Convexity Corrections via a Markov-functional approach

Urvashi Behary Paray

Thesis submitted for the degree of

Doctor of Philosophy in Statistics (Research)

University of Warwick
Department of Statistics

March 2023

Contents

List of Figures

Acknowledgements

Firstly, I would like to express my deepest gratitude to my supervisor Jo Kennedy for her guidance and patience throughout the years. I feel fortunate to have had the chance to work with her, a journey that started in 2017, when she supervised my master's thesis. I would like to thank her for introducing me to interest rate modelling, and also for her unwavering support and encouragement throughout.

Secondly, I would like to wholeheartedly thank the people who have supported me in this experience. I humbly apologise to whomever I may unintentionally forget in the following. I would like to thank Sam Forbes for keeping me sane during lockdown. I enjoyed the innumerable conversations (both deep and silly) we shared over our coffee breaks and walks around Coventry and Kenilworth. During my undergraduate time at Warwick, I was extremely lucky to have crossed paths with some incredibly smart and humble people whose friendship I cherish: Nazem Khan, Justin Gwee and Ruiqi Liu. My sincere gratitude goes to them for enduring me throughout the years, supporting me and making my PhD journey much more enjoyable. I would like to extend my gratitude to Nikolaos Constantinou for his kindness, support and the discipline he instilled in me at a time when I needed it the most. I am deeply indebted to Iris Li, for always being there for me and for believing in me. I am also thankful for the Badminton Rock-up community at Warwick, and special thanks goes to Omstavan Samant, for the fun games and encouragement. Lastly, I would like to thank the Stochastic Finance Reading Group at Warwick, for broadening my knowledge of the field.

Finally, I would like to thank my parents for their incredible support and love. To my brother, I express my heartfelt gratitude, for letting me steal his (bigger) room the moment he left home for his own studies. He has been an inspiration ever since, and his unfaltering encouragement throughout the years away from home brought us here. Special thanks goes to my sister too for her unconditional support. It goes without saying, I would not be where I am without them.

Declaration

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree. The work is original except where indicated by specific references in the thesis. The material in this thesis has not been submitted to any other publication.

Place: University of Warwick
Date: 28 March 2023

Urvashi Behary Paray

Abstract

In this thesis, we study convexity corrections in the context of pricing exotic European products whose payoff is a function of some forward swap rate, with Constant Maturity Swaps (CMS) being a relevant example we will focus on. To do so, we proceed in two stages: firstly, we develop and implement a Markov-functional model at a single time (single-time MFM) that is calibrated to an appropriate set of market-implied distribution of the swap rates in their own swaption measures. Knowledge about these marginal distributions can be recovered from actively-traded swaption prices. We want to incorporate this information into the pricing model and study the problem, specifically allowing for long payment dates. We use this model to analyse the properties of the (joint) distribution of the forward swap rates that matter when pricing a CMS. The single-time MFM is however too computationally expensive to be used as a practical pricing model, but it provides a flexible framework through which convexity corrections can be studied. Using the insights obtained from the single-time MFM, we move on to the second stage of the thesis: the development of a fast, efficient pricing model, which we label as the 'MF-Lite model' that is viable in practice, and takes into account the market features of the joint distributions of the swap rates that matter most when pricing convexity-related products. We develop the model first in a one-factor context, starting the setup under the swaption measure. We observe that the model is numerically close to the one-factor single-time MFM. We then reconfigure the approach to allow for a second factor. We propose a two factor MF-Lite model set up in the forward measure.

Glossary of Notation and Abbreviations

Notation	**Meaning**
$(N, \mathbb{N})$	general numéraire pair
D_{tT}	Value at time t of a unit of currency paid at time T
S_n/S_m	General single maturity
$T \equiv S_0 < S_1 < ... < S_n < \infty$	Tenor structure; an increasing sequence of maturities
α_i, for $i \in \{1, \ldots, n\}$	$S_i - S_{i-1}$; Time elapsed between two consecutive maturities
y_T^n	Forward swap rate with setting date T and maturity S_n
$\mathbb{N}^m$	Forward measure corresponding to taking $D_{.S_m}$ as numéraire
$\mathbb{S}^n$	Swaption measure with respect to which y^n is a martingale
$\mathbb{F}$	Forward measure corresponding to taking $D_{.T}$ as numéraire
K	Strike
$\mathbf{x}_T := (x_T^{(1)}, x_T^{(2)})$	Two-dimensional model driver
$\lambda_k, k \in \{1, 2\}$	Variance of $x_T^{(k)}$
$\hat{y}_T^n(\mathbf{x}_T)$	$:= \beta_n^{(1)} x_T^{(1)} + \beta_n^{(2)} x_T^{(2)}$
f^n	Monotonic increasing function (Markov-functional sweep constructed to calibrate the single-time MFM to the market-implied marginal distribution of y_T^n under $\mathbb{S}^n$
F_n^y	(Market-implied) Marginal distribution of y_T^n under $\mathbb{S}^n$
$V_0^n(K)$	Payer swaption price, with underlying rate y_T^n and strike K
$\bar{V}_0^n(K)$	Digital payer swaption price, with underlying rate y_T^n and strike K
$\tilde{V}_0^n(K)$	$\alpha_n V_0^n(K) + (1 + \alpha_n K)\bar{V}_0^n(K)$

$\mathcal{C}_{n,m}$	$:= \mathbb{E}_{\mathbb{N}^m}[F(y_T^n)] - \mathbb{E}_{\mathbb{S}^n}[F(y_T^n)]$; Convexity correction from pay-off $F : \mathbb{R} \mapsto \mathbb{R}$, based on reference swap rate y_T^n paid at time S_m
V_0^{CMS}	Value at time zero of a single payment of a CMS
a_{k1}	$:= \frac{1}{\sqrt{\tilde{M}}}$, for all $k \in \{1, \ldots, \tilde{M}\}$, $\tilde{M}$ denotes the maximum between the reference swap rate maturity index and the payment date index
a_{k2}	$a + b\exp(\lambda k)$, $a, b, \lambda \in \mathbb{R}$, $k \in \{1, \ldots, \tilde{M}\}$
η_n	One-dimensional model driver
X_n	Standard Normal random variable under $\mathbb{S}^n$
$g_n(X_n)$	$:= (F_n^y)^{-1}(\Phi(X_n))$; expressing y_T^n as a function of X_n
f_n	Postulated functional form for y_T^n
q_n	$:= (f_n)^{-1} \circ g_n$; $\eta_n(X_n) = q_n(X_n)$
$\hat{P}_T^n$	Postulated prior functional form for the PVBP P_T^n
a_n	Parameter that fixes the no-arbitrage condition; we have $P_T^n := a_n \hat{P}_T^n$
$\bar{F}_{\eta_n}$	Distribution function of η_n under $\mathbb{S}^n$
F_{η_n}	Distribution function of η_n under $\mathbb{F}$
$S_{n_k}, k \in \{1, 2\}$	The two general maturities for which we do the partial model setup
$\eta_{n_k}, k \in \{1, 2\}$	Corresponding model drivers for each maturity
Y	Standard Normal random variable under $\mathbb{F}$
$h_{n_k}(Y), k \in \{1, 2\}$	$:= (F_{\eta_k})^{-1}(\Phi(Y))$; expressing η_{n_k} as a function of Y
z_T^i, $i \in \{1, \ldots, n\}$	$\frac{\beta_i^{(1)} x_T^{(1)} + \beta_i^{(2)} x_T^{(2)}}{\sqrt{(\beta_i^{(1)})^2 \lambda_1 + (\beta_i^{(2)})^2 \lambda_2}} \sim \mathcal{N}(0, 1)$ under $\mathbb{F}$, $\beta_i^{(1)}, \beta_i^{(2)} \in \mathbb{R}$
$\hat{f}_i$, $i \in \{1, \ldots, n\}$	Postulated functional form for the prior swap rates used to construct the PVBP
$\hat{y}_T^i(z_T^i)$	$:= \hat{f}_i(z_T^i)$
$\hat{y}_T^{n:i}(z_T^i)$	$:= \hat{f}_i(h_n(z_T^i))$ (Refined prior model for the swap rates)
$(\tilde{x}_T^{(1)}, \tilde{x}_T^{(2)})$	Uncorrelated bivariate Gaussian random variable under $\mathbb{S}^n$
$\tilde{z}_T^i$	$\frac{\beta_i^{(1)} \tilde{x}_T^{(1)} + \beta_i^{(2)} \tilde{x}_T^{(2)}}{\sqrt{(\beta_i^{(1)})^2 \lambda_1 + (\beta_i^{(2)})^2 \lambda_2}} \sim \mathcal{N}(0, 1)$ under $\mathbb{S}^n$, $\beta_i^{(1)}, \beta_i^{(2)} \in \mathbb{R}$
$\tilde{y}_T^{n:i}(\tilde{z}_T^i)$	$\hat{f}_i(q_n(\tilde{z}_T^i))$
$\tilde{a}_n \tilde{P}_T^n$	Postulated form for the PVBP constructed using $\tilde{y}_T^{n:i}$ and $\tilde{a}_n \in \mathbb{R}$

Abbreviation	**Meaning**
LIBOR	London Inter-bank Offer Rate
LMM	LIBOR Market Model
LIA	LIBOR-in-arrears
CMS	Constant Maturity Swap
PDB	Pure discount bond
PVBP	Present value of basis point
MFM	Markov-functional model
1F/(2F naive) smcMFL	One-factor/(two-factor naive) swaption measure calibrated Markov-functional Lite
2F (refined) fmcMFL	Two-factor (refined) forward measure calibrated Markov-functional Lite

Table 1: Abbreviations

CHAPTER 1

Introduction

Recent statistics ("Bank for International Settlements" (2022)) indicate that in the global over-the-counter derivatives market, with a notional amount of 632 billion USD traded, interest rate contracts accounted for nearly 80% of the trades, with a significant contribution coming from swaps. Interest-rate markets undoubtedly form a large and important part of the global financial industry. Interest rate swaps are actively traded in the market, with prices quoted by market participants. More complex derivatives, tailor-made to suit the needs of clients are equally traded, but not on the market. This category of derivatives is referred to as *exotic* derivatives. There is no publicly available information about the prices of these products, so they need to be priced in-house. The pricing of exotic derivatives has attracted significant attention, from academics and practitioners alike and remains to this day, a challenging and complex problem in finance. They lack liquidity, hence trading them relies on quoting a price based on a pricing model. In this thesis, we aim to set up a pricing model for such products. We focus mainly on *European* products whose pricing require some form of *convexity correction*. We explain both concepts below. We provide in the preliminaries section at the end of this chapter some basic probability tools and an overview of interest-rate terms that we use in the discussion below and throughout the thesis. Readers unfamiliar with the topic can read through the preliminaries before carrying on.

Consider a product with a payoff $F(y_T^1, \ldots, y_T^n)$, set at time T and paid at time $S \geq T$, with $(y_T^1, \ldots, y_T^n)$ being a set of underlying forward rates (could be LIBOR or swap rates for example). This is a very general description of the payoff of European-type products, the characteristic feature of which is that we can observe the market at a single time T in order to determine the payments. In practice, n is often taken to be 1 or 2. In order to price this payoff today, we take an expectation of $F(y_T^1, \ldots, y_T^n)$. The problem arises when we have the expectation with respect to a measure under which the forward rates $y^1, \ldots, y^n$ are not martingales, hence some adjustments (convexity corrections) have to be made. We shall discuss some market products that fall into this category shortly. We now define convexity corrections.

1.1 Convexity Corrections

The need for convexity corrections arises when taking an expectation of an interest rate with respect to a measure under which it is not a martingale. Pelsser (2000) defined convexity correction as the adjustment that has to be made when certain interest rates are paid at the *'wrong'* time (or in the *'wrong'* currency) and can be thought of as the side-effect of a change of numéraire.

Suppose we are given a forward interest rate which is set at time T, denoted by y_T (it could be LIBOR or a swap rate). Suppose we have a contract whereby the payoff is given by $F(y_T)$ and is paid at some future time $S > T$. The value at time T of this contract will be:

$$V_T = F(y_T)D_{TS},$$

where D_{TS} denotes the value at T of a pure discount bond paying unity at S. Let us consider the simplest form of payoff function, where F is just the identity function. If we take y_T to be a forward swap rate, we will be valuing a single payment of a Constant Maturity Swap (CMS), an exotic European-type derivative that will be the main derivative we study in this thesis. We will elaborate further on CMS in the next section. In accordance to Theorem 1.5.1, which we review in the later Section 1.5, the value at time zero of this payoff is given by:

$$V_0 = N_0 \mathbb{E}_{\mathbb{N}} \left[\frac{y_T D_{TS}}{N_T} \right],$$

for a given numéraire pair $(N, \mathbb{N})$.

Remark 1: For any swap rate or LIBOR, there is a natural numéraire $\hat{N}$ and corresponding equivalent martingale measure (EMM) $\hat{\mathbb{N}}$ under which the forward rate is a martingale. We refer ahead to definition 1.5.2 where we introduced the swaption measure as the martingale measure for forward swap rates.

A first naive approach in valuing the payoff is to work in the EMM corresponding to $D_{\cdot S}$ as numéraire and disregard the fact that the forward rate is not a martingale under this measure, and to have:

$$V_0^{naive} = y_0 D_{0S}.$$

The error is then given by:

$$\begin{aligned} V_0 - V_0^{naive} &= \hat{N}_0 \mathbb{E}_{\hat{\mathbb{N}}}\left[\frac{y_T D_{TS}}{\hat{N}_T}\right] - y_0 D_{0S} \\ &= \hat{N}_0 \mathbb{E}_{\hat{\mathbb{N}}}\left[\frac{y_T D_{TS}}{\hat{N}_T}\right] - \hat{N}_0 \mathbb{E}_{\hat{\mathbb{N}}}[y_T] \mathbb{E}_{\hat{\mathbb{N}}}\left[\frac{D_{TS}}{\hat{N}_T}\right] \\ &= \hat{N}_0 \mathrm{cov}_{\hat{\mathbb{N}}}\left(y_T, \frac{D_{TS}}{\hat{N}_T}\right). \end{aligned}$$

Consider a one-factor model. In this case, one can write $\frac{D_{TS}}{\hat{N}_T}$ as a function of y_T. The error from using the naive approach arises because the term $\frac{D_{TS}}{\hat{N}_T}$ depends on y_T. Note that $\frac{D_{TS}}{\hat{N}_T}$ as a function of y_T may in fact be convex, concave, or neither, so the error may be positive or negative.

Suppose we now specialise to the case where $y_T = L_T := L_T[T, S]$, a LIBOR (see equation (1.3)). We define $\alpha := S - T$. In this case, the natural numéraire $\hat{N}$ is the bond maturing at S. When we take the payment date $S = T$, we obtain a LIBOR-in-arrears payment and the convexity correction in this case is:

$$\begin{aligned} V_0 - V_0^{naive} &= D_{0S} \mathbb{E}_{\hat{\mathbb{N}}}\left[L_T \frac{D_{TT}}{D_{TS}}\right] - D_{0S} \mathbb{E}_{\hat{\mathbb{N}}}[L_T] \mathbb{E}_{\hat{\mathbb{N}}}\left[\frac{D_{TT}}{D_{TS}}\right] \\ &= D_{0S} \mathbb{E}_{\hat{\mathbb{N}}}[L_T(1 + \alpha L_T)] - D_{0S} \mathbb{E}_{\hat{\mathbb{N}}}[L_T] \mathbb{E}_{\hat{\mathbb{N}}}[1 + \alpha L_T] \\ &= \alpha D_{0S} \big(\mathbb{E}_{\hat{\mathbb{N}}}[L_T^2] - (\mathbb{E}_{\hat{\mathbb{N}}}[L_T])^2\big). \end{aligned}$$

In this case, L_T^2 is convex in L_T, and by *Jensen's inequality*, the error is positive. This gives rise to the term 'convexity correction'.

1.2 Market Products

Past experience has taught the banking industry of the need to address the issue of convexity corrections appropriately and motivates the approach taken in this thesis. We mention three products here: the LIBOR-in arrears swap, the Constant Maturity Swap, and the TARN. We have already encountered the first two in the section above.

The LIBOR-in-arrears swap (LIA swap) is an instrument which has a payoff based on LIBOR as described above. As late as mid-1990's, these swaps were priced using the naive valuation, so not building in the optionality at all. Banks started to become aware of industry players exploiting the mispricing and seizing the arbitrage opportunities it created. This marked the beginning of the banking industry taking the problem of convexity corrections seriously.

A second product that drew the attention of the City in the late 1990s/early 2000s was the Constant Maturity Swaps (CMS). CMS derivatives are popular financial instruments that enable investors to take a long term view on the level (or change in case of CMS Spread derivatives) of the swap yield curve and hedge their exposure. A CMS is an interest rate swap where the interest in one leg is reset periodically with reference to a market swap rate, usually with a long-term maturity that could go up to 30 years. For the other leg of the swap, payments are either a short-term rate such as a 6-month LIBOR paid semi-annually, or a fixed rate. The standard swaptions traded between professional counterparties in the European markets are cash-settled swaptions. The payoff of a physically-settled swaption based on a forward swap rate y_T is of the form $P_T(y_T - K)_+$ (where y_T is as defined in equation (1.6) and P_T is the associated PVBP(present value of a basis point). For a cash-settled swaption, it is of the form $h(y_T)(y_T - K)_+$, where the term $h(y_T)$ is a function solely of the swap rate y_T chosen to mimic the more complicated term P_T (specifically, we have that $h(y_T) := \sum_{i=1}^{n} \frac{\alpha}{(1+\alpha y_T)^i} = \frac{1}{y_T}\left(1 - \frac{1}{(1+\alpha y_T)^n}\right)$, where n is the number of payments in the reference swap y_T and α is its accrual factor). When the payment date of a CMS cashflow is the start date of the reference swap, $S = T$, the liquid prices of cash-settled swaptions completely determine the market distribution of the swap rate y_T under the forward measure corresponding to taking the pure discount bond maturing at time T as numéraire. As a result, the CMS payment can be perfectly statically replicated with a portfolio of cash-settled swaptions. In the early 2000s, trades started to be placed on CMS swaps which exploited the fact that most pricing models for them in use at the time did not incorporate the correct market distribution for y_T (i.e, the correct skew in the volatility smile).

The third and last product we draw the reader's attention to here is the TARN (Targeted Accrual Redemption Note). In this trade, an investor hands over the notional amount. In return, she receives coupons. If the total amount received in coupons ever reaches a pre-specified maximum amount, say A, then the trade terminates; the current coupon is paid in part, to bring the total amount received in coupons up to the maximum level A, the notional amount is returned to the investor and the trade terminates. This is the early termination case. On the other hand, if the note reaches its full maturity before the maximum amount A is attained, all the coupons will have been paid in full and the notional amount is returned at the natural trade end date. In some versions of the TARN (relevant to us), in this natural maturity case, the final coupon is increased so that the coupon payments total A. One can see that in this case, the first coupon for example, is set early on in the trade and paid to the investor, but then the same coupon amount is reclaimed at a much later date. This repayment long after the coupon first set creates a large convexity effect. The TARN is a complex exotic product. It is clearly path-dependent and has embedded convexity, as analysed in this thesis. The relative importance of the path-dependence versus the convexity depends on the maximum amount A - when A is large, the convexity dominates. In practice, both parts are important. A typical TARN might be as follows:

* Maturity: 15 years.

* i^{th} coupon amount paid at S_i, $i = 1, \ldots, 15$, is $\max(y^{10}_{S_{i-1}} - y^2_{S_{i-1}} - K, 0) + c$, where c is a positive constant, $S_0 = 0, S_i = 1 + S_{i-1}$.

* TARN: if total coupon amount hits A before maturity, then the notional is returned, the current coupon is partially paid (to bring the total coupon payment to A), and the trade terminates. If A is not hit, then the trade continues until maturity, with notional returned in 15 years. The final coupon payment is increased so total of all coupons is A.

These are complicated to value and path dependent. To see that there is convexity, consider when A is large. In this case all coupons will most likely be paid in full. Furthermore, at maturity, A will be paid and all earlier coupons will be returned. So in the example, the investor will receive $\max(y^{10}_{S_{i-1}} - y^2_{S_{i-1}} - K, 0) + c$ at S_i but also pay $\max(y^{10}_{S_{i-1}} - y^2_{S_{i-1}} - K, 0) + c$ at 15 years.

The above discussion illustrates the need for a good understanding of convexity corrections in the commercial setting. This is a hard problem to solve in practice for several reasons.

(i) Any modelling approach must take into account, in an appropriate way, the market knowledge of the implied (joint) distributions of various swap rates. For example there is good market knowledge of the marginal distribution of the individual swap rates in their own swaption measure.
(ii) Any pricing model of practical use needs to have a fast implementation.
(iii) In the case of a long payment date the convexity correction will be large. Furthermore in this case a one factor model (as often used by practitioners for convexity calculations) will be inadequate. This has not been investigated in the literature nor yet addressed by practitioners.

As we shall see in the next section in some parts of the literature there is a focus on employing a full term-structure model for convexity corrections either to implement the full model directly or for the purposes of introducing some approximations to the model at a single time slice. This approach is inherently restrictive and limited in terms of what it can deliver on each of the points outlined above.

1.3 Literature

We look at some approaches that have been employed in the literature to tackle convexity corrections. The methodologies used can be broadly divided into two categories. In the first category are the models that have been developed in a full-term structure setting. In the second category are the approaches that take a more local view and aim to set up a model at a single time, incorporating features of the market relevant to pricing the European-type exotic product under consideration. We review the approaches involving a full-term structure model first.

A common modelling framework used in the interest rate markets in general is a short rate model. This class of models is parameterised in terms of the short rate, denoted by r and specified via an SDE:

$$\mathrm{d}r_t = \mu(r_t, t)\mathrm{d}t + \sigma(r_t, t)\mathrm{d}W_t,$$

where W is a Brownian motion in the risk neutral measure and the functions μ and σ are chosen to capture some particular behaviour (like mean-reversion). Short rate models are highly tractable, arbitrage-free and can be more easily implemented than other classes of term-structure models. The main drawback however lies in their calibration aspects. We shall discuss this issue in more depth when we review Murgoci and Gaspar (2016) in Chapter 2. Owing to their ease of implementation and tractability, short rate models remain quite a popular modelling choice for practitioners. In order to address the calibration shortfalls, new techniques are being proposed to calibrate short rate models to swaption prices (for e.g Russo and Torri (2019) & references therein). Di Francesco and Kamm (2022) use the *Gram-Charlier* Expansion to approximate the density of the reference swap rate under the forward measure $\mathbb{F}$ using Hermite polynomials. The technique depends on specifying the bond moments of order m, which they define as $\mathbb{E}_{\mathbb{F}}[(D_{TS_.})^m|\mathcal{F}_t]$ and they use a short rate model approach (they express the short rate dynamics as a difference of two CIR[1] dynamics plus a deterministic function whose role is to ensure a perfect fit to the initial term-structure/yield curve) to model the pure discount bonds. The model parameters can then be chosen to calibrate the model to swaption prices. However, as the authors themselves point out, work in this area is at an early stage and the method developed relies on several simplifying assumptions such as constant parameters and the Brownian framework that 'does not perfectly describe the real world'. The approach proposed by the authors also carries a high computational cost. We note that the goal for the authors is to calibrate the short rate model of choice to both the market term structure and the swaption surface, adapting it to price Bermudan swaptions. They employ their method to price a CMS as well, but due to

[1] Cox-Ingersoll-Ross model where the short rate r is modelled as follows: $dr_t = a(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t$, for a, θ and $\sigma \in \mathbb{R}$

the computational complexity of the model, they only report results for maturities up to 15 years, highlighting the observation that the use of a full-term structure model to price products with naturally long maturities, as is the case for a CMS, is difficult, especially when one tries to capture all the marginal distributions of the swap rates in their own swaption measures.

