerhofer. On the existence of non-central Wishart distributions. *ArXiv e-prints*, September 2010.

E. Mayerhofer, O. Pfaffel, and R. Stelzer. On strong solutions for positive definite jump-diffusions. *ArXiv e-prints*, October 2009.

A.J. McNeil, R. Frey, and P. Embrechts. *Quantitative risk management: concepts, techniques and tools*. Princeton series in finance. Princeton University Press, 2005. ISBN 9780691122557.

A. Mijatović and M. Urusov. On the martingale property of certain local martingales. *to appear in Probability Theory and Related Fields*, 2011.

K. R. Miltersen, K. Sandmann, and D. Sondermann. Closed form solutions for term structure derivatives with log-normal interest rates. *Journal of Finance*, 52(1):409–30, March 1997. URL `http://ideas.repec.org/a/bla/jfinan/v52y1997i1p409-30.html`.

J. Muhle-Karbe, O. Pfaffel, and R. Stelzer. Option pricing in multivariate stochastic volatility models of OU type. *ArXiv e-prints*, January 2010.

M. Musiela and M. Rutkowski. Continuous-time term structure models: Forward measure approach. *Finance and Stochastics*, 1(4):261–291, 1997. URL `http://ideas.repec.org/a/spr/finsto/v1y1997i4p261-291.html`.

C. Pigorsch and R. Stelzer. On the definition, stationary distribution and second order structure of positive semidefinite Ornstein–Uhlenbeck type processes. *ArXiv e-prints*, September 2009.

V. V. Piterbarg. A stochastic volatility model with time dependent skew. *Applied Mathematical Finance*, 12(2):147–185, 2005a.

V. V. Piterbarg. Time to smile. *Risk*, 18(5):71–75, 2005b.

J. Pitman and M. Yor. A decomposition of bessel bridges. *Probability Theory and Related Fields*, 59: 425–457, 1982. doi: 10.1007/BF00532802.

Dong Qu. Pricing Basket Options with Skew. *Willmott Magazine*, pages 58–64, 2010.

A. Quarteroni, R. Sacco, and F. Saleri. *Numerical mathematics*, volume 37 of *Texts in applied mathematics*. Springer, first edition, 2000.

D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1994. ISBN 3-540-57622-3.

A. Sepp. Fourier Transform for Option Pricing under Affine Jump- Diffusions: An Overview. *Unpublished Manuscript, available at www.hot.ee/seppar.*, 2003.

K. J. Singleton and L. Umantsev. Pricing Coupon-Bond Options and Swaptions in Affine Term Structure Models. *Mathematical Finance*, 12:427–446, 2002.

H. Wimmer. Applied linear algebra - matrix equations. unpublished lecture notes, Padova 2009, 2009.

B. Wong and C. C. Heyde. On the martingale property of stochastic exponentials. *Journal of Applied Probability*, 41:654–664, 2004.

L. Wu and F. Zhang. Libor market model with stochastic volatility. *Journal of industrial and management optimization*, 2(2):199–227, May 2006.

U. Wystup. *FX Options and Structured Products*. Wiley, 2006.

U. Wystup and D. Reiswich. A guide to fx options quoting conventions. *The Journal of Derivatives*, 18 (2):58–68, 2010.

F. Zhang. *Matrix theory: basic results and techniques*. Springer, first edition, 1999.