In order to overcome the calibration problems of short rate models, Market models were developed. Work in this area was pioneered by Miltersen, Sandmann, and Sondermann (1997) and Brace, Gatarek, and Musiela (1997). The defining feature of market models is that they are parameterised in terms of standard market rates such as LIBORs and swap rates, allowing efficient calibration to relevant market prices. In a LIBOR Market Model (LMM), the aim is to develop a full term structure model via the specification of an SDE for each forward LIBOR. The model dimension would then be the same as the number of LIBORs or swap rates being modelled, giving rise to high dimensionality, even if driven by a single Brownian motion, making the model computationally expensive and inefficient in such cases.

Wu and Chen (2010) opt for the LMM framework and approximate the swap rate by a log-normally distributed variable under the forward measure for pricing a CMS. The argument advanced by the authors is that a swap rate is roughly a weighted average of LIBORs, and LIBORs within the LMM framework are log-normally distributed under their associated forward measures. However, a sum of log-normal distributions has an unknown distribution but can still be realistically approximated by a log-normal distribution, an idea reinforced by Brigo and Mercurio (2006).

Henrard (2007) studied the pricing of a CMS payment within a one-factor separable[2] Gaussian HJM (Heath-Jarrow-Morton) and a one-factor separable LMM framework, choosing as numéraire the pure discount bond maturing at the start date T of the reference swap, with associated forward measure $\mathbb{F}$. Within the Gaussian HJM model, volatility is assumed to be deterministic and instantaneous forward rates normally distributed under their associated forward measure. Pure discount bonds are approximated as a function of a common normally distributed random variable which intuitively is the stochastic integral of the volatility along the underlying Brownian motion. Swap rates are then recovered using a Taylor expansion. A similar technique is applied within the one factor Gaussian LMM framework, where LIBORs are assumed to be normally distributed under their associated forward measures. Using the LMM framework is computationally expensive owing to the high-dimensional feature of market models and is restrictive in terms of modeling assumptions such as the choice of distribution of the forward swap rate (we are essentially locked into a normal assumption and the approach offers limited flexibility in moving away from such an assumption).

[2] We shall revisit the concept of separability in Chapter 2

When it comes to convexity corrections, practitioners need a solution that can be efficiently used in practice. One of the first methods available to them was an ad hoc approach proposed by Hull (1997). Hull used a second order Taylor expansion of the PVBP as a function of the reference swap rate around its initial value, and the martingale property, to obtain an expression for the expectation of the reference swap rate under the forward measure. Lognormal assumptions on the swap rate under the forward measure led to a formula easy to apply in practice. Benhamou (2000) analysed Hull's approach further. Hull's work marked the beginning of a literature which took a local approach to the problem and was aimed at practitioners.

We observe that the term-structure models described above, when employed for convexity corrections, are set up under an appropriate forward measure. As we have discussed, every forward swap rate has a natural associated (swaption) measure with respect to which it is a martingale. The practitioner literature tends to work in swaption measure.

Working in the swaption measure, Hunt and Kennedy (2000) proposed a terminal swap rate modelling approach that approximates the ratio of pure discount bonds over the PVBP as a function of the reference swap rate at a fixed time as a function of the reference swap rate. This approach provides an arbitrage-free model at a fixed time which can be used for convexity corrections and can incorporate directly the market implied distribution of the reference swap rate. Because of its ability to incorporate market smiles, this approach continues to be popular in the industry today. However it is not easy to see how to extend the method to include another factor.

This terminal swap rate model was later discussed in Chapter 16 of Andersen and Piterbarg (2010b) who point out that it can be considered as the one dimensional conditional expectation of a multifactor model. Cedervall and Piterbarg (2012) use this insight to develop an arbitrage-free approach for convexity corrections in the forward measure under normal and log-normal assumptions. This approach enables the inclusion of swap rate correlations but loses that ability to capture the implied market distributions accurately.

The terminal swap rate models were a precursor to Markov-functional models which were developed by Hunt, Kennedy, and Pelsser (2000) in the term structure setting and widely used because of their ability to calibrate to liquid market instruments. Bermin and Williams (2017) combine the local approach of modelling at a fixed time and the technique of a Markov-functional sweep to modify the distribution given to the swap rate by a term structure model of choice to price cash settled swaptions.

We return to discuss further the references above that take a local approach to convexity corrections in Chapter 3. We mention here the work by Chen, Oosterlee, and Van

Weeren (2010). This paper deals with the payment time being that of the first fixed coupon of the reference swap rate. They use a two-factor SABR model to characterise the distributions of rates at a fixed time. They note that the valuation of a standard CMS payment involves taking an expectation involving two rates: the reference swap rate and a LIBOR. At the reset date T, the pure discount bond maturing at the payment date of the CMS payment can be expressed as a function of a LIBOR. They assume that the PVBP can be approximated as a function of the reference swap rate. In their approach, they consider the correlation between the reference swap rate and the LIBOR. It is unclear how to extend the method beyond a single short payment date. In this thesis, we build upon the observation that correlation might matter for long payment dates and we will focus on correlation between relevant swap rates.

Finally we note that some recent papers are concerned with extending existing methods to multi-curves or LIBOR replacement. These are important effects that can be incorporated but do not fundamentally affect issues covered in this thesis. We omit them for clarity. The models developed here provide a basis which can be built upon to include these extra features.

The literature on approaches to convexity corrections, is still evolving to cope with the ever-changing nature of the markets. This prompts us to ask the question at this point: What really matters when it comes to the pricing and hedging of derivatives that require convexity corrections? How have practitioners' approaches so far coped and how can they be improved?

1.4 Objective and Structure

The first part of this thesis is dedicated to addressing point (iii) raised in Section 1.2. In particular, in Chapter 2, we introduce a single-time Markov-functional model (single-time MFM) set up under the forward measure $\mathbb{F}$, associated with taking $D_{\cdot T}$ as numéraire to investigate what matters most when pricing a CMS and some related options. Note that since the valuation of each payment of a CMS can be carried out independently, we will focus on a single payment of a CMS in the numerical analysis. Using a Markov-functional approach, as we will see in the chapter, provides us with a flexible framework to specify the 'target real world model' and explore the properties that matter when pricing a CMS. The development of the model at a fixed time enables us to have control over the marginal distributions of the swap rates we wish to capture. Incorporating a second factor extends the model flexibility by enabling us to model the joint distribution of the swap rates at a single time without undue compromises enforced by the martingale dynamics of full term-structure models. However, the single-time MFM is too computationally expensive to be attractive in practice. It nonetheless provides a rich framework through which convexity corrections can be studied, which we set out to do in Chapter 3. We use the single-time MFM to study the market properties that have a significant impact on convexity corrections. We firstly determine the market-implied marginal distributions of the swap rates in their own swaption measures that have a significant effect. This analysis is carried out in the one-factor context. We then introduce a second factor to analyse which aspects of the joint distribution of the swap rates matter. For this investigation, we assume that the swap rates are log-normal in their own natural measures and that the joint distribution of the model driver is bivariate Gaussian under the forward measure $\mathbb{F}$. Under these assumptions, the parameter that summarises the joint distribution is the correlation. We hence investigate the correlation effect on convexity corrections. We point out that the numerical analysis in this chapter is carried out under realistic but fairly stressed market conditions. The idea is that if the single-time MFM is able to identify the aspects of the joint distributions of the swap rates that matter in pricing a CMS, any model developed that captures this information will fare well under relaxed market conditions. Given that the single-time MFM fails to satisfy point (ii) discussed in Section 1.2, we use the insights obtained from Chapter 3 to formulate an appropriate practical approach, which we refer to as the 'MF-Lite model' that takes into consideration the properties that matter most when pricing swap-based convexity-related products, hence satisfying points (i) and (ii). In Chapter 4, we propose a computationally fast and efficient one-factor MF-Lite model, set up at a fixed time in the swaption measure. We then propose in Chapter 5 a practical two-factor Markov-functional model developed at a fixed/local time. For this approach, unlike the model developed in Chapter 4, we start off the modelling process in the forward measure. We shall explain later in this chapter the challenges of setting the two-factor approach in the swaption measure. We end Chapter 5 by proposing a simplified two-factor methodology in the swaption measure that still has some practical interest.

1.5 Preliminaries

1.5.1 No-arbitrage pricing framework

In this section, we provide a general, but concise overview of the no-arbitrage pricing framework, that will be used throughout this thesis. The no-arbitrage pricing framework, introduced by Harrison and Kreps (1979) is one that has been robustly discussed and explained in the literature (and extended by Schachermayer (1994), Delbaen and Schachermayer (1994) and others). We refer the interested reader to Andersen and Piterbarg (2010a) and Filipovic (2009), amongst others for a detailed discussion, and its application to interest rate modelling. Whilst we do not explore the basics of stochastic calculus here, we refer the reader to Karatzas and Shreve (1991) or Revuz and Yor (2013). For this thesis, we follow closely the framework as established in Hunt and Kennedy (2004).

An economy, which we shall denote by $\mathcal{E}$, is characterised by two components: a (finite) number of assets with a model for the evolution of their prices $(A^{(i)})_{i=1}^n$, and a set of strategies $(\phi^{(i)})_{i=1}^n$ (we will require that the strategies satisfy certain conditions which we will get to shortly) in the economy. We start off with a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leq t}, \mathbb{Q})$, where the filtrations $\{\mathcal{F}_t\}$ satisfy the usual conditions i.e (i) $\mathcal{F}_0$ contains all $\mathbb{Q}$-null sets of the complete space $(\Omega, \mathcal{F}, \mathbb{Q})$ and $\mathcal{F}_t$ is right-continuous $\mathcal{F}_t = \cap_{s>t}\mathcal{F}_s$. The asset price processes are modelled as continuous semimartingales on the filtered probability space. We introduce the notion of asset filtration, which we denote by $\{\mathcal{F}_t^A\}$. We skip the mathematical detail, but we provide an intuitive explanation. The asset filtration can be interpreted as all the information that is available to us by observing the evolution of asset prices over time. We will assume $\mathcal{F}_t = \mathcal{F}_t^A, \forall t \geq 0$.

We can interpret $\phi_t^{(i)}$ as the number of units we hold for asset i with price process $A_t^{(i)}$ at time t. We impose that decisions about holdings of the asset is made an instant before subsequent trading occurs, i.e ϕ is $\{\mathcal{F}_t^A\}$-predictable. Another restriction is that we do not allow money to be injected or removed from a trading portfolio (this is known as the self-financing property; we assume that all strategies are self-financing). The notion of *no-arbitrage* is central to derivatives pricing. An arbitrage is a self-financing trading strategy that either (a) starts with a negative wealth and ends with non-negative wealth with probability one or (b) starts with non-positive wealth and ends with non-negative profit with probability one or strictly positive wealth with strictly positive probability. If there exists no such strategy, the economy is said to be arbitrage-free. A third condition that we impose on the strategies that we allow within the economy is that of *admissibility*. It is a technical requirement, one we shall not get into, which is imposed to remove troublesome strategies from the economy.

We introduce some probabilistic tools required for the discussion that follows. We define the concept of a numéraire.

Definition 1.5.1. Let measure $\mathbb{N}$ be defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leq t<\infty})$. The measure $\mathbb{N}$ is an equivalent martingale measure (EMM) if $\mathbb{N} \sim \mathbb{Q}$ and there exists an almost surely strictly positive generalized price process, referred to as the numéraire N, such that the numéraire-rebased asset price process, $A^N := \frac{A}{N}$, is an $\{\{\mathcal{F}_t\}, \mathbb{N}\}$-martingale. The pair $(N, \mathbb{N})$ is referred to as a numéraire pair.

Observe that the positivity assumption on the numéraire guarantees that the ratio A^N is well-defined. The economy $\mathcal{E}$ is arbitrage-free if there exists at least one EMM $\mathbb{N} \sim \mathbb{Q}$, associated with a numéraire N such that numéraire-rebased asset price processes are martingales under $\mathbb{N}$.

Remark 2: The converse is not true in a continuous time setting. We refer to Delbaen and Schachermayer (1999a) and Delbaen and Schachermayer (1999b) for a detailed *discussion.*

We introduce the concept of completeness. In an arbitrage-free economy, we can price a derivative (contingent claim) by constructing an admissible strategy that replicates the derivative's payoff. The derivative is then said to be *attainable*. An economy $\mathcal{E}$ is complete if any derivative (contained in a suitable set) is attainable. The theory of completeness of the economy gives us sufficient and necessary conditions to know that such a portfolio exists. Let the economy $\mathcal{E}$ admit at least one EMM. The economy $\mathcal{E}$ is said to be complete if and only if the EMM is unique.

Theorem 1.5.1. *Let $\mathcal{E}$ admit a numéraire pair $(N, \mathbb{N})$ and let V_T be the value at time T of an attainable contingent claim. Then, for $0 \leq t \leq T < \infty$, the time-t value of the claim, V_t is given by:*

$$V_t = N_t \mathbb{E}_{\mathbb{N}} \left[\frac{V_T}{N_T} \middle| \mathcal{F}_t \right]. \tag{1.1}$$

Existence of a numéraire pair guarantees our economy is arbitrage-free and equation (1.1) follows from the martingale property of numéraire-rebased price processes.

The measures $\mathbb{Q}$ and $\mathbb{N}$ can be related to each other, the result of which is given by the *Radon-Nikodỳm Theorem.*

Theorem 1.5.2. *Let $\mathbb{Q}$ and $\mathbb{N}$ be probability measures on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t < \infty})$, and suppose $\mathbb{N} \sim \mathbb{Q}$ with respect to $\mathcal{F}$. Then, $\forall t \in [0, \infty)$*

$$\zeta_t := \left. \frac{\mathrm{d}\mathbb{N}}{\mathrm{d}\mathbb{Q}} \right|_{\mathcal{F}_t}$$

defines an a.s strictly positive uniformly integrable $\{\mathcal{F}_t\}$-martingale under $\mathbb{Q}$. The process $\{\zeta_t : t \geq 0\}$ is referred to as the Radon-Nikodỳm derivative.

Remark 3: Theorem (1.5.2) holds for $t = \infty$ as well. However, we are restricting ourselves to a finite time horizon here.

Under the assumption of completeness, the Radon-Nikodỳm derivative is simply the ratio of numéraires.

Theorem 1.5.3. *Assume the economy $\mathcal{E}$ is complete. Let $0 \leq t < \infty$. Let $\mathbb{Q}$ be the equivalent martingale measure with respect to the numéraire Q. Let $\mathbb{N}$ be the equivalent martingale measure with respect to the numéraire N. The Radon-Nikodỳm derivative that allows us to change from the measure $\mathbb{Q}$ to $\mathbb{N}$ is given by*

$$\left. \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{N}} \right|_{\mathcal{F}_t} = \frac{Q_t}{N_t} \frac{N_0}{Q_0}. \tag{1.2}$$

For the rest of this thesis, we assume that the economy is complete.

1.5.2 Interest rates and standard derivatives

Let $S_0 < S_1 < \ldots < S_n < \infty$ be an increasing sequence of maturity dates, referred to as a *tenor structure*. We adopt the convention that $S_0 \equiv T$, where T is a fixed time and we define the accrual factor $\alpha_i := S_i - S_{i-1}$, for $i = 1, \ldots, n$, as the time elapsed between S_{i-1} and S_i. We introduce the basic fundamental asset of the economy, which is the Pure Discount Bond (PDB). A pure discount bond is a financial instrument which pays the holder one unit of currency at a pre-determined maturity. We denote by D_{tT}, $t \le T$, the value at time t of the pure discount bond with maturity T (Note that $D_{TT} = 1$). Going forward, we will assume we have specified an arbitrage-free and complete economy consisting of PDBs $D_{\cdot S_i}$, $i = 0, \ldots, n$.

We can define the forward *LIBOR* (London Inter-Bank Offer Rate) from the pure discount bonds as follows: Denote by $L_t^i := L_t[S_{i-1}, S_i]$, the forward LIBOR for the period $[S_{i-1}, S_i]$, and we have that

$$L_t^i := \frac{D_{tS_{i-1}} - D_{tS_i}}{\alpha_i D_{tS_i}}. \tag{1.3}$$

Remark 4*: Throughout this thesis, we will assume a single-curve framework, in that we can move from the yield curve to the forward curve via the above relationship. However, this does not necessarily hold, which led to the development of the multi-curve framework, an interesting discussion of which can be found in Henrard (2014).*

We now formally introduce the forward swap rate, which will be the crucial interest rate we will focusing on in this thesis. We do so by analysing a basic interest rate instrument, the *interest rate swap.*

Valuing a Payer's Swap

In a physically settled vanilla payer's interest rate swap, a floating short term rate, for example the LIBOR, is received in exchange for paying a fixed rate K. Suppose for a given swap starting at time T, cashflows arise at times $S_1, \ldots, S_n$. The i^{th} cashflow for the fixed leg arising at time S_i is given by $\alpha_i K$, and that of the floating leg is given by $\alpha_i L_{S_{i-1}}[S_{i-1}, S_i]$. Hence the value at time t for the i^{th} cashflow of the fixed leg is given $\alpha_i D_{tS_i} K$. The value of the floating portion is $\alpha_i D_{tS_i} L_t^i$. Figure 1.1 below illustrates the payout structure of the interest rate swap.

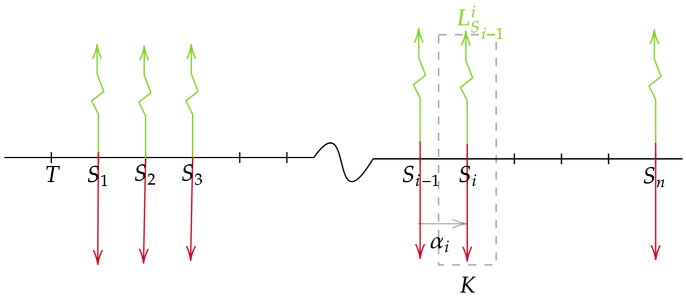

Figure 1.1: Payers Interest Rate Swap

Taking into consideration all cashflows, the value of the fixed leg is given by:

$$V_{fix}(t) := K \sum_{i=1}^{n} \alpha_i D_{tS_i}.$$

The value of the floating leg is given by:

$$\begin{aligned} V_{flt}(t) &:= \sum_{i=1}^{n} \alpha_i D_{tS_i} L^i_t \\ &= D_{tT} - D_{tS_n}. \end{aligned} \tag{1.4}$$

Equation (1.4) follows from the definition of LIBOR.

The net value of the payer's swap at some time t is then equal to:

$$V^S_t := V_{flt}(t) - V_{fix}(t) = D_{tT} - D_{tS_n} - K \sum_{i=1}^{n} \alpha_i D_{tS_i}. \tag{1.5}$$

Definition 1.5.2. The *Forward Par Swap Rate*, denoted by y^n_t is the fixed rate K that values the payer swap to zero at time t.

$$y^n_t := \frac{D_{tT} - D_{tS_n}}{\sum_{i=1}^{n} \alpha_i D_{tS_i}}. \tag{1.6}$$

The denominator term $\sum_{i=1}^{n} \alpha_i D_{tS_i} =: P^n_t$ is referred to as the *present value of a basis point* (PVBP). The PVBP, being a strictly positive process and a linear combination of pure discount bonds, can be chosen as numéraire, and from (1.6), we observe that the swap rate is of form assets $(D_{tT} - D_{tS_n})$ over numéraire. Associated with P^n as numéraire, we have the EMM $\mathbb{S}^n$, commonly referred to as the swaption measure, with respect to which the forward par swap rate is a martingale.

Using (1.6), equation (1.5) can be re-written as:

$$V^S_t = P^n_t (y^n_t - K).$$

Valuing physically-settled swaptions and digital swaptions

A swaption is an option on an interest rate swap. These derivatives are actively traded on the market, and the prices carry information on the probability distribution of the underlying swap rate under its associated swaption measure. There are two types of swaptions: payer's swaption and receiver's swaption. In the former case, at the pre-determined exercise date (usually it will coincide with the swap start date T) of the option, the holder has the right, but not the obligation to enter a swap whereby he pays the fixed rate, and receives the market swap rate in exchange. The reverse is true in the latter case. We proceed with the pricing of a payer's swaption. Assume the payer's swaption is based on the payer's swap defined above. At time T, the effective payoff to the option holder is $\max(P_T^n(y_T^n - K), 0) = P_T^n(y_T^n - K)_+$. Working under the swaption measure $\mathbb{S}^n$, by the valuation formula, the value at some given time t of the swaption is given by

$$V_t^n := P_t^n \mathbb{E}_{\mathbb{S}^n}[(y_T^n - K)_+ | \mathcal{F}_t].$$

Suppose we set up a Black's model for the swap rate, (i.e we assume that y^n is a log-normal martingale under $\mathbb{S}^n$ with constant volatility σ_n), we will obtain the following explicit expression for the price of a swaption:

$$V_t^n = P_t^n(y_t^n \Phi(d_1) - K\Phi(d_2)),$$

where $d_1 := \frac{\log(y_t^n) - \log(K)}{\sigma_n\sqrt{T-t}} + \frac{1}{2}\sigma_n\sqrt{T-t}$ and $d_2 := d_1 - \sigma_n\sqrt{T-t}$.

In setting up a pricing model for swap-based products, it is essential that the model is calibrated to the relevant vanilla instruments i.e to the vanilla swaptions. Equivalent to calibrating a model to vanilla swaption prices, is calibrating the model to market prices of *digital swaptions*. Similar to swaptions, there exists two types of digital swaptions: payer's digital swaption and receiver's digital swaption. In the former, the holder receives P_T^n at T if the underlying swap rate is *above* the strike K at setting date T, and in the latter case, the holder receives P_T^n at T if the underlying swap rate is *below* the strike K at setting date T. The value at time t of a payer's digital swaption with exercise date T is given by

$$\bar{V}_t^n := P_t^n \mathbb{E}_{\mathbb{S}^n}[\mathbb{1}(y_T^n > K) | \mathcal{F}_t].$$

In the Black's model example given above, we have that $\bar{V}_t^n = P_t^n \Phi(d_2)$.
Differentiating V_t^n with respect to K, we can deduce the relationship between swaption and digital swaption prices:

$$\bar{V}_t^n(K) = -\frac{\mathrm{d}}{\mathrm{d}K} V_t^n(K).$$

CHAPTER 2

A Markov-functional approach for investigating convexity corrections

As discussed in the introduction, we are mainly focused on studying the valuation of European-type products whose payoffs can be viewed as a function of a set of underlying swap rates with an 'unnatural' payment schedule[1](Pelsser (2000)). In this chapter, we set out to develop a Markov-functional model through which we can study the pricing problem. We motivate the rest of this chapter by providing some high-level insights to naturally arising questions at this point.

∗ *What is a Markov-functional model?*
Generally speaking, any model whereby the pure discount bond prices can be expressed as a function of some process (which is Markovian with respect to some martingale measure - hence the term *Markov*) belongs to the class of Markov-functional models (MFM). By that description, commonly encountered interest-rate models like the short-rate model or market models are Markov-functional, but they fail to satisfy certain properties/requirements that would be crucial for the study we are interested in, in this thesis, which we shall touch upon shortly.

We focus on the modelling framework developed by Hunt and Kennedy (2000). In Section 2.1, we provide a formal, general formulation of the model. The basic underlying idea in this case is that we aim to keep the dimensionality of the model low, whilst being able to effectively calibrate the model to market prices. We will consider a swap version

[1]This leads to the need for some form of adjustments/convexity corrections

of the model. The model, as described in Hunt and Kennedy (2000) is developed in a full term-structure setting. We will show in Section 2.1 how we can adapt the setup to model the forward swap rates at a given fixed time, hence the name 'single-time' MFM. We could potentially extend the model to earlier times, but this will be an unnecessary step for the valuation of European-type products which the single-time MFM is designed to price.

∗ *Why a Markov-functional approach?*
The aim here is to set up a model that provides a rich and flexible framework through which convexity corrections can be studied, and provides a benchmark against which existing approaches can be compared. We want to ensure that the model is properly calibrated to closely-related liquid instruments and captures the features of the market which are relevant. A Markov-functional approach can be set up to satisfy the above, unlike other approaches, and we will discuss this further in Section 2.1.1. We could have set up a short rate model. We would start off by specifying an SDE for the short rate, from which we can obtain the PDBs and the forward swap rate at T. There is a complicated relationship between the short rate and the swap rates, and any variation introduced in the modelling of the short rate will have a non-transparent effect on the forward swap rate. We discuss this point at the end of Section 2.1.1.

Given that we can obtain a market-implied distribution for the swap rate in its own swaption measure from publicly available information on swaption prices, it would make sense to construct a model that captures this information directly. We could turn instead to a swap market model. But this approach suffers three particular drawbacks: one is the high-dimensionality issue which makes the model computationally expensive (hence, not one that would be particularly attractive for an investigative analysis). Secondly, it is a full-term structure model, that will be too extensive for the pricing problem we are interested in. We could introduce some approximations to overcome this, but that could potentially result in significant arbitrage introduced in the model. Thirdly, it is not clear how one could introduce a general non-Gaussian copula to model the swap rate distributions in the approach.

The single-time MFM, on the other hand, gives us considerable flexibility - for instance, in contrast to the short rate model, the choice of the model driver is disassociated from the modelling of the forward swap rates, in the sense that we can vary the distribution of the model driver without distorting the marginal distribution of the swap rates under their respective swaption measures. It is also low-dimensional, hence addressing the shortcomings of market models.

∗ *What are the modelling assumptions?*
We will set up a two-factor single-time MFM. Hence, we assume that there exists a two-dimensional driver that summarises the state of the economy at time T, whose joint

distribution under the forward measure $\mathbb{F}$ is a modelling choice. We provide in Chapter 3 a justification for the choice of two factors.

We also assume that we know the market prices of a suitable set of payer swaption prices for different strikes, from which we can determine the market-implied marginal distributions of the forward swap rates in their own swaption measures.

Finally, we introduce the notion of a prior model that allows for efficient calibration of the Markov-functional model. Indeed, the ability to efficiently capture the market-implied distributions lies in the fact that the MFM bypasses the SDE formulation of the forward rates, but instead calibrates the model by numerically constructing a functional form for the forward swap rate in terms of the driver. This in turn makes the model less transparent in its properties. However, this can be overcome if we can formulate a transparent prior model that would usually be of low-dimension, but admits arbitrage. The idea is that the prior model gives us a rough approximation of the forward swap rates we want to model and we introduce a perturbation on the prior (in the form of a monotonic increasing function - this is the third model assumption) that would remove the arbitrage and calibrate the model to the desired marginal distributions. In Section 2.2, we discuss in depth the concept of a prior model, and we provide an example of the prior model setup based on local volatility separable market models. In Chapter 3, we show, in the specific case of Gaussian assumption on the model driver and log-normal assumptions on the swap rates in their own swaption measures, we can use the swap market model as the basis for the prior model setup and lean on information from LIBORs (the context for LIBORs will become clearer in Chapter 3) to inform the choice of prior model for the single-time MFM.

$*$ *What can we gain from the single-time MFM?*
The single-time MFM gives us a rigorous and flexible mathematical platform through which convexity corrections can be studied. Indeed, in Chapter 3, we demonstrate how we can use the model to study convexity corrections through the MF lens. The model however is too computationally expensive to be attractive and viable in practice. It nonetheless provides a basis from which we can formulate faster, efficient Markov-functional based approaches that are numerically close to the single-time MFM. This will constitute the focus of Chapter 4 and Chapter 5.

2.1 A single-time Markov-functional model (MFM)

We want to set up a two-factor model designed for the accurate pricing of a CMS for which the payoff depends on a set of swap rates with varying end dates and a fixed start date. To do so, we use a Markov-functional approach. A characterising feature of a Markov-functional model that allows for efficient practical implementation is the ability to express the pure discount bonds at any given time as a function of some low-dimensional process. Hunt, Kennedy, and Pelsser (2000) propose the Markov-functional modelling framework in a full-term structure setting, with the aim to price multi-temporal derivatives. We provide a formal, yet very general definition of the Markov-functional model, taken from Kennedy (2010):

Definition 2.1.1. Let T' be a finite time horizon, and assume that the underlying assets of an economy are pure discount bonds with maturities from a non-empty set $\mathcal{T} \subset (0, T']$. A model of the economy is said to be Markov-functional if there exists a numéraire pair $(N, \mathbb{N})$ and an n-dimensional $\{\mathcal{F}_t\}_{t\in[0,T']}$- adapted process x such that

(a) x is a Markov process under $\mathbb{N}$

(b) The pure discount bond prices at time $t \leq T$, for all $T \in \mathcal{T}$ are of form:

$$D_{tT} = D_{tT}(x_t).$$

We point out that, by standard theory which we have seen in the preliminaries, the existence of a numéraire pair ensures that the model is arbitrage-free. The Markov process $\{x_t, t \geq 0\}$ is referred to as the driving process (or simply driver) of the model. The number of factors in the model is equal to the dimension of the driving process. Any interest rate model used in practice, be it a short rate model, or a market model are Markov-functional. The definition gives us the general properties of the model we aim to develop, but not the setting up of the model itself. To ensure efficient implementation of the model, the aim is to keep the dimensionality of the driver low, a property not possessed by market models, whilst being able to accurately capture the market-implied distributions of the forward rates. It is the freedom to choose the functional forms of the model that enables accurate calibration to those distributions. We shall see in the next section how we choose to calibrate the model and construct the required functional forms.

As discussed in Hunt, Kennedy, and Pelsser (2000), setting up a Markov-functional model in a full term structure setting under the terminal measure involves starting from the last payment date and working recursively backwards to determine functional forms of pure discount bonds at earlier times. A market rate can be chosen to calibrate to at each time slice and correlations can be controlled through the specification of the driving process. However, it is only possible to calibrate to the market distribution of a single rate for each time slice, an important limitation as we shall see when pricing a CMS. Indeed, for exotic European derivatives, since payments can be determined by observing the market at a single pre-determined time, we will apply the Markov-functional technique in a local setting (i.e, calibrating at a single time). We illustrate in the diagram below the general idea underpinning the Markov-functional approach we take here.

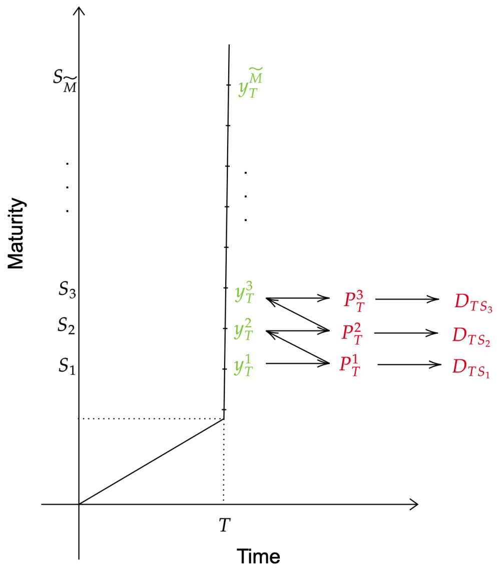

Figure 2.1: Single-time MFM: A simple illustration

For our application, we only need to define the model at a fixed time T (at the start date of the reference swap rate), so the driving process reduces to an n-dimensional random variable (Since we are interested in setting up a two-factor model, $n = 2$. We will later provide a justification as to why we deem a two-factor model appropriate here). We are aiming to model a finite number of pure discount bonds at a fixed time, with varying maturities. At each maturity, denoted by $S_j, j \in \{1, \dots, \tilde{M}\}$, $\tilde{M} < \infty$, we would want the model to be able to capture the market implied distribution of the swap rates under their respective swaption measures. We calibrate the model to the known market prices via the *Markov-functional sweep* technique. We unfold the model recursively forward in maturity, whereby at each maturity S_j the market prices determine the functional form of the swap rate y_T^j (illustrated as a green variable in figure 2.1), and this determines the functional forms of the PVBP and the pure discount bonds (the red variables in figure 2.1). To see this, we point out that we can express the PVBP and the pure discount bonds in terms of swap rates only. By definition stated in 1.5.2, we have that:

$$y_T^j = \frac{D_{TT} - D_{TS_j}}{\sum_{k=1}^{j} \alpha_k D_{TS_k}} = \frac{1 - D_{TS_j}}{P_T^j}.$$

Hence,

$$D_{TS_j} = 1 - y_T^j P_T^j. \tag{2.1}$$

First, it can be observed that

$$P_T^j = P_T^{j-1} + \alpha_j D_{TS_j}, \tag{2.2}$$

with $P_T^0 \equiv 0$. Substituting the expression for D_{TS_j} from equation (2.1) in the expression for P_T^j given in equation (2.2), we have

$$\begin{aligned}
P_T^j &= P_T^{j-1} + \alpha_j(1 - y_T^j P_T^j) \qquad (2.3)\\
&\Rightarrow P_T^j = \frac{\alpha_j}{(1+\alpha_j y_T^j)} + \frac{P_T^{j-1}}{(1+\alpha_j y_T^j)}\\
&= \frac{\alpha_j}{1+\alpha_j y_T^j} + \frac{\alpha_{j-1}}{(1+\alpha_j y_T^j)(1+\alpha_{j-1} y_T^{j-1})} + \frac{P_T^{j-2}}{(1+\alpha_j y_T^j)(1+\alpha_{j-1} y_T^{j-1})}\\
&= . \quad . \quad .\\
&= \sum_{k=1}^{j} \alpha_k \prod_{m=k}^{j} \frac{1}{1+\alpha_m y_T^m}. \qquad (2.4)
\end{aligned}$$

We can now work out an expression for the pure discount bonds in terms of the swap rates as follows:

$$D_{TS_j} = 1 - y_T^j \sum_{k=1}^{j} \alpha_k \prod_{m=k}^{j} \frac{1}{1+\alpha_m y_T^m}. \qquad (2.5)$$

Remark 5: We can see from the equality in (2.3) *and the implication following it, at maturity* S_j*, assuming we have derived all the required functional forms at time* S_{j-1}*, we only need knowledge of* y_T^j *to specify* P_T^j*. Then from equation* (2.1)*, we can work out the functional form of* D_{TS_j}*, thereby explaining the recursive step forward in maturity, starting at* S_1*, in unfolding the model. Alternatively, we can work out the functional forms of pure discount bonds and the PVBP from those of the swap rates at earlier maturities by using the functional relationships defined above.*

So far, we started off with a general description of a Markov-functional model in the usual full term structure setting, and proposed an approach built upon the properties of the model, at a local (fixed) time. We have only commented on a very basic description of the setup of the Markov-functional approach without explaining how to calibrate the model and thus get hold of the functional forms themselves. The next sections are dedicated to this particular task.

2.1.1 General model setup and assumptions

We discuss the model setup. We fix some time T. Assume we have a finite tenor structure $T < S_1 < S_2 < \ldots < S_{\tilde{M}} < \infty$, with the time elapsed between S_{j-1} and S_j given by α_j. We work under the forward measure $\mathbb{F}$, associated with taking $D_{\cdot T}$ as numéraire. We lay out the modelling assumptions.

A1 Assume there exists a bivariate random variable $\mathbf{x}_T$, where $\mathbf{x}_T = (x_T^{(1)}, x_T^{(2)})$ which summarises the state of the economy at time T (so $\mathbf{x}_T$ is taken to be the model driver) and pure discount bonds at time T can be written in the form $\{(D_{TS_j}(\mathbf{x}_T))_{j=1}^{\tilde{M}} : T < S_j < \infty\}$.

We can specify the joint distribution of $\mathbf{x}_T$ under $\mathbb{F}$ by a chosen copula and associated marginal distributions for $x_T^{(1)}$ and $x_T^{(2)}$. We assume that the marginal density functions with respect to the Lebesgue measure exist.

At each maturity date $S_j, j \in \{1, \ldots, \tilde{M}\}$, we would naturally want the model to be able to correctly price liquid vanilla instruments most closely related to the product we are interested in pricing. The calibrating instruments we use are swaptions (we can equally use digital swaptions, the prices of which can be recovered from swaptions. In practice, some *smart* interpolation technique is used to construct a curve for the remaining strikes in a way that does not allow for arbitrage and digital swaption prices can be obtained by differentiating the interpolation function with respect to strike). To this end, we make the following assumption:

A2 Assume we are given a suitable set of payer swaption prices of different strikes K, on the forward par swap rate $\{y_T^j, j \in \{1, \ldots, \tilde{M}\}\}$, which pay $P_T^j(y_T^j - K)_+$ at time T. This is equivalent to knowing the prices of digital swaptions with same underlying and strike. These prices would determine the distribution of y_T^j under the associated swaption measure $\mathbb{S}^j$.

The techniques developed in Hunt, Kennedy, and Pelsser (2000) for calibrating a Markov-functional model relied extensively on the one-dimensionality property of the driver. Hunt and Kennedy (2000) discussed the extension to multi-dimensional Markov-functional models. The authors argue that in order to make the functional fitting efficient when moving to a higher dimension, through an *informed* choice, we can collapse the dimension of the driver to a one dimensional variable and apply the univariate techniques for the Markov-functional fitting. Through this, it is possible to generate a wide range of models that calibrates to the desired marginal and joint distributions. The choice of the method to collapse the dimension comes from a *prior model*, which generally speaking is a rough approximation of the swap rate in terms of the driver $\mathbf{x}_T$. At a given maturity S_j, say we have an approximation for the swap rate y_T^j, which we denote by $\hat{y}_T^j(\mathbf{x}_T)$. We assume that the functional dependence of the swap rate on $\mathbf{x}_T$ is only via the prior model. The prior model is chosen to capture the joint distribution of the swap rates, but it may admit significant arbitrage. To overcome this, we introduce a Markov-functional sweep in the form of a monotonic function that acts as a small perturbation on the prior model and removes arbitrage. We shall elaborate on the choice of prior model in Section 2.2.

A3 We assume we can express the swap rate y_T^j as a monotonic increasing function of the prior $\hat{y}^j_T(\mathbf{x}_T)$:

$$y^j_T(\mathbf{x}_T) := f^j(\hat{y}^j_T(\mathbf{x}_T)), \tag{2.6}$$

where $\hat{y}^j_T : \mathbb{R}^2 \to \mathbb{R}$ and $f^j : \mathbb{R} \to \mathbb{R}$ is some monotonic increasing function.

Remark 6*: Given the model driver $\mathbf{x}_T$ and a chosen prior $\hat{y}^j_T$, provided that we start the functional fitting at the first maturity date S_1 and we unfold the model forward in maturity, the functional form f^j is uniquely determined by the market-implied marginal distribution of the swap rate under its associated swaption measure. The functional form y^j_T in equation* (2.6) *is unaltered by a monotonic transformation of the prior $\hat{y}^j_T$.*

To see this, first assume that we have collapsed the dimension of the driver via the prior. The functional fitting now relies on the one-dimensionality property of the prior. Define $\tilde{y}^j_T = g^j(\hat{y}^j_T)$, where $g^j : \mathbb{R} \to \mathbb{R}$ is a deterministic monotonic function. We use this modified prior and calibrate the model, we would get $y^j_T(\mathbf{x}_T) := \tilde{f}^j(\tilde{y}^j_T(\mathbf{x}_T)) = \tilde{f}^j(g^j(\hat{y}^j_T(\mathbf{x}_T))$, for some monotonic function $\tilde{f}^j : \mathbb{R} \to \mathbb{R}$. For any $x \in \mathbb{R}$, we define

$$F^y_j(x) := \mathbb{E}_{\mathbb{S}^j}[\mathbb{1}\{y^j_T \le x\}],$$

*where F^y_j is known under the swaption measure $\mathbb{S}^j$ (in line with assumption **A2**). We also define:*

$$\hat{F}^y_j(x) = \mathbb{E}_{\mathbb{S}^j}[\mathbb{1}\{\hat{y}^j_T \le x\}].$$

Since we are calibrating the model to the given marginal distribution under $\mathbb{S}^j$, we have chosen f^j such that:

$$\begin{aligned} \mathbb{E}_{\mathbb{S}^j}[\mathbb{1}\{f^j(\hat{y}^j_T) \le x\}] &= \mathbb{E}_{\mathbb{S}^j}[\mathbb{1}\{\hat{y}^j_T \le (f^j)^{-1}(x)\}] = F^y_j(x) \\ &\implies \hat{F}^y_j((f^j)^{-1}(x)) = F^y_j(x). \end{aligned}$$

By the same line of reasoning, for the modified prior, we would obtain:

$$\hat{F}^y_j((\tilde{f}^j \circ g^j)^{-1}(x)) = F^y_j(x).$$

Equating the LHS, we therefore have that:

$$\begin{aligned} &\hat{F}^y_j((f^j)^{-1}(x)) = \hat{F}^y_j((\tilde{f}^j \circ g^j)^{-1}(x)) \\ &\implies (f^j)^{-1}(x) = (\tilde{f}^j \circ g^j)^{-1}(x), \end{aligned}$$

by monotonicity of $\hat{F}^y_j$, or equivalently, we have that $f^j \equiv \tilde{f}^j \circ g^j$.

We end this section with two observations: Firstly, the model thus set up allows us to vary the joint distribution of $\mathbf{x}_T$, without inadvertently changing the market-implied distribution of the swap rates at time T under their associated swaption measure. Secondly, viewing things from the Markov-functional perspective, it is clear that given market swaption prices, the distribution of the driver in a one-factor model has no effect on the model, and therefore CMS prices. To illustrate the second point further, consider, for simplicity, a one-factor short-rate model. One would start off by formulating an SDE for the short rate, from which one can recover the pure discount bond at any given fixed time T and maturity as a function of the short rate r_T. One can thus work out the functional form for the swap rate from those of the pure discount bonds with varying maturities. This enables us to find an approximate distribution of the swap rate under its associated swaption measure. Given the intricate relationship between the swap rate and the short rate, the choice we make to model the short rate will have a non-transparent effect on the distribution the model ascribes to the swap rate at time T. At this point, we have lost control over the choice of distribution of the swap rate at a fixed time. Murgoci and Gaspar (2016) study closed form convexity corrections in an affine term structure set up. The numerical study they carry out compares one factor Vasicek and CIR models for LIBOR-in-Arrears(LIA) and CMS convexity adjustments. The authors observe for example that higher speed of mean reversion lowers the convexity correction and recommend that if one believes that interest rates exhibit mean reversion then this should be included in the convexity adjustment alongside the volatility of the short rate. In fact, it is not the effect of the mean-reversion on the dynamics of the short-rate that matters, but rather its effect on the joint distribution of all the swap rates at time T. A higher mean-reversion yields lower swap rate volatilities for long dated swaps relative to short-dated swaps at any fixed time T.

It is informative to re-examine the findings through the insights available from taking a Markov-functional approach. We could view the short rate as the driver of a one factor Markov-functional model. If were to do a Markov-functional sweep on an affine term structure model as our prior model in order to calibrate the model to the market-implied distributions of relevant swap rates at time T, the effect of the choice of driver, Vasicek or CIR, would totally disappear- there is a unique one factor Markov-functional model that calibrates to these market prices. The comparison between Vasicek and the CIR models made in Murgoci and Gaspar (2016) is a comparison between two one factor models with different distributions for the swap rates at time T in their respective swaption measures. In that sense it is not the properties of the driver that matter but the resulting distributions it prescribes at time T.

2.1.2 Constructing the functional forms

In Section 2.1, we gave a brief description of the model setup, but we did not discuss how we choose to calibrate the model to the market information. In this section, we describe how the assumptions given above can be used to construct a Markov-functional model at a single time under the forward measure $\mathbb{F}$, that is calibrated to a given set of swaption prices, driven by a bivariate random variable, whose joint distribution is determined by a copula.

We recall we have a finite tenor structure $T < S_1 < S_2 < \ldots < S_{\tilde{M}} < \infty$. Fix $j \in \{1, \ldots, \tilde{M}\}$. Assume we have derived functional forms of y_T^k, for $k \in \{1, \ldots, j-1\}$. Equivalently, from equation (2.4), we know P_T^k, and from equation (2.5), we know D_{TS_k}, for $k \in \{1, \ldots, j-1\}$. We now need to construct the functional form of y_T^j. To do so, we use the alternative prices defined in Proposition 2.1.1.

Proposition 2.1.1. *Given a suitable set of payer swaption prices, with underlying rate y_T^j and strike $K \in \mathbb{R}_+$, denoted by $V_0^j(K)$ and digital swaption prices with same underlying and strike, denoted by $\overline{V}_0^j(K)$, we define the alternate prices $\tilde{V}_0^j(K)$:*

$$\tilde{V}_0^j(K) := \alpha_j V_0^j(K) + (1 + \alpha_j K)\overline{V}_0^j(K).$$

We can express the decreasing càdlàg function $\tilde{V}_0^j(\cdot)$ as an expectation with respect to the swaption measure $\mathbb{S}^j$ as follows:

$$\tilde{V}_0^j(K) = P_0^j \mathbb{E}_{\mathbb{S}^j}[\mathbb{1}(y_T^j > K)(1 + \alpha_j y_T^j)]. \tag{2.7}$$

Proof.

$$\begin{aligned}
\tilde{V}_0^j(K) &:= \alpha_j V_0^j(K) + (1 + \alpha_j K)\overline{V}_0^j(K) \\
&= \alpha_j P_0^j \mathbb{E}_{\mathbb{S}^j}[(y_T^j - K)_+] + (1 + \alpha_j K) P_0^j \mathbb{E}_{\mathbb{S}^j}[\mathbb{1}(y_T^j > K)] \\
&= P_0^j \mathbb{E}_{\mathbb{S}^j}[\alpha_j (y_T^j - K)\mathbb{1}(y_T^j > K) + (1 + \alpha_j K)\mathbb{1}(y_T^j > K)] \\
&= P_0^j \mathbb{E}_{\mathbb{S}^j}[(1 + \alpha_j y_T^j)\mathbb{1}(y_T^j > K)].
\end{aligned}$$

□

Remark 7: Note that $\tilde{V}_0^j(K)$ is the value at time zero of a portfolio comprising α_j units of a payer's swaption with strike K and $(1+\alpha_j K)$ units of a digital swaption with strike K.

The alternative prices $\tilde{V}_0^j(\cdot)$ enable us to construct the functional form f^j in a recursive manner. Indeed, from equation (2.7), by a change of numéraire to $\mathbb{F}$, we get

$$\begin{aligned}
\tilde{V}_0^j(K) &= P_0^j \mathbb{E}_{\mathbb{F}}\left[\left.\frac{\mathrm{d}\mathbb{S}^j}{\mathrm{d}\mathbb{F}}\right|_{\mathcal{F}_T} \mathbb{1}(y_T^j > K)(1 + \alpha_j y_T^j) \right] \\
&= P_0^j \mathbb{E}_{\mathbb{F}}\left[\frac{P_T^j}{D_{TT}} . \frac{D_{0T}}{P_0^j} \mathbb{1}(y_T^j > K)(1 + \alpha_j y_T^j) \right] \\
&= D_{0T} \mathbb{E}_{\mathbb{F}}[\mathbb{1}(y_T^j > K) P_T^j (1 + \alpha_j y_T^j)].
\end{aligned} \tag{2.8}$$

We do not know the functional form of P_T^j yet, but we can omit this term using the fact that:

$$P_T^j(1+\alpha_j y_T^j) = P_T^{j-1} + \alpha_j. \tag{2.9}$$

This follows from rearranging equation (2.3). Hence, using the expression from equation (2.9) in equation (2.8), we have

$$\tilde{V}_0^j(K) = D_{0T}\mathbb{E}_{\mathbb{F}}[\mathbb{1}(y_T^j > K)(P_T^{j-1} + \alpha_j)]. \tag{2.10}$$

We recall the role of the prior model in assumption **A2**: we stated that the prior model informs the choice to collapse the dimension of the driver to a univariate random variable.

Let $y^* \in \mathbb{R}$ and we define $J_0^j(y^*)$ as

$$J_0^j(y^*) := D_{0T}\mathbb{E}_{\mathbb{F}}[\mathbb{1}(\hat{y}_T^j(\mathbf{x}_T) > y^*)(P_T^{j-1} + \alpha_j)],$$

where $\hat{y}_T^j : \mathbb{R}^2 \to \mathbb{R}$. Since P_T^{j-1} is already determined in the previous step, J_0^j is well defined, and can be computed numerically given the distribution of $\mathbf{x}_T$. We observe that $\tilde{V}_0^j$ and J_0^j differ only in the indicator function. By the monotonicity assumption of f^j in **A2**, we can find a unique $K^* \in \mathbb{R}$, such that the set identity holds:

$$\{\hat{y}_T^j > y^*\} = \{y_T^j > K^*\}.$$

For a given y^*, finding that unique K^* such that the above set identity holds is equivalent to knowing $f^j(y^*)$ in the sense that:

$$f^j(y^*) = K^* := \sup\{K \geq 0 : \tilde{V}_0^j(K) \geq J_0^j(y^*)\}. \tag{2.11}$$

Following the steps above allows us to construct a functional form for y_T^j that gives us the desired marginal distribution under the swaption measure $\mathbb{S}^j$. We subsequently know the functional forms of P_T^j and D_{TS_j}.

2.1.3 Specifying the distribution of the driver

We recall that the joint distribution of the components of the driver is a modelling choice. We can choose the marginal distribution of the components of the driver, and specify their joint distribution via a copula function. To do so, we need *Sklar's Theorem*:

Theorem 2.1.2 (Sklar's Theorem)**.** *Consider a d- dimensional cumulative distribution function F with marginals $F_1, \ldots, F_d$. There exists a copula C, such that*

$$F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)), \forall x_i \in \mathbb{R}, i = 1, \ldots, d, \tag{2.12}$$

where $F_i(x_i) \sim U(0,1)$. If F_i are continuous $\forall i = 1, \dots, d$, then C is unique. Conversely, consider a copula C and univariate CDFs $F_1, \dots, F_d$. Then F as defined in (2.12) is a multivariate CDF with marginals $F_1, \dots, F_d$.

We now consider the case where the prior $\hat{y}^j_T$ only depends on $\mathbf{x}_T$ through a linear combination of the components, $\beta^{(1)}_j x^{(1)}_T + \beta^{(2)}_j x^{(2)}_T$. This is sufficient for the cases examined in this thesis. We can now derive a more explicit form for J^j_0.

Building on Theorem 2.1.2, we transform $x^{(i)}_T$ as follows: Let

$$\begin{aligned} \hat{X} &:= F_1(x^{(1)}_T) \\ \hat{Y} &:= F_2(x^{(2)}_T), \end{aligned}$$

where $F_1(\cdot)$ and $F_2(\cdot)$ are the marginal distributions of $x^{(1)}_T$ and $x^{(2)}_T$ respectively (we suppress the time notation T for simplicity). Let $\tilde{P}^{j-1}_T(\mathbf{x}_T) := P^{j-1}_T(\mathbf{x}_T) + \alpha_j$. Assuming $\beta^{(2)}_j > 0$, we have that

$$\begin{aligned} &\mathbb{E}_{\mathbb{F}}[\mathbb{1}(\hat{y}^j_T(\mathbf{x}_T) > y^*)\tilde{P}^{j-1}_T(\mathbf{x}_T)] \\ &\quad = \mathbb{E}_{\mathbb{F}}[\mathbb{1}(\beta^{(1)}_j x^{(1)}_T + \beta^{(2)}_j x^{(2)}_T > y^*)\tilde{P}^{j-1}_T(x^{(1)}_T, x^{(2)}_T)] \\ &\quad = \mathbb{E}_{\mathbb{F}}\Big[\mathbb{E}_{\mathbb{F}}[\mathbb{1}(\beta^{(1)}_j x^{(1)}_T + \beta^{(2)}_j x^{(2)}_T > y^*)\tilde{P}^{j-1}_T(x^{(1)}_T, x^{(2)}_T)\big|\hat{X}]\Big] \\ &\quad = \mathbb{E}_{\mathbb{F}}\left[\mathbb{E}_{\mathbb{F}}\left[\mathbb{1}\left(x^{(2)}_T > \frac{y^* - \beta^{(1)}_j F^{-1}_1(\hat{X})}{\beta^{(2)}_j}\right)\tilde{P}^{j-1}_T(F^{-1}_1(\hat{X}), F^{-1}_2(\hat{Y}))\big|\hat{X}\right]\right] \\ &\quad = \mathbb{E}_{\mathbb{F}}\left[\mathbb{E}_{\mathbb{F}}\left[\mathbb{1}\left(\hat{Y} > F_2\left(\frac{y^* - \beta^{(1)}_j F^{-1}_1(\hat{X})}{\beta^{(2)}_j}\right)\right)\tilde{P}^{j-1}_T(F^{-1}_1(\hat{X}), F^{-1}_2(\hat{Y}))\big|\hat{X}\right]\right]. \end{aligned} \tag{2.13}$$

We look at the inner expectation defined by

$$h^j(y^*, \hat{X}) := \mathbb{E}_{\mathbb{F}}\left[\mathbb{1}\left(\hat{Y} > F_2\left(\frac{y^* - \beta^{(1)}_j F^{-1}_1(\hat{X})}{\beta^{(2)}_j}\right)\right)\tilde{P}^{j-1}_T(F^{-1}_1(\hat{X}), F^{-1}_2(\hat{Y}))\big|\hat{X}\right].$$

We further define for $\hat{x} \in (0,1)$, the function $g^j : \mathbb{R}^2 \to [0,1]$,

$$g^j(y^*, \hat{x}) = F_2\left(\frac{y^* - \beta^{(1)}_j F^{-1}_1(\hat{x})}{\beta^{(2)}_j}\right).$$

We can then evaluate

$$h^j(y^*, \hat{x}) = \int_{g^j(y^*,\hat{x})}^1 \tilde{P}^{j-1}_T(F^{-1}_1(\hat{x}), F^{-1}_2(\hat{y}))c_{\hat{Y}|\hat{X}}(\hat{y}|\hat{x})\mathrm{d}\hat{y},$$

where $c_{\hat{Y}|\hat{X}}(.|.)$ is the chosen conditional copula density. Once we get hold of $h^j(y^*, \hat{x})$, plugging it back in equation (2.13), we can evaluate $J_0^j(y^*)$ as follows:

$$\begin{aligned} J_0^j(y^*) &= D_{0T} \int_0^1 h^j(y^*, \hat{x}) d\hat{x} \\ &= D_{0T} \int_0^1 \int_{g^j(y^*,\hat{x})}^1 \tilde{P}_T^{j-1}(F_1^{-1}(\hat{x}), F_2^{-1}(\hat{y})) c_{\hat{Y},\hat{X}}(\hat{y}, \hat{x}) \mathrm{d}\hat{y}\mathrm{d}\hat{x}. \end{aligned} \tag{2.14}$$

Remark 8: The above equation follows from the assumption that the parameter $\beta_j^{(2)}$ is strictly positive. However, $\beta_j^{(2)} \in \mathbb{R}\backslash\{0\}$ and suitable adjustment has to be made, in particular, reversing the inequality sign if $\beta_j^{(2)} < 0$, and consequently the order of the inner integration.

2.2 General prior model setup

A key feature of a Markov-functional model is its ability to separate the driver and the marginal distributions we want to capture. As touched upon in Section 2.1.1, given the model driver, and a prior model, the functional forms of the swap rates are uniquely determined by the swaption prices when the functional fitting starts at the first maturity date S_1. The ability to choose the functional forms allows us to capture the desired market-implied marginal distributions. The joint distribution of the swap rates is determined by the choice of prior model. Since the specification of the joint distribution of $\mathbf{x}_T$ via a copula and the prior model are both modelling choices, the Markov-functional approach provides considerable flexibility in specifying the joint distribution. As such, the model setup described is very general. If we believe that the state of the economy at a single time can be summarised by a bivariate random variable, all models for which this holds true can be represented by the Markov-functional approach. We would be able to set up a prior model to capture any desired joint distribution of the swap rates and carry out a Markov-functional sweep to calibrate to the market-implied marginal distributions. The last step depends on the monotonicity assumption we made in **A2**, which in essence is not a restrictive assumption. We discuss below a prior model setup based on a local volatility separable swap market model as a concrete example. While this choice is broad in itself, we point out that we could have chosen any low-dimensional model as our base model for the prior model setup.

2.2.1 Example: A prior model setup for local volatility separable market models

We now elaborate on the prior model setup. We can make an informed choice by exploiting the link between Markov-functional models and *separable* market models. The *separability* concept was introduced by Pietersz, Pelsser, and Van Regenmortel (2004) as a method to approximate high-dimensional LIBOR market models (LMM) by a low-dimensional Markov process. Bennett and Kennedy (2005) showed that under the separability condition, a one-dimensional LMM and a one-factor Markov-functional model are numerically indistinguishable. The authors believe that similar observations would hold for swap based market models, and the results would carry over to higher factor models. This allows us to borrow the structure of an SDE formulation of swap rates and carry over the intuition to the Markov-functional framework via the prior model.

We consider a two-factor separable local volatility swap market model. For any given $j \in \{1, \ldots, \tilde{M}\}$, when modelling the swap rate y^j, we assume that the volatility function can be expressed as the product of a bounded, time-dependent deterministic function denoted by $\bar{\sigma}^{i,j}(t)$, for $i \in \{1, 2\}$, and a time-homogeneous (possibly, non-linear) function of the swap rate, denoted by $\phi(\cdot)$. Furthermore, by the separability condition, we can express $\bar{\sigma}^{i,j}(t)$ as the product:

$$\bar{\sigma}^{i,j}(t) = \beta_j^{(i)} \sigma_t^{(i)},$$

where $\sigma_t^{(i)} : [0, T] \to \mathbb{R}$ and $\beta_j^{(i)} \in \mathbb{R}$. Under a general EMM $(N, \mathbb{N})$, the swap rate y^j satisfies an SDE of form:

$$\mathrm{d}y_t^j = \phi(y_t^j)(\beta_j^{(1)} \sigma_t^{(1)} \mathrm{d}W_t^1 + \beta_j^{(2)} \sigma_t^{(2)} \mathrm{d}W_t^2) + \mu_t^j \mathrm{d}t, \tag{2.15}$$

where $\phi : \mathbb{R} \to \mathbb{R}$, and for $i = \{1, 2\}$, $\beta_j^{(i)} \in \mathbb{R}$, $\sigma_t^{(i)}$ is a deterministic function of time and μ_t^j is determined by the no-arbitrage condition. Andersen and Andreasen (2000) showed that under certain conditions on the local volatility function ϕ, the SDE admits a unique non-negative solution, provided $y_0^j \geq 0$, with the solutions being strictly positive if $y_0^j > 0$, for $j \in \{1, \ldots, \tilde{M}\}$.

Some common choices of local volatility functions include $\phi(x) = x - \theta, \theta \geq 0$ which leads to the displaced-diffusion (shifted log-normal) swap market model, and for $\theta = 0$, we obtain the log-normal swap market model. We also have the Constant Elasticity of Variance (CEV) local volatility function with $\phi(x) = x^\gamma$, $\gamma \in (0, 1)$, discussed in Andersen and Andreasen (2000).

We aim to set up a prior model for the Markov-functional approach when mirroring a local volatility swap market model. Note that the marginals given by the local volatility model should be the same marginals to which we calibrate the Markov-functional model. By Itô's lemma, assuming the function $h : \mathbb{R} \to \mathbb{R}$ is twice continuously differentiable, we have that:

$$\begin{aligned}
\mathrm{d}h(y_t^j) &= \frac{\partial h}{\partial y} \mathrm{d}y_t^j + \frac{1}{2} \frac{\partial^2 h}{\partial y^2} \mathrm{d}[y^j]_t \\
&= \frac{\partial h}{\partial y} (\phi(y_t^j)(\beta_j^{(1)} \sigma_t^{(1)} \mathrm{d}W_t^1 + \beta_j^{(2)} \sigma_t^{(2)} \mathrm{d}W_t^2) + c_j(t) \mathrm{d}t,
\end{aligned}$$

where $c_j(\cdot)$ is the finite variation process that groups together the drift and the quadratic variation component. For our purposes, we can ignore this term since any arbitrage introduced by taking an approximation for this term will be removed by the Markov-functional sweep. We note however, in general, if time t is large, significant arbitrage could be introduced in the drift approximation model, and can be enlarged if we are considering a full-term structure approach to price a path-dependent product, as highlighted in Bennett and Kennedy (2005). However, we are not considering this case here.

Choosing the function h such that it satisfies:

$$h(x) = \int \frac{1}{\phi(x)} \mathrm{d}x,$$

we obtain:

$$\mathrm{d}h(y_t^j) = \beta_j^{(1)} \sigma_t^{(1)} \mathrm{d}W_t^1 + \beta_j^{(2)} \sigma_t^{(2)} \mathrm{d}W_t^2 + c_j(t)\mathrm{d}t.$$

We note that if $\phi(x)$ is chosen to be strictly positive, then, by definition h will be a monotonic increasing function. For a given time t, we therefore have that:

$$h(y_t^j) = \beta_j^{(1)} \int_0^t \sigma_s^{(1)} \mathrm{d}W_s^1 + \beta_j^{(2)} \int_0^t \sigma_s^{(2)} \mathrm{d}W_s^2 + \int_0^t c_j(s)\mathrm{d}s.$$

If we apply the common technique of freezing the drift term to its initial value, denoting the constant by $\tilde{c}_j(0)$, and we express the components of the driver of the full term-structure Markov-functional model as $x_t^{(1)} := \int_0^t \sigma_1(s) dW_s^1$ and $x_t^{(2)} := \int_0^t \sigma_2(s) dW_s^2$, we obtain:

$$h(y_t^j) \approx \beta_j^{(1)} x_t^{(1)} + \beta_j^{(2)} x_t^{(2)} + \tilde{c}_j(0),$$

or equivalently, we have:

$$y_t^j \approx h^{-1}(\beta_j^{(1)} x_t^{(1)} + \beta_j^{(2)} x_t^{(2)} + \tilde{c}_j(0)). \tag{2.16}$$

By monotonicity of h, we know the inverse function exists. Fixing the time $t = T$, we observe that we can mirror the local volatility separable market model in the Markov-functional approach and this is achieved by taking a linear combination. We can thus define

$$\hat{y}_T^j(\mathbf{x}_T) := \beta_j^{(1)} x_T^{(1)} + \beta_j^{(2)} x_T^{(2)}.$$

Alternatively, we could have taken the prior model to be the RHS expression of equation (2.16), whereby we fix the time $t = T$ and take a monotonic transformation of the linear combination of the components of the driver. This follows from the fact that the functional fitting is unique, as discussed in assumption **A2**. It does not matter whether the functional fitting is done to $\hat{y}_T^j$ or a monotonic transformation of $\hat{y}_T^j$; they would result in the same functional forms, calibrated to the same set of marginal distributions.

We take as a motivating example, a *modified* version of a swap market model with the local volatility function given by:

$$\phi(y_t^j) = (y_t^j)^\gamma,$$

where γ is a positive constant. This is the CEV model introduced by Cox and Ross (1976). We note that $\gamma = 0$ corresponds to the Bachelier model, giving us Gaussian marginals on the swap rates, and $\gamma = 1$, corresponds to the swap market model, with log-normal marginals. Allowing for general value of γ enables us to incorporate a volatility skew in the model (hence assuming fatter tails in the marginal distributions of the forward rates relative to the Gaussian assumption). Choosing $h(x) := x^{1-\gamma}$, we have that:

$$\begin{aligned} y_T^j &\approx \left[(1-\gamma)\left(\beta_j^{(1)} \int_0^T \sigma_1(t)\mathrm{d}W_t^1 + \beta_j^{(2)} \int_0^T \sigma_2(t)\mathrm{d}W_t^2\right) + \tilde{c}_j(0)\right]^{\frac{1}{1-\gamma}} \\ &= \left[\beta_j^{(1)} x_T^{(1)} + \beta_j^{(2)} x_T^{(2)} + \tilde{c}_j(0)\right]^{\frac{1}{1-\gamma}}, \end{aligned}$$

where $x_T^{(i)} := (1-\gamma)\int_0^T \sigma_i(t)\mathrm{d}W_t^i$, for $i = 1, 2$. We obtain a prior model expressed as a function of a linear combination of the components of the driver.

The Markov-functional approach described in this chapter is general enough to represent a large set of interest rate models for which the state of the world can be summarised by a two-dimensional driver. In the above example, we have shown how the Markov-functional model could be set up if we assume that the distribution of the swap rates exhibit volatility skews. Even if we were to consider stochastic volatilities, with the underlying prior model being a SABR-type model, the Markov-functional approach proposed here would still be applicable. In a one-factor SABR-type swap market model, (y, σ) is Markovian. However, we recall we are setting the model up at a fixed time. By assumption **A3**, if we are given a set of suitable swaption prices obtained from SABR marginals, we can perform a Markov-functional sweep and capture the swap rate marginals. We would only need to set up a prior model for the swap rates to model their joint distribution. In that sense, the contribution of the volatility disappears - this has already been taken into account when calibrating the model. On the other hand, the setup would be more involved for applications where one needs a full term-structure model. In this case, we refer to the interested reader to Kennedy, Mitra, and Pham (2012) Kaisajuntti and Kennedy (2014), and Guo (2016) for the development of a full term-structure stochastic Markov-functional model.

CHAPTER 3

What matters when pricing a CMS and CMS Option: A numerical investigation

In this chapter, we explore numerically via the single-time MFM developed in Chapter 2 the properties that matter most when pricing convexity-related exotic products. We will focus mainly on the pricing of a CMS (and related options), but we point out that the model could be used to price and study other swap-based products (even extending beyond single-rate products to include for example spread options). We expand briefly on what we mean by the properties we want to investigate. Observe firstly that if we want to price a CMS payment with underlying forward swap rate y_T^N and payment made at S_M, we would set up a single-time MFM up to $S_{\tilde{M}}$, $\tilde{M} = \max\{N, M\}$. The model will be calibrated to the full set of market-implied marginal distributions of the swap rates $y_T^i, i = 1, \ldots, \tilde{M}$. The first question therefore is which of these market-implied distributions matter most when pricing a CMS. We study this question using the single-time MFM in the one-factor context. We explain the choice of a one-factor model for the basis of this investigation in Section 3.1. In Section 3.2, we provide some numerical results that allow us to determine the impact that each forward swap rate has on convexity corrections, and how the behaviour changes as the shape of the volatility smile changes. We consider three shapes: flat, skew and smile. We observe that at maturities that are of primary importance to practitioners ($M = 0, 1$ or $M \gg N$), the marginal distribution of the reference swap rate and the marginal distribution of the swap rate whose end date coincide with the payment date (payment swap rate) have the greatest impact on convexity corrections. We equally observe that as the payment date gets further away from the reference swap rate maturity, the effect of the payment swap rate outweighs the effect of all the other rates. Finally, we see that the other forward swap rates have a non-negligible effect on convexity corrections as well, but we point out that the investigation is carried out under fairly stressed market conditions. Nonetheless, the observation that the payment swap rate becomes a significant variable in the valuation remains a salient feature that will persist even under normal market conditions when $M > N$.

Having established the above, we are now interested in the effect of the joint distribution of the reference and payment swap rates on convexity corrections. We study this problem using a two-factor (2F) single-time MFM. For the numerical analysis, we assume the model driver is bivariate Gaussian (under the forward measure), and the swap rates are log-normally distributed in their own swaption measures. Their joint dependence can thus be summarised in terms of the correlation. In Section 3.3.1, we show how to set up the prior model for the 2F single-time MFM based on the assumptions stated above. These assumptions enable us to borrow the structure and information from market models, and mirror the intuition and understanding in the single-time MFM. In Sections 3.3.3 and 3.3.4, we set out to study the effect that the correlation has on convexity corrections. We observe that the correlation between the reference swap rate and the payment swap rate does have a significant effect (especially for long payment dates) on convexity corrections when pricing a CMS and a CMS caplet. Note that if we are working with a copula other than Gaussian, properties of the joint distributions other than correlation might matter, but we do not consider this case in the thesis.

3.1 Establishing the context for a numerical investigation

We use the single time MFM developed in the previous chapter as a platform to study convexity corrections. The numerical analysis carried out in this chapter focuses mainly on the single payment of a Constant Maturity Swap (CMS). We start off by analysing a single payment of the *exotic leg* of a CMS. Let y_T^N be a given forward swap rate with start date $T \equiv S_0$ and $S_1 < S_2 < \ldots < S_N$ be the sequence of the reset dates of the corresponding swap. A payment of the CMS is for the amount y_T^N and is made at some fixed time $S_M \geq T$. We therefore refer to y_T^N as the *reference* swap rate. The value of the payoff at time T is given by:

$$V_T^{CMS} = y_T^N D_{TS_M}.$$

It follows, by Theorem 1.5.1, that the value at time zero is given by:

$$V_0^{CMS} = N_0 \mathbb{E}_{\mathbb{N}}\left[y_T^N \frac{D_{TS_M}}{N_T} \right], \tag{3.1}$$

for some numéraire pair $(N, \mathbb{N})$. Using $D_{\cdot S_M}$ as numéraire with associated forward measure $\mathbb{N}^M$, we have that:

$$V_0^{CMS} = D_{0S_M}\mathbb{E}_{\mathbb{N}^M}[y_T^N].$$

However, the process y^N is not a martingale under $\mathbb{N}^M$, but is so under its associated swaption measure $\mathbb{S}^N$. This gives rise to the need to *adjust* the forward swap rate to account for the unnatural payment schedule. This adjustment, as we have discussed before is referred to as the convexity correction and is defined as follows:

$$\begin{aligned} \mathcal{C}_{N,M} &:= \mathbb{E}_{\mathbb{N}^M}[y_T^N] - \mathbb{E}_{\mathbb{S}^N}[y_T^N] \\ &= \mathbb{E}_{\mathbb{N}^M}[y_T^N] - y_0^N, \end{aligned} \tag{3.2}$$

where the second equality follows as a consequence of the martingale property of y^N under $\mathbb{S}^N$.

It would be natural to consider the valuation of the CMS payment with respect to the swaption measure. From a suitable set of swaption prices over a range of strikes, which we can obtain from the market, at a given fixed time T, one can recover the implied marginal distribution of y_T^N under $\mathbb{S}^N$. Viewed from the swaption measure, the valuation of the payoff requires the knowledge of the joint law of y_T^N and $\frac{D_{TS_M}}{P_T^N}$, which has to be modelled. We shall discuss shortly the approach commonly taken in this case. In this chapter, we choose to work with the pure discount bond maturing at time T as numéraire, with associated forward measure $\mathbb{F}$ and we therefore have:

$$V_0^{CMS} = D_{0T}\mathbb{E}_{\mathbb{F}}[y_T^N D_{TS_M}].$$

Following equation (3.2), by a change of measure to $\mathbb{F}$, the convexity correction is given by:

$$\mathcal{C}_{N,M} = \frac{D_{0T}}{D_{0S_M}}\mathbb{E}_{\mathbb{F}}[y_T^N D_{TS_M}] - y_0^N. \tag{3.3}$$

We observe that for the pricing of the CMS payoff, we would need to consider the joint distribution of the reference swap rate and the pure discount bond maturing at the payment date under the forward measure $\mathbb{F}$. As we have seen in the earlier chapter, D_{TS_M} is itself a function of a set of swap rates $(y_T^1, y_T^2, \ldots, y_T^M)$. We recall:

$$D_{TS_M}(y_T^1, y_T^2, \ldots, y_T^M) := 1 - y_T^M \sum_{k=1}^{M} \alpha_k \prod_{l=k}^{M} \frac{1}{1 + \alpha_l y_T^l}.$$

Hence, in principle, in order to price a CMS, one would need to find an appropriate model for the joint distribution of the swap rates $y_T^1, \ldots, y_T^M$ and the reference swap rate y_T^N, which turns out to be a non-trivial task. A natural question at this stage is to determine what properties of the joint distribution have a significant impact on the valuation problem; in other words, what is important to get right? In the current practical literature, emphasis is placed mainly on getting the marginal distribution of the reference swap rate under its associated swaption measure accurately modelled. This is a reasonable approach in the particular context where the payment date lies within the length of the reference swap rate maturity. Any model setup in this case has to satisfy a consistency condition specified below, that puts a fairly strong restriction on the model. We consider the valuation at a set of payment dates coinciding with the payment dates of the reference swap rate. Taking a weighted sum of the expected payoffs, we have that:

$$\begin{aligned}\sum_{i=1}^{N} \alpha_i \mathbb{E}_{\mathbb{F}}[y_T^N D_{TS_i}] &= \mathbb{E}_{\mathbb{F}}[y_T^N P_T^N] \\ &= \frac{1 - D_{0S_N}}{D_{0T}}.\end{aligned}$$

By definition of convexity correction given in equation (3.3), it follows that:

$$\begin{aligned}\sum_{i=1}^{N} \alpha_i D_{0S_i}(\mathcal{C}_{N,i} + y_0^N) &= D_{0T} \sum_{i=1}^{N} \alpha_i \mathbb{E}_{\mathbb{F}}[y_T^N D_{TS_i}] \\ &= 1 - D_{0S_N} \\ \implies \sum_{i=1}^{N} \alpha_i D_{0S_i}\mathcal{C}_{N,i} + y_0^N P_0^N &= 1 - D_{0S_N} \\ \implies \sum_{i=1}^{N} \alpha_i D_{0S_i}\mathcal{C}_{N,i} &= 0. \qquad (3.4)\end{aligned}$$

We observe that if we knew the convexity correction at the setting time T (which we do in the European markets as it can be derived from market prices of cash-settled swaptions), the convexity correction at the reference swap maturity, and we assume that the convexity corrections behave roughly linearly with respect to the payment dates, we could estimate the convexity correction at the intermediate times $S_i, i \in \{1, \ldots, N-1\}$. We do not know the convexity correction $\mathcal{C}_{N,N}$ but the consistency condition would give us a weighted average and this gives us some control for when the payment date is earlier than the reference swap rate maturity.

However, recent products, typically of spread type have pushed the payment dates beyond the reference swap rate maturity, a setup not currently explored or discussed in the literature, except incidentally in Cedervall and Piterbarg (2012) who consider the possibility in their numerical analysis, but do not expand on the idea, nor explore the context in depth. When the payment date goes beyond the length of the reference swap rate, we lose control over convexity corrections given by the consistency condition. Viewed from the forward measure, we would want the model to take into account the decorrelation between the pure discount bond D_{TS_M} and the reference swap rate y_T^N. We would intuitively expect the correlation to be large and negative at first, but to increase as the payment date increases beyond the swap rate maturity. So in this case, it is not only the level of rates that matter, but we would also want to capture the slope, hence motivating the need for a two-factor model.

In light of the above, in this chapter, we set out to study what matters most in the pricing problem. The general question we want to address via the single-time MFM is: What matters when it comes to effectively pricing a CMS? We break down the general question into two specific questions that can be numerically investigated:

(a) When pricing a CMS, which marginal distributions of the swap rates under their respective swaption measures have a significant impact?
We study the first question using a 1F single-time MFM. There is a unique one-factor model that would calibrate to the whole set of marginal distributions of the swap rates under their respective swaption measures. As we shall demonstrate in the numerical sections that follow, we observe from the single-time one-factor MFM that, as the payment date is taken further away from the reference swap rate maturity, the marginal distribution of payment swap rate, together with that of the reference swap rate, have a significant impact on convexity corrections, while the other rates have comparatively much weaker effect.

We note that we could equally have used the two-factor version of the model to study this question. We chose not to do so, firstly because a two-factor model would be cumbersome in the specific context we are in; it would be harder to dissociate the marginal distribution effect of the swap rates from their joint dependence. Secondly, we do not expect the conclusions we draw, particularly with regard to those rates which have little impact, to alter in a two-factor setting. Neither the payoff $y_T^N D_{TS_M}$ nor its distribution under the forward measure $\mathbb{F}$ have a strong dependence on rates other than y^N and y^M. (Note that for any n, the Radon-Nikodỳm derivative $\frac{\mathrm{d}\mathbb{S}^n}{\mathrm{d}\mathbb{F}}\big|_{\mathcal{F}_T} = P_T^n$ depends predominantly on y_T^n, but not on other rates y_T^m, $m \neq n$. Thus the rates that we observe to be insignificant for the accurate pricing of a CMS in a one-factor model, will remain so in a two-factor context.

(b) Which aspects of the joint distribution of the swap rates matter when valuing convexity corrections? We investigate this using a 2F single-time MFM. A first element that plays a significant role in establishing the joint dependence between the swap rates is the PVBP. We recall briefly the model setup discussed in Section 2.1.1 and illustrated in Figure 2.1 of chapter 2. At any fixed maturity S_j, the calibration step involves the PVBP P_T^{j-1} which is constructed using market information of all the swap rates from previous maturities. Hence, the unfolding of the model forward in maturity results in the PVBP being a complicated function of the swap rates, and this information is reflected in the construction of the functional form of y_T^j. We shall explore in more detail in Chapter 5, the importance of modelling the PVBP under the forward measure appropriately, as this variable fixes up a lot of structure in the joint dependence between the swap rates, a point we will come to later in this section when discussing the approach taken by Cedervall and Piterbarg (2012). A second element determining the joint distribution between the swap rates under the forward measure $\mathbb{F}$ is the copula. The Markov-functional modelling approach allows us to choose a copula function governing the joint distribution of the components of the driver. Once we choose the copula, along with the market-implied marginal distributions of the swap rates, we have specified a joint distribution for the swap rates under the forward measure. The modelling approach further enables us to vary the copula without losing the ability to calibrate the model. So, a sub-question here is: What is the effect of the joint distribution of the driver on the valuation problem? We expect that the copula would not have a significant effect in the case of a CMS, but would be significant in a CMS spread option. To motivate this argument, we point out that the CMS is a single-rate product, but as we will learn from the investigation in the one-factor case, we need to factor in the market-implied marginal distribution of the payment swap rate, on top of that of the reference swap rate for the appropriate valuation of the expected payoff. Assume we start off by evaluating the expected payoff in a two-factor setting using a Gaussian copula on the driver. If we were to change the copula, but keeping the correlation between the swap rates the same, we would expect the two model results to not be significantly different from each other. On the other hand, spread-based products, whose payoff are a function of more than one swap rate, might be more sensitive to the choice of copula.

The rest of this section is devoted to an in-depth analysis of some current approaches for valuation of a CMS. The approaches we discuss here are relevant and important from a practitioner's point of view. The models we shall explore are set up with the view that we are interested in modelling swap rates at a single time, and they are calibrated to the relevant market information. As such, we are not focused on a full term-structure setting, as we have seen in some approaches discussed in the Introduction and Chapter 2. Term-structure models have limited flexibility and control over the model at a single time, a key drawback when pricing European derivatives.

We begin with the Terminal Swap Rate (TSR) models commonly used in practice. This modelling approach was developed by Hunt and Kennedy (2000) and is typically used to model and price European interest rate derivatives. The idea behind the approach is that any pure discount bond of a given maturity S, observed at a fixed time T and the numéraire N_T (often defined as a function of pure discount bonds) can be reasonably approximated as a function of a single variable most closely related to the product we are interested in pricing. In the case of a CMS, the natural choice is the reference swap rate. A TSR model therefore makes the assumption:

$$\begin{aligned} D_{TS} &= D_{TS}(y_T^N), \qquad S \geq T, \\ N_T &= N_T(y_T^N), \end{aligned}$$

where $D_{TS}(\cdot)$ and $N_T(\cdot)$ is a collection of exogenous, pre-determined functions of the reference swap rate. For instance, we could assume that D_{TS} rebased by the numéraire P^N is linear in y_T^N, giving us the *linear swap rate model*:

$$\frac{D_{TS}}{P_T^N}(y_T^N) := A + B_S y_T^N,$$

where $A, B_S \in \mathbb{R}$. Other classes of terminal swap rate models include the *exponential swap rate model*, whereby the pure discount bonds are modelled as decaying exponential functions and the *geometric swap rate model.* In a full term-structure setting, the relationship between the pure discount bond and the swap rate will be induced by the model itself. Since the functional forms in a terminal swap rate model are determined exogenously, some conditions have to be placed on the functional forms to ensure the model is consistent, arbitrage-free and realistic. We refer the interested reader to Hunt and Kennedy (2000) for a detailed overview.

Going back to the valuation problem, from equation (3.1), if we work under the swaption measure corresponding to taking P^N as numéraire, we have that:

$$V_0^{\text{CMS}} = P_0^N \mathbb{E}_{\mathbb{S}^N}\left[y_T^N \frac{D_{TS_M}}{P_T^N}\right].$$

By the Tower property of expectations, Andersen and Piterbarg (2010b) observe that the valuation can be expressed as follows:

$$\begin{aligned} V_0^{\text{CMS}} &= P_0^N \mathbb{E}_{\mathbb{S}^N}\left[y_T^N \frac{D_{TS_M}}{P_T^N}\right] \\ &= P_0^N \mathbb{E}_{\mathbb{S}^N}\left[\mathbb{E}_{\mathbb{S}^N}\left[y_T^N \frac{D_{TS_M}}{P_T^N}\middle| y_T^N\right]\right] \\ &= P_0^N \mathbb{E}_{\mathbb{S}^N}[y_T^N \alpha_{M,N}(y_T^N)], \end{aligned}$$

where

$$\alpha_{M,N}(y_T^N) := \mathbb{E}_{\mathbb{S}^N}\left[\frac{D_{TS_M}}{P_T^N}\middle| y_T^N\right]$$

is referred to as the *annuity mapping function*.

Once we get hold of the annuity mapping function, which could be approximated using a TSR model, the valuation problem boils down to an expectation of some function of the underlying swap rate under its associated swaption measure (They use the *replication* argument to evaluate the expectation, by breaking down the payoff function into a weighted sums of standard European swaptions). Hence, the authors believe that the market-implied marginal distribution of the reference swap rate is the key information they aim to get right in their approach to price a CMS. They also make a connection between the model parameters and the mean-reversion parameter of a full-term structure model (Gaussian one-factor model), pointing to the belief that the model evolution at earlier times influences the choices at a single fixed time. This misses the point that for European derivatives, we aim to appropriately model the underlying rates at a single fixed time, and our beliefs on model dynamics at earlier times do not affect the choices we make at the fixed time.

Building on the above, Cedervall and Piterbarg (2012) extend the approach to account for the fact that the payoff of a CMS depends not only on the reference swap rate; the authors argue that for accurate pricing and risk management, it is the joint distribution of the reference swap rate, the associated PVBP and the discount bond maturing at the payment date that matters. To account for this, they introduce a different approximation for the annuity function as follows:

$$\alpha_{\tilde{M},N}(y_T^N) := f(Y_{1,N}, Y_{2,N}, \ldots, Y_{\tilde{M},N}),$$

where

$$Y_{j,N}(x) := \mathbb{E}_{\mathbb{F}}[y_T^j | y_T^N = x], \tag{3.5}$$

for $j \in \{1, \ldots, \tilde{M}\}$, $\tilde{M} = \max\{N, M\}$, and $f : \mathbb{R}^{\tilde{M}} \to \mathbb{R}$ is some deterministic function. It is this representation of the annuity function that allows them to incorporate correlation and volatilities of all the swap rates into the payoff function. However, as we can immediately observe from equation (3.5), evaluating the annuity function requires knowledge of the joint distribution of the swap rates under the forward measure. The authors provide an approximation that is derived from the joint Gaussian distribution assumption on the (log of) swap rates: They assume that the swap rates are Normal under their respective swaption measures, but can be reasonably approximated as jointly shifted log-normal variables under $\mathbb{F}$ with some drift terms that are fixed using no-arbitrage conditions. The volatility of the swap rates under $\mathbb{F}$ is approximated using the market-implied volatility known only under the swaption measure. The swap rate approximations under the forward measure are used as building blocks for the PVBP, the variable that enables them to move the valuation problem back to the swaption measure. As we shall explore later on in Chapter 5, considerable care has to be taken when we set a model up with respect to a measure under which we have no information about the marginal distribution of the swap rates. The authors do not consider how sensitive their models are to a change in the parameters used to approximate the swap rates under the forward measure, that would have a direct impact on the PVBP. Furthermore, the assumptions rely heavily on Gaussian assumptions and do not provide flexibility for varying the model choices without significantly altering the model structure. While the authors argue that their approach accounts for the correlation between all the swap rates, the authors single out the correlation between the reference swap rate and the payment swap rate when considering the CMS convexity corrections for which the payment date is taken far away from the reference swap rate maturity, but they do not elaborate on this choice. In that sense, they recognise the influence of the payment swap rate on the valuation of a CMS. However, they do not have control over the marginal distribution of the payment swap rate in their approach. They are mainly focused on setting up a single-time model aimed at modelling the reference swap rate directly, with the goal of pricing single rate derivatives.

In a more recent paper, Bermin and Williams (2017) explored the use of a Markov-functional technique to price cash-settled swaptions and CMS. More precisely, the authors set up a full-term structure model of choice (the authors opt for a Quadratic Gaussian Model (QGM)) and perform a *Markov-functional sweep* on the functional form of the reference swap rate, in order to match the model-induced distribution to the market-implied distribution at a given fixed time T. The authors take the view that the market-implied marginal distribution of the reference swap rate is relevant and significant to the accurate pricing of a CMS. In so doing, they perturb the model by modifying the functional form of the reference swap rate but they keep the functional forms of the PVBP and pure discount bonds the same. In light of the known functional relationship between the swap rate and pure discount bonds, it is straightforward to argue that the technique has inadvertently introduced arbitrage in the model. The only criteria mentioned in the paper is that the model must be arbitrage free in the sense that the fixed and floating leg of a swap are priced correctly. A second potential issue relates to the choice of the base (full term-structure) model. The QGM is chosen as it highlights relevant modeling aspects whilst still being tractable. How the choice of the starting model ultimately affects CMS pricing is however not approached.

The literature studied above allows us to appreciate the complexity of the valuation problem of a CMS, or any CMS-related derivatives, and points out the modelling aspects that one might need to consider in order to correctly price these derivatives. Considerable effort goes into capturing the market-implied distribution of the reference swap rate, a reasonable approach when the payment date varies within the length of the reference swap rate maturity. In this chapter, we explore the impact of relaxing this payment date assumption and determine the features of the joint distribution of the swap rates that matter when pricing a CMS (and related products).

3.2 Numerical Investigation: The one-factor case

As we have discussed in the previous section, we use a one factor single-time MFM to investigate the impact of the market-implied marginal distributions of the swap rates on convexity corrections arising from pricing a single payment of a CMS. We choose to work under the forward measure $\mathbb{F}$ corresponding to taking $D_{\cdot T}$ as numéraire. The questions we are interested in are, (i) whether all market-implied marginal distributions of the swap rates contribute equally to the convexity corrections and (ii) will the behaviour change if the shape of the volatility smile changes. In the next section, we lay out the single-time one factor MFM setup.

3.2.1 Single-time one-factor MFM setup and initial conditions

We use the same tenor structure setup as in Section 2.1.1. We assume that the setting date T is 20 years from now, and the *maximum* payment time $S_{\tilde{M}}$ is 30 years from T. The accrual factor, $\alpha_i = 1$, for $i \in \{1, \dots, \tilde{M}\}$. In order to study the payment time effect on convexity corrections arising from the pricing of a CMS based on the forward rate y_T^N (the valuation of which is given in equation (3.3)), we allow the payment date S_M, for $M \in \{0, \dots, \tilde{M}\}$, to vary within the given tenor up to $S_{\tilde{M}}$ and we define y_T^M as the *payment swap rate* - i.e the forward swap rate with maturity coinciding with the payment date. We consider two cases, one where we take $S_N = 2$ years from T and $S_N = 10$ years from T. The initial conditions are given in table 1 in Appendix A. We assume that the model driver is a standard Gaussian random variable under the forward measure. As discussed before, we could have made any other choice for the driver since once we calibrate the model, the effect of the driver would disappear, owing to the uniqueness property of the one-factor MFM. We chose a Gaussian variable for ease of numerical implementation.

3.2.2 Log-normal market-implied distributions

We begin the investigation with the assumption that the swap rates are log-normally distributed under their respective swaption measures, with $y_0^i = 0.07$, for all i (we can ignore the subscript for ease of notation) and log-normal volatilities σ_i, $i \in \{1, \dots, \tilde{M}\}$ as given in Table 1 in Appendix A. We shall explore in detail how we choose to set these values up when we discuss the two-factor version of the model. We first compute and plot the convexity corrections as a function of payment date below.

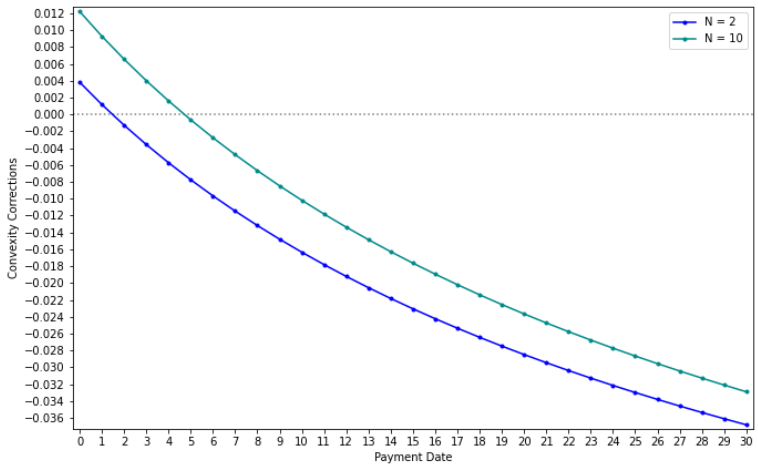

Figure 3.1: Convexity Corrections against payment date under log-normal market-implied distributions
$\tilde{M} = 30, D_{0T} = 1.0, y_0 = 0.07, \sigma_2 = 0.149, \sigma_{10} = 0.143$

In order to investigate the marginal distributions that have a significant impact in determining the value of convexity corrections, the vega profile is computed as detailed below.

We first fix the reference swap rate maturity S_N. For $i \in \{1, \ldots, 30\}$, the i^{th} log-normal volatility is increased by 1 basis point (bp) (i.e a percentage point increase of 0.01%) and the convexity correction at a given payment date S_M, for $M \in \{0, \ldots, \tilde{M}\}$, denoted by $(\mathcal{C}_{N,M})^{i+}$ is computed. Similarly, the i^{th} log-normal volatility is decreased by 1 bp and the convexity correction $(\mathcal{C}_{N,M})^{i-}$ is computed. The i^{th} vega associated with payment date S_M, denoted by $v_{M,i}$ is calculated as follows:

$$v_{M,i} = \frac{(\mathcal{C}_{N,M})^{i+} - (\mathcal{C}_{N,M})^{i-}}{10^{-4} \text{ x } 2}.$$

Remark 9: The vega also depends on the reference swap rate maturity S_N, but we have suppressed the dependence on S_N for ease of notation.

Figures 3.2 and 3.3 below illustrate the vega profile. On the x-axis is the payment date and on the y-axis are the vega values $v_{(.,.)}$ for a fixed reference swap rate maturity. At any given payment date, each coordinate represents the i^{th} vega corresponding to the log-normal volatility of y_T^i for $i \in \{1, \ldots, 30\}$.

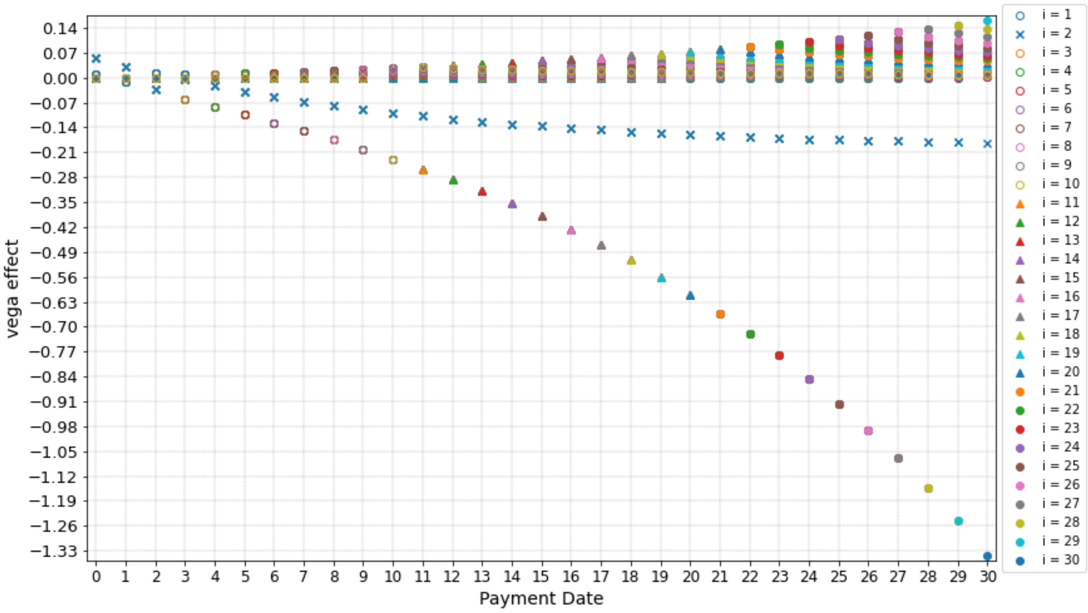

Figure 3.2: Plot illustrating the vega profile
$N = 2, \tilde{M} = 30, D_{0T} = 1.0, y_0 = 0.07, \sigma_2 = 0.149$

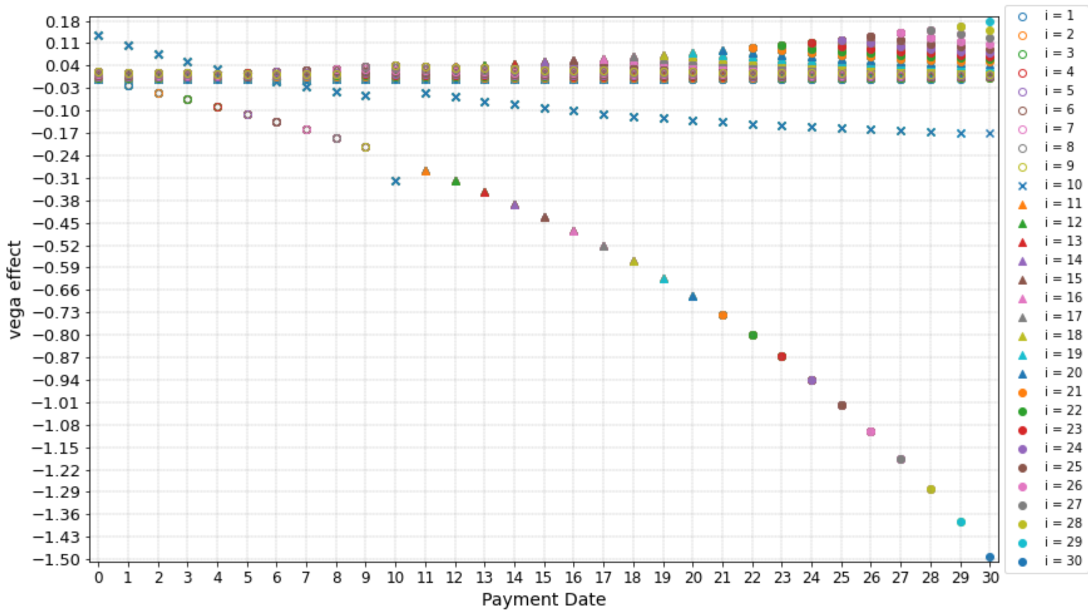

Figure 3.3: Plot illustrating the vega profile
$N = 10, \tilde{M} = 30, D_{0T} = 1.0, y_0 = 0.07, \sigma_{10} = 0.143$

We note the following:

The N^{th} vega, $v_{M,N}$ (the vega associated with the log-normal volatility of the reference swap rate; illustrated as crosses in the plots) starts off positive and decreases to some negative value as S_M increases.

The vega $v_{M,M}$, for $M \in \{1, \ldots, 30\}$, is always negative and decreases faster than

$v_{M,N}$ to some negative value bigger in magnitude than $v_{30,N}$.

We equally observe that the vega at the payment date coinciding with the reference swap rate maturity registers a much bigger effect. This is so because in this case, since $y_T^M \equiv y_T^N$, the change in log-normal volatility of the reference swap rate is complemented by the log-normal volatility effect from the pure discount bond D_{TS_N}(itself being a function of y_T^N).

Finally, for large payment dates, the vega associated with the i^{th} volatility for $i \in \{1, \dots \tilde{M}\} \backslash \{N, M\}$ is comparable to that of the N^{th} vega. We should point out however that the vega is computed under fairly stressed market conditions. Nonetheless, we still expect $v_{M,N}$ and $v_{M,M}$ to be significant under stable market conditions.

We particularly focus on small payment dates ($M = 0, 1$) and $M = 30$. We note that when $M = 0$ or 1, the effect from the reference swap rate dominates all the other effects. When $M = 30$, the vega associated with the payment swap rate far outweighs that associated with the reference swap rate. This shows that in the log-normal case, together with the marginal distribution of the reference swap rate, the marginal distribution of the payment swap rate in its own swaption measure, has a significant effect on the valuation of a CMS and its associated convexity correction. Whether this observation would hold if we relax the log-normal assumption on the marginals can be investigated if we move away from a flat implied volatility to a skew, generated by assuming that the density function of y_T^i, for $i \in \{1, \dots, \tilde{M}\}$ is given by a mixture of normal and log-normal distributions, which we shall study in the next section.

3.2.3 Skew effect on convexity corrections

We move away from log-normal assumptions on the market-implied marginal distributions of the swap rates and allow for a skew in the shape of the implied volatility curve. We do so by assuming that for $i \in \{1, .., \tilde{M}\}$, the marginal distribution of y_T^i under its associated swaption measure $\mathbb{S}^i$ is given by a mixture of normal and log-normal distributions.

Let $X_1 \sim \mathcal{N}(y_0^i, \eta_i^2)$ and denote the density function of X_1 as f_1
$\log(X_2) \sim \mathcal{N}(\log y_0^i - \frac{1}{2}\sigma_i^2 T, \sigma_i^2 T)$, with density function given by f_2.
The density function of y_T^i, denoted by f_y^i is then given by:

$$f_y^i(x) = (1 - \gamma^i) f_1(x) + \gamma^i f_2(x), \quad 0 \leq \gamma^i \leq 1. \tag{3.6}$$

The parameter σ_i is the log-normal volatility and the parameter η_i is determined by matching ATM swaption prices under the Gaussian assumption to that of the log-normal case, the computation of which is given in the appendix B.

We assume for all i, the weight parameter is constant, so $\gamma^i = \gamma$. We illustrate the implied volatility curve for y_T^N, $N = 2, 10$ below and understand the impact that γ has on the shape of the implied volatility curve:

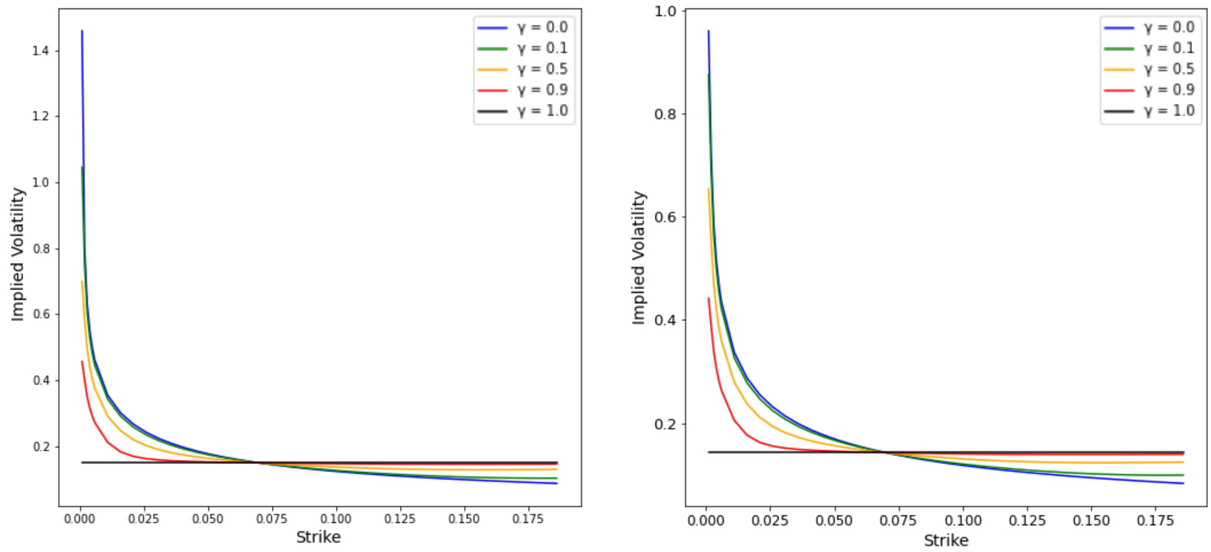

Figure 3.4: Implied volatility given γ
Left - $\sigma_2 = 0.149$; Right - $\sigma_{10} = 0.143$

For any given $\gamma \in [0, 1)$, we denote the implied volatility as σ_{imp}^{γ}, the log-normal volatility as σ_{LN} and the ATM strike as K_{ATM}. We observe the following from Figure 3.4,

For $K \leq K_{ATM}$,

For any fixed $\gamma \in [0, 1), \sigma_{imp}^{\gamma} > \sigma_{LN}$

For any fixed K, $\gamma_1 > \gamma_2 \implies \sigma_{imp}^{\gamma_1} < \sigma_{imp}^{\gamma_2}$

For $K > K_{ATM}$,

For any fixed $\gamma \in [0, 1), \sigma_{imp}^{\gamma} < \sigma_{LN}$

For any fixed K, $\gamma_1 > \gamma_2 \implies \sigma_{imp}^{\gamma_1} > \sigma_{imp}^{\gamma_2}$

We now explore the skew effect on convexity corrections. We consider 3 values : $\gamma = 0.1, 0.5$, and 0.9. We only include the plots when $\gamma = 0.5$ here. The results for the two other cases can be found in the appendix B.1. We also provide some further detail to explain the behaviour of the skew effect.

For a given $i \in \{1, \ldots, 30\}$, the parameter γ is increased (decreased) by 1 bp and the convexity correction, denoted by $(\mathcal{C}_{N,M})^{\gamma^i+}((\mathcal{C}_{N,M})^{\gamma^i-})$ is computed. The change in convexity correction with respect to a change in γ^i, denoted by $\Gamma_{M,i}$ is calculated as follows:

$$\Gamma_{M,i} := \frac{(\mathcal{C}_{N,M})^{\gamma^i+} - (\mathcal{C}_{N,M})^{\gamma^i-}}{10^{-4} \times 2}.$$

The plots below illustrate the effect of γ on convexity corrections. Similar patterns can be observed for both reference swap rate maturities considered, whereby the changes in the marginals of the reference swap rate and that of the payment swap rate via the weight γ are more pronounced than the rest.

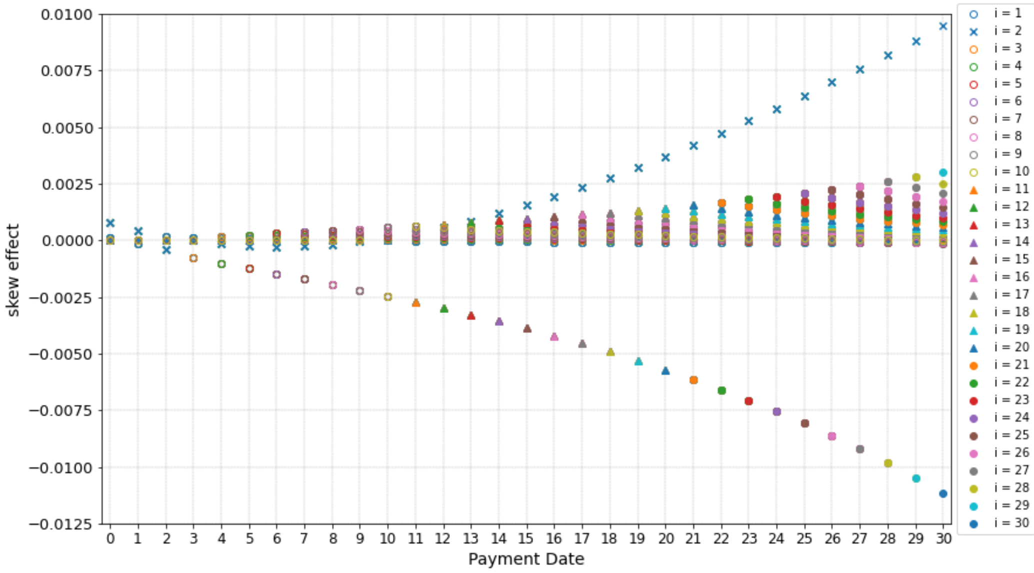

Figure 3.5: Skew effect
$\gamma = 0.5, N = 2, \tilde{M} = 30, y_0 = 0.07, D_{0T} = 1.0$

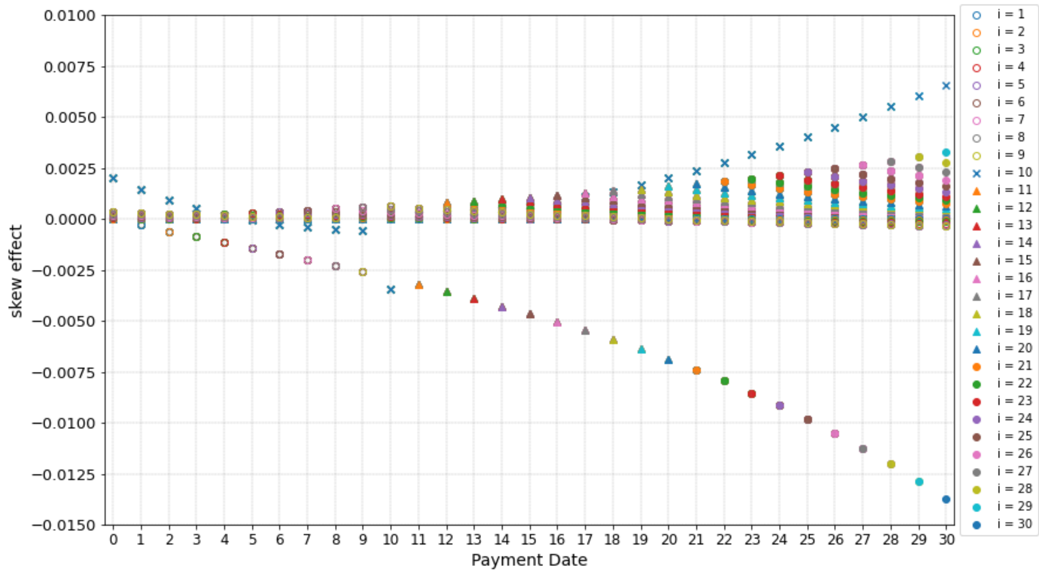

Figure 3.6: Skew effect
$\gamma = 0.5, N = 10, \tilde{M} = 30, y_0 = 0.07, D_{0T} = 1.0$

The payment dates that we are most interested in and that are of practical importance, are $M \in \{0, 1\}$ (M small) and $M = 30$. In both Figures 3.5 and 3.6, we observe that for $M \in \{0, 1\}$, $\Gamma_{M,N}$ registers the biggest effect, dominating all the other skew effects. When $M = 30$, we note that the skew effect from the payment swap rate, $\Gamma_{M,M}$ registers the biggest effect, outweighing that of the reference swap rate, which itself is more significant than $\Gamma_{M,i}$, for $i \in \{1, \ldots, \tilde{M} - 1\} \backslash \{N\}$.

In line with the effects plotted in Figures 3.5 and 3.6 we can argue that if γ^N is increased, keeping all other weights constant, the convexity correction would decrease in magnitude as the payment date increases. On the other hand, if γ^M is increased, keeping all other weights constant, the convexity corrections would increase in magnitude as payment time increases. We provide a quantitative argument to emphasize the point. We focus on the case $N = 10$ and $M = 30$. If we assume that the convexity correction behaves linearly as a function of γ (this is only a rough approximation, but we do not expect it to be too far off since $\Gamma_{30,30}$ is approximately the same value for all three cases of γ values considered when $N = 10$), we would have that:

$$\begin{aligned} \mathcal{C}_{10,30}(\gamma) &= \gamma \frac{\mathrm{d}\mathcal{C}_{10,30}}{\mathrm{d}\gamma} + c \\ &= \gamma \Gamma_{30,30} + c. \end{aligned}$$

It follow that $\mathcal{C}_{10,30}(0) = c$ and $\mathcal{C}_{10,30}(1) = \Gamma_{30,30} + c$. Going back to Figure 3.1, in the log-normal case, the convexity correction at payment date S_{30} when $N = 10$ is about -0.034. The analysis above tells us that a slight deviation from the log-normal distribution assumption on the payment swap rate will change the convexity correction by roughly 1.4%, which is very significant. Given that a change in the market-implied distribution of the payment swap rate via γ outweighs that of the reference swap rate as the payment date is pushed away from the reference swap rate maturity, a model that fails to account for the marginal distribution of the payment swap rate would lead to a consistent overestimation of its convexity corrections. We can push this experiment further to understand in more depth the implied volatility effect by moving from a skew to a smile. We impose a smile free from static-arbitrage on the distributions of the swap rates under their respective swaption measure using the method proposed by Gatheral and Jacquier (2014).

3.2.4 Smile effect on convexity corrections

For a fixed setting date T, in line with the definition of implied volatility given by Gatheral and Jacquier (2014), for $i \in \{1, \dots, \tilde{M}\}$, we define the following:

- strike, denoted by $K = y_0^i \exp(\hat{K})$, for $\hat{K} \in \mathbb{R}$
- $\sigma_{BS}^i(\hat{K}, T)$ is the implied volatility of y_T^i for a given $\hat{K} \in \mathbb{R}$ and
- The '*total implied variance*' is denoted by $\omega^i(\hat{K}, T)$ and is defined by $(\sigma_{BS}^i(\hat{K}, T))^2 T$

Remark 10*:*

(a) The 'total implied variance' $\omega^i(\hat{K}, T)$ refers to the variance of the log of the swap rate y_T^i, if we were to assume y_T^i is log-normally distributed under its associated swaption measure $\mathbb{S}^i$

(b) Borrowing the terms from Gatheral and Jacquier (2014), the map $(\hat{K}, t) \to \omega^i(\hat{K}, t)$, for $t > 0$, is referred to as the volatility surface. Since we are only interested in fitting a smile at a given fixed time T, we only focus on the map $\hat{K} \to \omega^i(\hat{K})$, representing a slice of the volatility surface. So we drop the dependence on T in the notation. We also fit the same smile to all the forward swap rates, hence we suppress the index i for ease of notation.

The Surface Stochastic Volatility Inspired (SSVI) is a class of parameterised volatility surfaces that allows one to fit a smile that can be arbitrage-free by imposing some conditions on the parameters. Using the SSVI parameterisation, we define the total implied variance as follows:

$$\omega(\hat{K}) := \frac{\theta}{2}\left[1 + \rho\phi(\theta)\hat{K} + \sqrt{(\phi(\theta)\hat{K} + \rho)^2 + (1 - \rho^2)}\right], \tag{3.7}$$

where θ is defined as the ATM implied total variance i.e $\theta = \omega(0)$ and $\phi(\theta) : \mathbb{R}_+ \to \mathbb{R}_+$ is some deterministic function of θ.

In order to investigate the impact of the smile on convexity corrections, we want to gain control over the left and right wings of the curve separately and study the effect of each on convexity corrections to determine which side brings a greater contribution to convexity corrections. To do so, we use the mapping to the SVI-JW parameterisation, argued to be more intuitive to traders than the SSVI parameterisation. The SVI-JW parameterisation is attributed to Tim Klassen by Gatheral and discussed in Gatheral and Jacquier (2014). The mapping to SVI-JW parameters are as follows:

$\nu_T = \frac{\theta}{T}$	Gives the ATM variance
$\psi(T) = \frac{1}{2}\rho\sqrt{\theta}\phi(\theta)$	Gives the ATM skew
$\mathcal{L}_T = \frac{1}{2}\sqrt{\theta}\phi(\theta)(1-\rho)$	Gives the slope of the left wing of smile
$\mathcal{R}_T = \frac{1}{2}\sqrt{\theta}\phi(\theta)(1+\rho)$	Gives the slope of right wing of smile
$\tilde{\nu}_T = \frac{\theta}{T}(1-\rho^2)$	Gives the minimum implied variance

Remark 11*: The parameters ν_T and $\tilde{\nu}_T$ have an explicit dependence on expiry. We can drop the dependence on T for the other parameters.*

Note that there is no one-to-one correspondence between the SVI-JW parameterisation and that of the SSVI. The relationship between the parameters would depend on the properties we want to capture. Since we are interested in gaining control over the left and right wings of the smile, we choose to work with the following correspondence:

$$\begin{aligned}
\theta &= \nu_T.T \\
\phi(\theta) &= \frac{\mathcal{L}+\mathcal{R}}{\sqrt{\nu_T.T}} \\
\rho &= \frac{\mathcal{R}-\mathcal{L}}{\mathcal{R}+\mathcal{L}} \\
1-\rho^2 &= \frac{\tilde{\nu}_T}{\nu_T}
\end{aligned}$$

Plugging the expression on the RHS for each parameter back in the total implied variance given in (3.7), we obtain the same smile as the SSVI, but parameterised differently. We can now investigate the effect of the slope of the smile for smaller strikes by controlling $\mathcal{L}$ and that of larger strikes through $\mathcal{R}$. In order to get a realistic smile curve, we make the following choices of parameters: $\theta = 0.2$ (This sets the ATM implied volatility to 0.1), $\mathcal{L} = 0.502$ and $\mathcal{R} = 0.512$.

Before investigating the effect of the smile on convexity corrections, we look at how the parameters $\mathcal{R}$ and $\mathcal{L}$ control the smile curve.

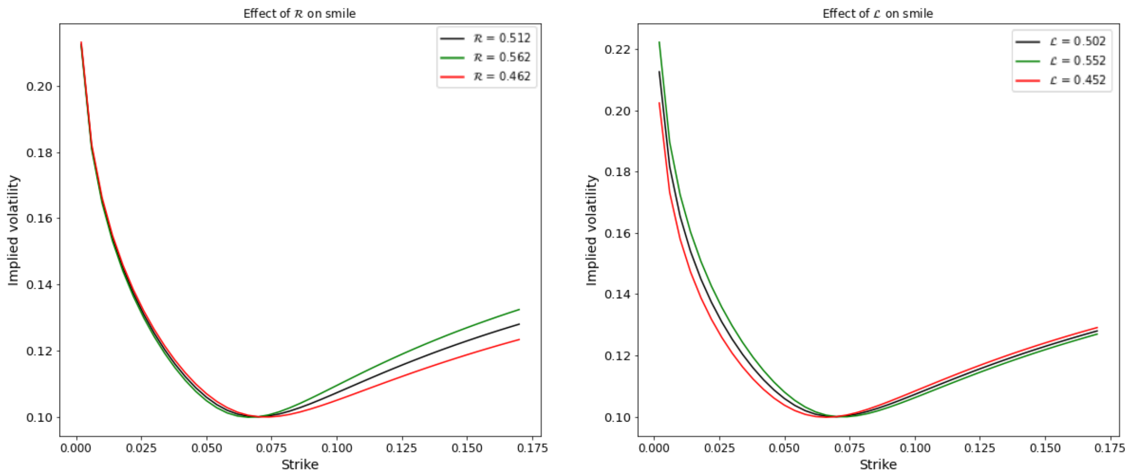

Figure 3.7: Effect of parameters $\mathcal{R}$ and $\mathcal{L}$ on smile
$y_0 = 0.07$, ATM implied volatility = 0.1

We investigate the right-wing effect on convexity corrections. We define:

$$\delta_{\mathcal{R}}(M, i) := \frac{(\mathcal{C}_{N,M})^{\mathcal{R}^{i+}} - (\mathcal{C}_{N,M})^{\mathcal{R}^{i-}}}{10^{-4} \text{ x } 2},$$

where $(\mathcal{C}_{N,M})^{\mathcal{R}^{i+}}$ is the convexity correction at payment date S_M when the smile curve of the swap rate y_T^i, for a given $i \in \{1, \ldots, \tilde{M}\}$ is modified via the parameter $\mathcal{R}$, which is increased by 0.01%, and $(\mathcal{C}_{N,M})^{\mathcal{R}^{i-}}$ is the analogous result, but now we decrease $\mathcal{R}$ by 0.01%.

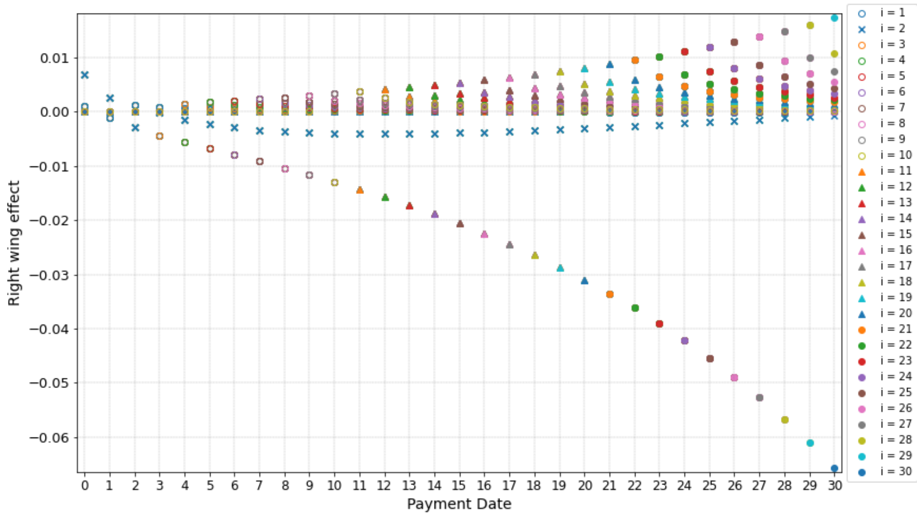

Figure 3.8: Right-wing effect on convexity corrections
$N = 2, \tilde{M} = 30, y_0 = 0.07$, *ATM implied volatility = 0.1*, $\mathcal{L} = 0.502$, $\mathcal{R} = 0.512$

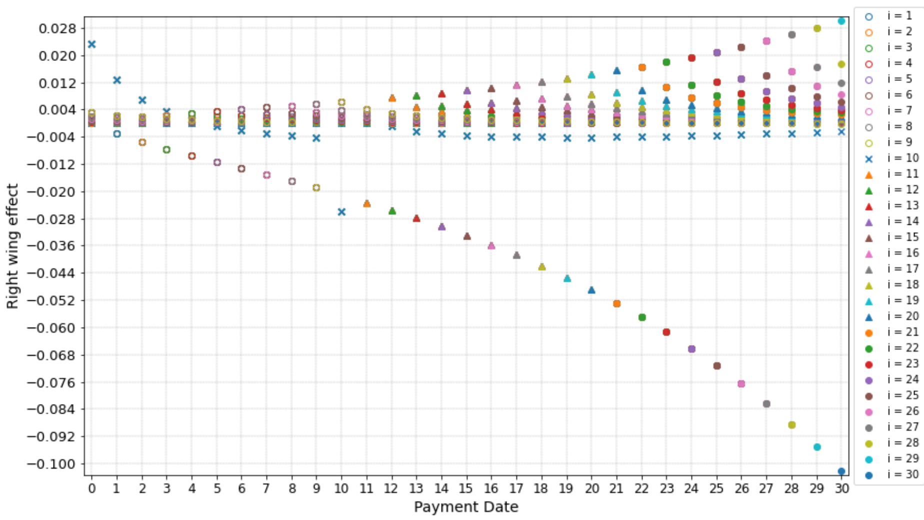

Figure 3.9: Right-wing effect on convexity corrections
$N = 10, \tilde{M} = 30, y_0 = 0.07$, *ATM implied volatility = 0.1*, $\mathcal{L} = 0.502, \mathcal{R} = 0.512$

Figures 3.8 and Figures 3.9 illustrate the change in convexity correction with respect to a change in the right wing of the smile curve. We observe that as the payment date is pushed further away from the reference swap rate maturity, the effect associated with the smile curve of the payments swap rate grows larger and outweighs the effect of all the other swap rates. In Figure 3.7, we can see the effect on the smile for a change in the parameter $\mathcal{R}$ of 0.1. For $M = 30$, from Figure 3.9, we see that the approximate derivative of convexity correction with respect to $\mathcal{R}$ is -0.1, which implies that a change of 0.1 in $\mathcal{R}$ moves the convexity correction by roughly 0.01, i.e 1%. The effect for a corresponding change in the parameter $\mathcal{L}$ is a factor of ten less (see Figure 3.11), although the pattern of the smile of the payment swap rate having the most significant effect as the payment date gets further away remains.

We can conclude from the effect plots, that again, in either case, it is the smile associated with the payment swap rate that has biggest impact on convexity corrections, but the effect is only (potentially) significant for a shift in the right wing of the smile.

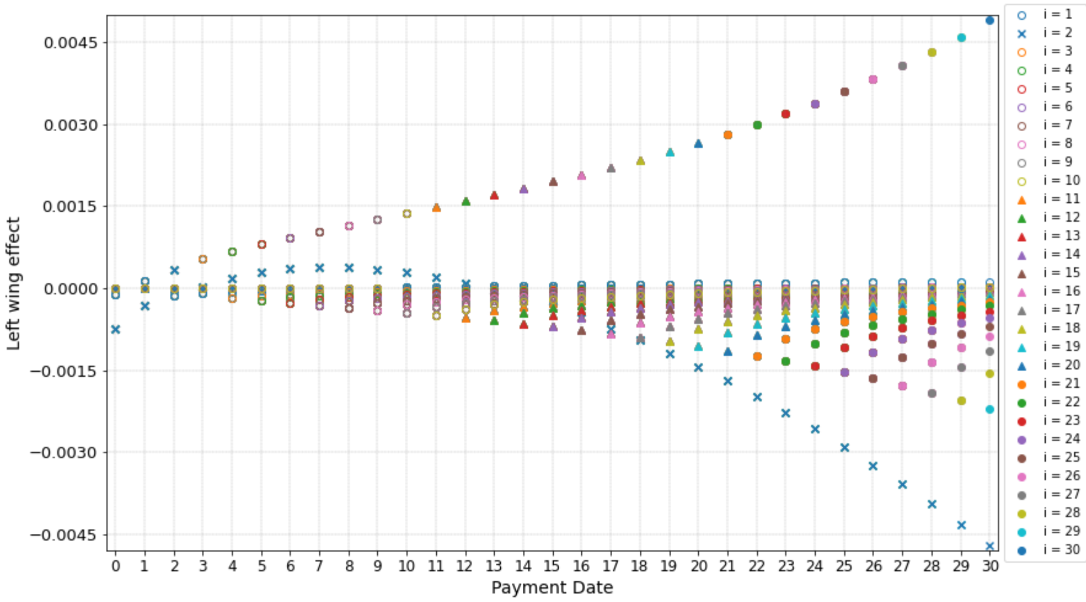

Figure 3.10: Left wing effect on convexity corrections
$N = 2, \tilde{M} = 30, D_{0T} = 1, y_0 = 0.07,$ *ATM implied volatility = 0.1,* $\mathcal{L} = 0.502, \mathcal{R} = 0.512$

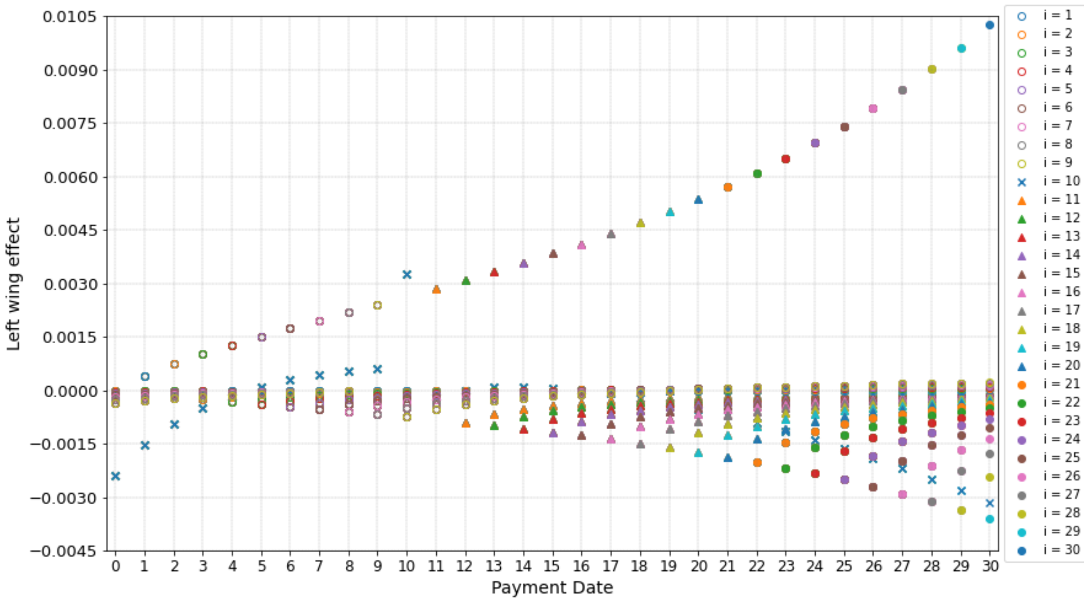

Figure 3.11: Left wing effect on convexity corrections
$N = 10, \tilde{M} = 30, D_{0T} = 1, y_0 = 0.07,$ *ATM implied volatility = 0.1,* $\mathcal{L} = 0.502, \mathcal{R} = 0.512$

We end this analysis with the observation that the investigation was carried out for an ATM implied volatility of 0.1. Working with the smile given by a higher ATM vol of 0.15 illustrates a potential modelling issue which has been pointed out by Pelsser (2000). We elaborate on this: Set $N = 2$. We fit a smile to the swap rate y_T^2 given by the following parameters: $\theta = 0.45$ (giving us an ATM implied volatility of 0.15), $\mathcal{R} = 0.512$ and $\mathcal{L} = 0.502$. Using the single time MFM, we look at $y_T^2(X)$, for $X \in \mathcal{N}(0,1)$ under $\mathbb{F}$. In particular, we plot $\log(y_T^2(x))$ against $x \in \mathbb{R}$.

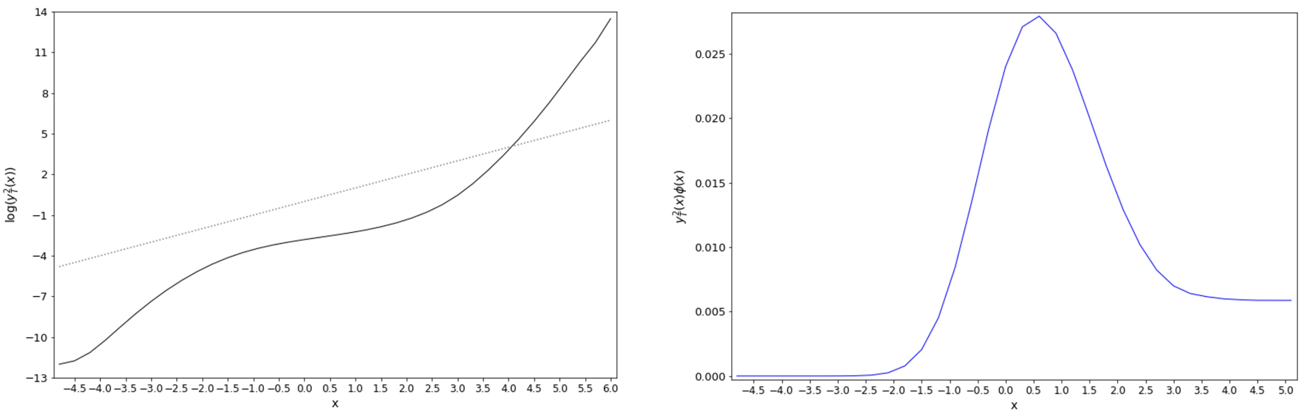

Figure 3.12: Left: Functional form of $\log(y_T^2)(x)$
Right: Integrand function $y_T^2(x)\phi(x)$
$N = 2, D_{0T} = 1, y_0 = 0.07,$ *ATM implied volatility = 0.15,* $\mathcal{L} = 0.502, \mathcal{R} = 0.512$

The straight line in the left plot is used to illustrate by how much we are deviating from the log-normal distribution.

Suppose we are evaluating the convexity correction at time T, i.e $M = 0$. We have that:

$$\mathcal{C}_{2,0} = \mathbb{E}_{\mathbb{F}}[y_T^2] - y_0^2,$$

where

$$\mathbb{E}_{\mathbb{F}}[y_T^2] = \int_{-\infty}^{\infty} y_T^2(x)\phi(x)\mathrm{d}x.$$

From Figure 3.12, we observe that a significant contribution to the integrand and hence the convexity correction is coming from values of x greater than two standard deviations away from the mean. Put differently, the smile is putting too much weight on unrealistic values of the swap rate. In practice, one would not model the tail of the distribution in such a way that either (or both) wings is making a big contribution to the convexity corrections.

3.3 A two-factor single-time Markov-functional model

The numerical investigation in the one-factor case revealed the significance of the market-implied marginal distributions of the reference swap rate and the payment swap rate when pricing a single payment of a CMS. We now aim to answer the second question raised in Section 3.1, namely, what aspects of the joint distribution of the swap rates might one want to take into account? As explained then, there are three components of the model determining the joint dependence of the swap rates: the market-implied marginal distributions, the modelling of the PVBP, and the copula. In the analysis that follows, we assume that the swap rates are log-normally distributed under their respective swaption measures. In the single-time MFM, by unfolding the model forward in maturity, we feed in the market information of the swap rates in the construction of the PVBP using the functional relationship between the PVBP and the swap rates as we saw in equation (2.4). Finally, we assume that the driver is jointly Gaussian under the forward measure $\mathbb{F}$ (hence fitting a Gaussian copula to the components of the driver). We point out that the single-time MFM can be adapted to take into account other choices of marginal distributions on the swap rates, as we have discussed in Chapter 2. The log-normal assumption on the swap rates, together with the Gaussian copula fitted on the components of the model driver allows us to borrow information from already well-established results of market models when constructing the prior model, and carry over the intuition and understanding to the less transparent Markov-functional model. We would expect the same qualitative behaviour if we were to consider different market-implied distributions. We do not expect the copula to have a significant effect in the CMS case, so we stick with a Gaussian copula for ease of numerical implementation. In

Section 2.2, we discussed how to set up a general prior model for the single-time MFM. In the next section, we shall discuss the prior model setup under the specific log-normal assumption. We then provide and discuss the numerical results in Section 3.3.2.

3.3.1 The prior model setup

The common approach used when setting up a Markov-functional model driven by a multi-dimensional driver (in this case, we are taking a two-dimensional driver) is to set up a *prior model* that expresses the forward rate as a function of its components. The role of the prior model is two-fold: firstly, we want to retain the univariate and monotonicity properties of a one-factor model that allows for efficient functional fitting, and secondly, that the covariance structure is realistic and desirable. We define the prior model as follows:

$$\log(\hat{y}^j_T) := \beta^{(1)}_j x^{(1)}_T + \beta^{(2)}_j x^{(2)}_T, \tag{3.8}$$

where $\beta^{(i)}_j \in \mathbb{R}$, for $i \in \{1, 2\}$, $j \in \{1, \ldots, \tilde{M}\}$ and under the forward measure $\mathbb{F}$,

$$\begin{pmatrix} x^{(1)}_T \\ x^{(2)}_T \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\right).$$

Remark 12*: We pointed out in remark 6 in Chapter 2, that we could take a monotonic transformation of the prior model, without affecting the calibration of the model. Here, we are taking a log-transform to linearise the prior model in terms of the components of the driver.*

Once we have chosen the parameters $\beta^{(i)}_j$, under the assumptions we have made on the driver and the marginal distributions of the swap rates being log-normal under their respective swaption measures, the separable swap market model (the benchmark model from which we have formulated the prior model) and the single-time MFM are numerically close to each other, a result following Bennett and Kennedy (2005), as discussed in Chapter 2. However, the prior model admits significant arbitrage, which is removed by performing a *Markov-functional sweep* that acts as a perturbation on the prior model. Given the close numerical relationship to the separable swap market model, we expect the Markov-functional sweep not to have to modify the functional form significantly in order to remove arbitrage. This slight adjustment needed implies that it would be reasonable to approximate the correlation between the log of any two given swap rates under the forward measure as follows:

$$
\begin{aligned}
\operatorname{corr}(\log(y_T^j), \log(y_T^m)) &\approx \frac{\mathbb{E}_{\mathbb{F}}[(\beta_j^{(1)} x_T^{(1)} + \beta_j^{(2)} x_T^{(2)})(\beta_m^{(1)} x_T^{(1)} + \beta_m^{(2)} x_T^{(2)})]}{\sqrt{\operatorname{var}(\log y_T^j)\operatorname{var}(\log y_T^m)}} \\
&= \frac{\beta_j^{(1)}\beta_m^{(1)}\lambda_1 + \beta_j^{(2)}\beta_m^{(2)}\lambda_2}{\sqrt{(\beta_j^{(1)})^2\lambda_1 + (\beta_j^{(2)})^2\lambda_2}\sqrt{(\beta_m^{(1)})^2\lambda_1 + (\beta_m^{(2)})^2\lambda_2}},
\end{aligned}
\tag{3.9}
$$

for $j, m \in \{1, \ldots, \tilde{M}\}$.

We now want to assign a structural meaning to the $\boldsymbol{\beta}^{(i)}$ vectors, for $i \in \{1, 2\}$. For this purpose, we borrow the information from a LIBOR Market model. Extensive research has been carried out in the LMM framework and there are robust methodologies developed for the calibration of models and capturing suitable aspects of market information relevant to pricing specific products. Our aim here is to borrow the existing knowledge about the covariance of the log of the LIBORs from a LMM and reflect it in the prior model setup.

As our starting point, we use the relationship between the swap rate and a set of forward LIBORs. For a given maturity S_j, $j \in \{1, \ldots, \tilde{M}\}$, we have that:

$$
\begin{aligned}
y_T^j &:= \frac{1 - D_{TS_j}}{\sum_{l=1}^j \alpha_l D_{TS_l}} \\
&= \frac{1 - D_{TS_1} + D_{TS_1} - D_{TS_2} + D_{TS_2} - \ldots - D_{TS_j}}{\sum_{l=1}^j \alpha_l D_{TS_l}} \\
&= \frac{\alpha_1 D_{TS_1} L_T^1 + \alpha_2 D_{TS_2} L_T^2 + \ldots + \alpha_j D_{TS_j} L_T^j}{\sum_{l=1}^j \alpha_l D_{TS_l}} \qquad \text{follows from : } L_T^k := \frac{D_{TS_{k-1}} - D_{TS_k}}{\alpha_k D_{TS_k}} \\
&= \sum_{k=1}^j \left(\frac{\alpha_k D_{TS_k}}{\sum_{l=1}^j (\alpha_l D_{TS_l})} \right) L_T^k =: \sum_{k=1}^j \omega_k^j(T) L_T^k.
\end{aligned}
\tag{3.10}
$$

We apply a one-order Taylor expansion to y_T^j about $(L_0^1, L_0^2, \ldots, L_0^j)$ and we obtain:

$$
\begin{aligned}
y_T^j &\approx y_0^j + \sum_{k=1}^j \left(\left(\sum_{p=1}^j L_0^p \left(\frac{\partial \omega_p^j(T)}{\partial L_T^k} \right)_{T=0} (L_T^k - L_0^k) \right) + \omega_k^j(0)(L_T^k - L_0^k) \right) \\
&= y_0^j + \sum_{k=1}^j \left(\left[\omega_k^j(0) + \sum_{p=1}^j L_0^p \left(\frac{\partial \omega_p^j(T)}{\partial L_T^k} \right)_{T=0} \right] (L_T^k - L_0^k) \right) \\
&= y_0^j + \sum_{k=1}^j \tilde{\omega}_k^j(0)(L_T^k - L_0^k),
\end{aligned}
\tag{3.11}
$$

where

$$\tilde{\omega}_k^j(0) = \omega_k^j(0) + \sum_{p=1}^{j} L_0^p \left(\frac{\partial \omega_p^j(T)}{\partial L_T^k} \right)_{T=0}.$$

We observe that a consequence of the Taylor expansion is that the weight $\omega_k^j(\cdot)$ and its partial derivative with respect to the LIBORs are evaluated at their time-zero values. Rebonato (1998) was one of the first to explicitly point out that swap rates could be expressed as a weighted sum of LIBORs. The weights are themselves dependent on the LIBORs, so we do not have a proper weighted average. However, empirical studies have shown that the variability of the weights is small enough compared to that of the LIBORs that the weights and their partial derivatives can be approximated by their known time-zero values. This technique has been extensively applied in the literature, see for example, Rebonato (2002), Brigo and Mercurio (2006) and Andersen and Piterbarg (2010b).

Considering a first-order Taylor expansion of $\log(y_T^j)$ about $\log(y_0^j)$, we have:

$$\log(y_T^j) \approx \log(y_0^j) + \frac{1}{y_0^j}(y_T^j - y_0^j) + \ldots$$

Similarly, applying the above technique to $\log(L_T^k)$ about $\log(L_0^k)$, we have:

$$\log(L_T^k) \approx \log(L_0^k) + \frac{1}{L_0^k}(L_T^k - L_0^k) + \ldots \tag{3.12}$$

Rearranging the two expressions above, and plugging back into equation (3.11), we get:

$$\begin{aligned}\log(y_T^j) &\approx \log(y_0^j) + \sum_{k=1}^{j} \frac{\tilde{\omega}_k^j(0)}{y_0^j} L_0^k(\log(L_T^k) - \log(L_0^k)) \\ &= \log(y_0^j) + \sum_{k=1}^{j} \xi_k^j(0)(\log(L_T^k) - \log(L_0^k)), \end{aligned} \tag{3.13}$$

where

$$\xi_k^j(0) := \frac{\tilde{\omega}_k^j(0) L_0^k}{y_0^j}.$$

Remark 13*: The exact form of the weight $\xi_k^j(0)$ is given in Appendix B.2.*

We observe from equation (3.13), that we have expressed the log of the swap rate as a weighted sum of the log of the LIBORS. At this point, we are able to incorporate information from a LMM within the prior model setup that will determine the correlation structure between the log of the swap rates.

The instantaneous correlation matrix of a LMM can be obtained from market data. A suitable parameterisation for the instantaneous volatility function which reflects market features such as the *'Rebonato hump'* (Rebonato (2005)) can also be chosen and the final calibrated LMM used to provide the covariance matrix $\mathbf{Q}$ for the log of the LIBORs at time T under the forward measure $\mathbb{F}$. We refer to $\mathbf{Q}$ as the integrated covariance matrix.

A Principal Component Analysis (PCA) approach, first applied for the calibration of LMMs by Pedersen (1998), has been extensively used in the literature. An interpretation of the principal components of the *instantaneous* covariance matrix of the log of the LIBORs have been derived and studies have shown they can be described in terms of the *level of rates*, *slope* and *curvature*, with the first two components making up more than 90% of the variability [Choy, Dun, and Schlögl (2003), Rebonato (2005), Brigo and Mercurio (2006) and Lord and Pelsser (2007)]. It is not possible to use market data to study the integrated covariance matrix $\mathbf{Q}$. However using the theory of *Total Positivity*, Lord and Pelsser (2007) make the case that for correlation matrices we would expect to see for the log of the LIBORs at time T, we will still have the interpretation of level, slope and curvature for the first three principal components. We define:

$$\hat{\mathbf{X}}_T := \mathbf{D}^{-1}(\log(L_T^1), \log(L_T^2), \ldots, \log(L_T^{\tilde{M}}))^\top,$$

where the diagonal matrix $\mathbf{D} := \operatorname{diag}\left(\sqrt{\operatorname{var}(\log(L_T^1))}, \sqrt{\operatorname{var}(\log(L_T^2))}, \ldots, \sqrt{\operatorname{var}(\log(L_T^{\tilde{M}}))}\right)$.

Denote by $\hat{\mathbf{Q}}$, the positive-definite, symmetric, $\tilde{M} \times \tilde{M}$ covariance matrix of $\hat{\mathbf{X}}_T$. Note that this is also a correlation matrix. The matrix $\hat{\mathbf{Q}}$ can be therefore be decomposed in terms of its eigenvectors and eigenvalues as follows:

$$\hat{\mathbf{Q}} = \hat{\mathbf{A}}\Lambda\hat{\mathbf{A}}^\top$$

where the columns of orthonormal matrix $\hat{\mathbf{A}}$ correspond to the eigenvectors of $\hat{\mathbf{Q}}$ and Λ is the diagonal matrix of the corresponding eigenvalues $(\lambda_i)_{i=1}^{\tilde{M}}$ arranged in descending order.

The principal component transform of $\hat{\mathbf{X}}_T$ is then given by

$$\hat{\mathbf{Y}}_T := \hat{\mathbf{A}}^\top(\hat{\mathbf{X}}_T - \mathbb{E}_{\mathbb{F}}[\hat{\mathbf{X}}_T]), \tag{3.14}$$

where $\mathbb{E}_{\mathbb{F}}[\hat{\mathbf{X}}_T]$ denotes the $\tilde{M}$-dimensional column vector with entries given by $\mathbb{E}_{\mathbb{F}}\left[\frac{\log(L_T^i)}{d_{ii}}\right]$, where d_{ii} is the i^{th} diagonal entry of $\mathbf{D}$, for $i \in \{1, \ldots, \tilde{M}\}$. We can observe from (3.14), that

$$\begin{aligned}\mathbb{E}_{\mathbb{F}}[\hat{\mathbf{Y}}_T] &= 0\\ \mathrm{cov}_{\mathbb{F}}[\hat{\mathbf{Y}}_T] &= \hat{\mathbf{A}}^\top\hat{\mathbf{Q}}\hat{\mathbf{A}} = \hat{\mathbf{A}}^\top\hat{\mathbf{A}}\Lambda\hat{\mathbf{A}}^\top\hat{\mathbf{A}}.\end{aligned}$$

Rearranging equation (3.14), we obtain

$$\hat{\mathbf{X}}_T = \mathbb{E}_{\mathbb{F}}[\hat{\mathbf{X}}_T] + \hat{\mathbf{A}}\hat{\mathbf{Y}}_T.$$

Equivalently, setting $\mathbf{X}_T = (\log(L_T^1), \ldots, \log(L_T^{\tilde{M}}))^\top$, we have

$$\mathbf{X}_T = \mathbb{E}_{\mathbb{F}}[\mathbf{X}_T] + \mathbf{D}\hat{\mathbf{A}}\hat{\mathbf{Y}}_T. \tag{3.15}$$

Remark 14*: The formal solution at a given time T to the general SDE of the LIBOR L^i, for $i \in \{1, \ldots, \tilde{M}\}$, given by the LMM is formulated as follows:*

$$L_T^i = L_0^i \exp\left(\int_0^T \mu_i(s)\mathrm{d}s + \mathrm{f.v} + \sum_{k=1}^{d}\int_0^T \sigma_i^k(s)\mathrm{d}W_s^k\right),$$

where $d \in \mathbb{N}$ represents the number of factors used to model the LIBOR, $\sigma_i^k : [0, T] \to \mathbb{R}$ is the volatility function (assumed to be deterministic; we can further make the separability assumption on the volatility function as discussed in Section 2.2.1, the 'f.v' term is the finite variation term, a by-product of Itô's formula, and $\mu_i(\cdot)$ is the drift term that is set using the no-arbitrage condition, and is a function of the instantaneous volatilities, the correlation between them and the forward LIBOR rates, making it a stochastic term. In this section we are not after the evaluation of the drift term. We can use the common technique of freezing the LIBORs appearing in the drift term to their time-zero value, reducing the drift to a simple deterministic function $\mu_i^0(\cdot)$, without significant loss of accuracy. From this we can work out that

$$\hat{\mu}_T^i =: \big(\mathbb{E}_{\mathbb{F}}[\mathbf{X}_T]\big)_i \approx \log(L_0^i) + \int_0^T \mu_i^0(s)\mathrm{d}s + \mathrm{f.v.}$$

Going back to equation (3.15), based on the idea that the first two principal components capture most of the variability, we can ignore the other principal components and express:

$$\begin{pmatrix} \log(L_T^1) \\ \log(L_T^2) \\ \cdot \\ \cdot \\ \cdot \\ \log(L_T^{\tilde{M}}) \end{pmatrix} \approx \begin{pmatrix} \hat{\mu}_T^1 + a_{11}Y_1 + a_{12}Y_2 \\ \hat{\mu}_T^2 + a_{21}Y_1 + a_{22}Y_2 \\ \cdot \\ \cdot \\ \cdot \\ \hat{\mu}_T^{\tilde{M}} + a_{\tilde{M}1}Y_1 + a_{\tilde{M}2}Y_2 \end{pmatrix}, \tag{3.16}$$

where $a_{ij} := d_{ii}\hat{a}_{ij}$ for $i \in \{1, \dots, M\}$, $j \in \{1, 2\}$, and Y_1 and Y_2 are independent, normally distributed random variables with $Y_k \sim \mathcal{N}(0, \lambda_k)$ for $k = 1, 2$.

We could choose the first component of the driver $x_T^{(1)}$ to capture the level of rates, and $x_T^{(2)}$ is set to capture the slope. We set the variance of $x_T^{(i)}$ to be λ_i. Note that we will still get the level and slope interpretation, in the sense that the terms d_{ii} are positive and will preserve any sign change in the shape of the eigenvectors $(\hat{a}_{ij})_{i=1}^{\tilde{M}}$, $j \in \{1, 2\}$.

We now want to understand how this structure would carry over to the swap rates. From equation (3.13), we can now work out that:

$$\begin{aligned} \log(y_T^j) &\approx \log(y_0^j) + \sum_{k=1}^{j} \xi_k^j(0)(a_{k1}x_T^{(1)} + a_{k2}x_T^{(2)} + \hat{\mu}_T^k) \\ &= \log(y_0^j) + \beta_j^{(1)}x_T^{(1)} + \beta_j^{(2)}x_T^{(2)} + C, \end{aligned} \tag{3.17}$$

where

$$\beta_j^{(1)} = \sum_{k=1}^{j} \xi_k^j(0)a_{k1} \qquad\qquad \beta_j^{(2)} = \sum_{k=1}^{j} \xi_k^j(0)a_{k2}, \tag{3.18}$$

and C is some constant (we recall that the Markov functional sweep will remove the arbitrage introduced by the drift approximation technique). We observe that we have constructed a prior model for the swap rates as a linear combination of the components of the model driver with a designated definition for the weight parameters $\boldsymbol{\beta}^{(i)}$, as desired.

The prior model setup allows us to control the model correlation (recall we have an approximate analytical formula given in equation (3.9)), an important modelling aspect that enables us to carry out the numerical investigation in the later sections. We have shown here how we chose to set up a two-factor single-time Markov-functional approach for the pricing of a CMS. The log-normal assumption on the distribution of the swap rates under their respective swaption measures, together with the driver chosen to be Gaussian, enabled us to lean on the construction of a separable swap market model to derive a prior model for the single-time MFM. The correlation structure between the log of the swap rates was justified through widely accepted information about the covariance matrix of the log of the LIBORs integrated within the prior model setup. Even though the PCA technique is usually used in the context of instantaneous volatilities and correlation, the broad conclusions carry over to the terminal correlation matrix, justified through Lord and Pelsser (2007).

3.3.2 Numerical Investigation: The two-factor case

Initial conditions

For the numerical analysis, we use the same tenor structure as discussed in the one-factor case, and the same initial conditions as given in Table 1 in Appendix A. For a given $j \in \{1, \ldots, \tilde{M}\}$, for all $k \in \{1, \ldots, j\}$, we assume $L_0^k = 0.07$. Setting $T = 0$ in equation (3.10), we have that:

$$y_0^j = \sum_{k=1}^{j} \omega_k^j(0) L_0^k.$$

We can take the constant LIBOR term out of the summation, and observe that $\sum_{k=1}^{j} \omega_k^j(0) = 1$; hence $y_0^j = 0.07$. We assume that each swap rate y_T^j is log-normally distributed under its associated swaption measure $\mathbb{S}^j$. We use the relationship between the log of the swap rates and the log of the LIBORs, as given in equation (3.13) to determine the log-normal volatility of the swap rates. For this task, we first need to make some assumptions on the LIBORs: we take the caplet implied volatility to be constant and equal to 0.15. We assume the correlation between L_T^i and L_T^j, for $(i, j) \in \{1, \ldots, \tilde{M}\}^2$, is given by $\exp(-0.03|i - j|)$. The decaying exponential coupled with a humped shape is a common parameterisation with some desirable features applied in practice to model LIBOR correlations, and has been discussed in Rebonato (1998) and Brigo and Mercurio (2006). From equation (3.13) we can work out the volatility of the swap rates from the fact that:

$$\mathrm{var}(\log(y_T^j)) := \sigma_j^2 T \approx 0.15^2 T \left(\sum_{k=1}^{j} \left(\xi_k^j(0) \right)^2 + 2 \sum_{k=1}^{j} \sum_{p>k}^{j} \xi_k^j(0) \xi_p^j(0) \exp(-0.03|k - p|) \right). \tag{3.19}$$

From equation (3.19), we can recover the log-normal volatility of the swap rates under their associated swaption measures, the values of which are given in Appendix A.

Remark 15: *The values obtained for the log-normal volatility are the ones used in the one-factor case as well.*

In this study, we are interested in determining the impact of correlation on convexity corrections. We therefore want to get a handle on the correlation structure. Looking closely at the approximate expression given in equation (3.9), we can control the correlation structure via the parameter λ_2. One however needs to be careful when varying λ_2. In particular, we observe that, from the approximate equation obtained for the log of the LIBORs via the PCA in equation (3.16), the correlation between the log of the LIBORs given by: for $j, k \in \{1, \dots, \tilde{M}\}$,

$$\operatorname{corr}(\log(L_T^j), \log(L_T^k)) \approx \frac{a_{j1}a_{k1}\lambda_1 + a_{j2}a_{k2}\lambda_2}{\sqrt{a_{j1}^2\lambda_1 + a_{j2}^2\lambda_2}\sqrt{a_{k1}^2\lambda_1 + a_{k2}^2\lambda_2}}, \tag{3.20}$$

which depends on λ_1 and λ_2. In line with constructing a realistic model, we would want the model to ascribe non-negative correlation to the log of the LIBORs. This therefore gives us an upper bound on λ_2. WLOG, taking $\lambda_1 = 1$, we have that:

$$\operatorname{corr}(\log(L_T^j), \log(L_T^k)) > 0 \implies a_{k1}a_{j1} + a_{j2}a_{k2}\lambda_2 > 0.$$

An upper limit on λ_2 is then given by:

$$\lambda_2 < -\frac{\min_j\{a_{j1}\}\max_k\{a_{k1}\}}{\min_j\{a_{j2}\}\max_k\{a_{k2}\}}.$$

Remark 16*: Note that in order to derive an expression for the variance of the log of the swap rates, we chose a specific parameterisation for the correlation between the log of the LIBORs, which is different from equation* (3.20)*, obtained via the PCA approach.*

Choice of eigenvectors

We chose to set the $\beta^{(i)}$ parameters, for $i \in \{1, 2\}$ by linking them to the PCA decomposition of the covariance matrix of the log of the LIBORs as given in equation (3.18). This requires us to specify the eigenvectors $(a_{ki})_{k=1}^{\tilde{M}}$, $i \in \{1, 2\}$. which are set up as follows:

$$\begin{aligned} a_{k1} &= \frac{1}{\sqrt{\tilde{M}}} \\ a_{k2} &= a + b\exp(\lambda k), \end{aligned} \tag{3.21}$$

for $a, b \in \mathbb{R}$ chosen such that we obtain an orthonormal set of eigenvectors, and $\lambda \in \mathbb{R}$ is a free parameter that controls the shape of the second eigenvector. Recall that we are assuming that the first component of the model driver is capturing the level of rates, hence the first eigenvector is chosen to be flat, and the second component $x_T^{(2)}$ is set to capture the slope, so we choose the second eigenvector to exhibit one sign change. We illustrate below the eigenvectors; we consider two values for λ: 0.1 and 0.3.

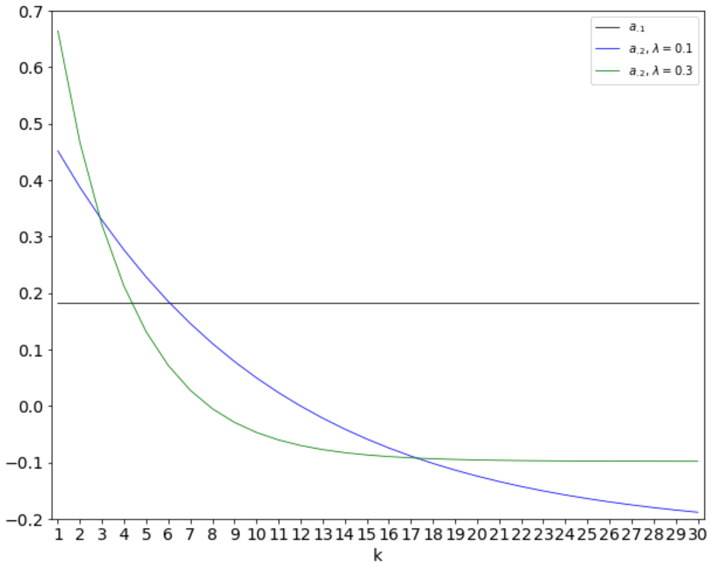

Figure 3.13: Choice of eigenvectors
when $\lambda = 0.1$, $a = -0.225, b = 0.677$
when $\lambda = 0.3$, $a = -0.098, b = 0.762$

3.3.3 The correlation effect when pricing a CMS

We fix a correlation, which we denote by ρ, and we find a value for λ_2 such that for a given reference swap rate maturity and a given payment date, the correlation between the log of the reference swap rate and the log of the payment swap rate is equal to ρ. We plot the convexity correction against payment date for each fixed ρ.

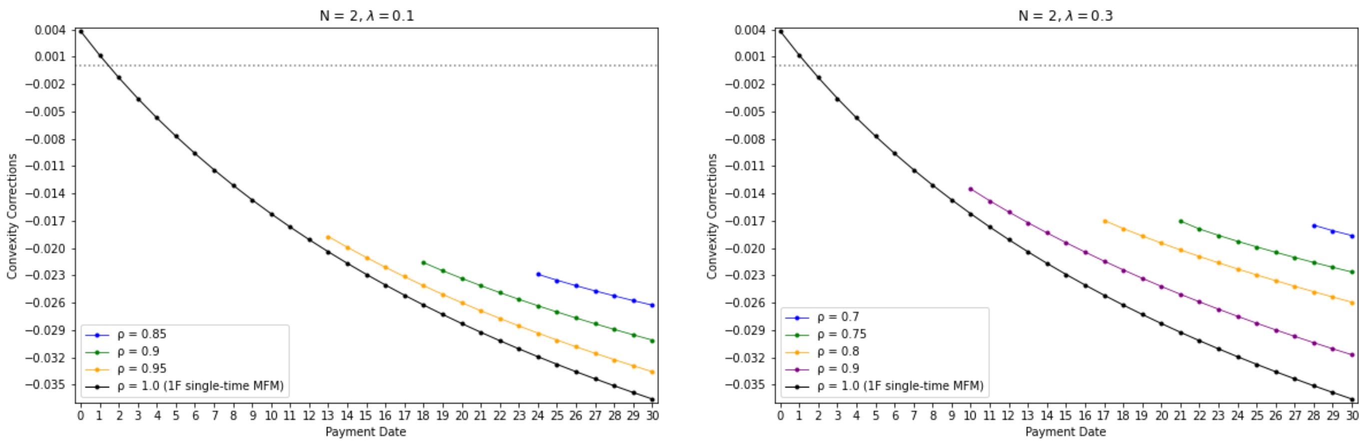

Figure 3.14: Convexity corrections against payment date for fixed correlation
$corr(\log(y_T^N), \log(y_T^M)) = \rho$
$N = 2, \tilde{M} = 30, y_0 = 0.07 D_{0T} = 1.0, \lambda_1 = 1$

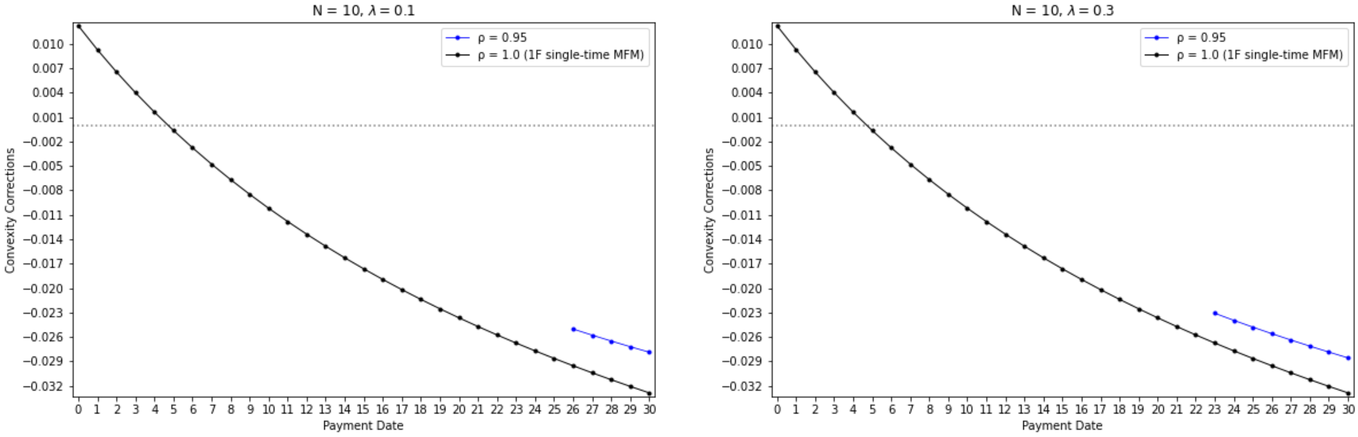

Figure 3.15: Convexity corrections against payment date for fixed correlation $corr(\log(y_T^N), \log(y_T^M)) = \rho$
$N = 10, \tilde{M} = 30, y_0 = 0.07, D_{0T} = 1.0, \lambda_1 = 1$

In Figures 3.14 and 3.15, we first note that for each ρ, there is a minimum payment index preceding which we have no results. This happens because for the choices we have made (on the eigenvectors and the initial conditions), there exists no λ_2 that would equate $\mathrm{corr}(\log(y_T^N), \log(y_T^M))$ to ρ. But we focus our attention on large M, i.e payment dates far from the reference swap rate maturity. From Figure 3.14, we observe that when $\lambda = 0.3$, the convexity correction almost halves in value as correlation decreases from 1 to 0.7. Similarly, from figure 3.15, we can see that the convexity correction is sensitive to correlation. A decrease in correlation by a factor of 0.05 from 1 (so the correlation is still quite high and close to 1) results in a decrease of roughly 0.4% in convexity correction when $M = 30$. In the next part, we employ a different method to measure the significance of correlation in the pricing problem.

To do so, for a given payment date S_M and a reference swap rate maturity S_N we fix the correlation in the same way as we did above, and we compute the convexity correction. We then turn to a 1F single-time MFM and we aim to find an implied volatility, denoted by $\hat{\sigma}_M$, that will yield the same convexity correction. If the difference between the initial log-normal volatility σ_M and $\hat{\sigma}_M$ is significant, this would indicate that the correlation does have an effect on convexity correction and should not be overlooked. (Were it small, a small uncertainty as to the correct volatility input into a one-factor model would have the same effect.) For $M \in \{1, \ldots, \tilde{M}\}$, we define:

$$\Delta_M = \hat{\sigma}_M - \sigma_M. \tag{3.22}$$

We focus on the case when $N = 2$ and $\lambda = 0.3$.

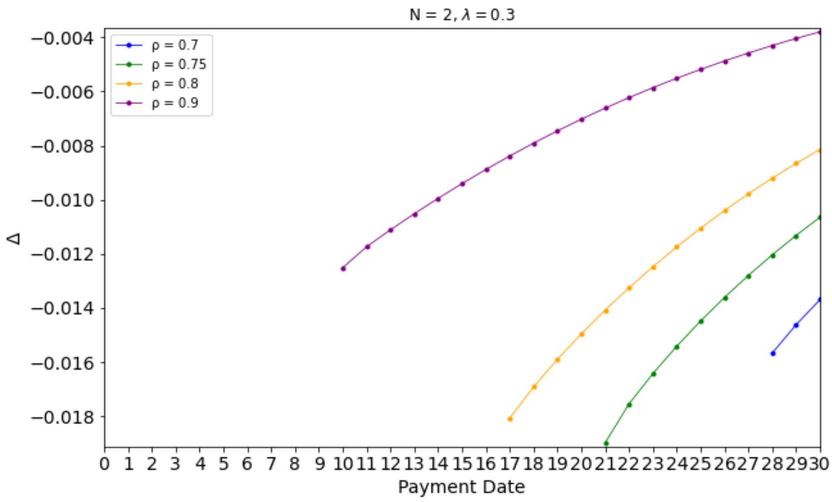

Figure 3.16: CMS: Significance of correlation
$N = 2, \tilde{M} = 30, y_0 = 0.07, D_{0T} = 1.0, \lambda_1 = 1$

Firstly, from Figure 3.16, we observe that the difference is always negative, implying that for any payment date S_M, $\hat{\sigma}_M < \sigma_M$. This is an expected result given that, from Figure 3.14, we know that the convexity correction decreases in magnitude as correlation decreases (i.e, there is an inverse relationship between correlation and convexity corrections). Combined with the fact that the vega is always negative (both vegas associated with the reference swap rate and the payment swap rate), in order to achieve a lower convexity correction in the one-factor model, the volatility has to decrease to reflect the correlation effect in a one-factor model. We observe that the volatility has to be modified by nearly 2% to reflect the correlation effect.

Another feature of the result is that the size of the difference decreases as the payment date increases. This might seem counter-intuitive at first, but is again an expected effect if we take into account the vega. We know from Figure 3.2, that as the payment date increases, the vegas (associated with both the reference swap rate and the payment swap rate) get bigger in magnitude. In other words, for larger payment dates, the convexity correction is more sensitive to the volatility, in the sense that a small change in the volatility will result in a bigger change in the convexity correction. However, while the size of the difference reported in Figure 3.16 might decrease as we increase the payment date, we would still argue that the correlation has a significant effect on convexity corrections.

We now consider a convex payoff (a CMS caplet) and we carry out the same analysis as above. The goal is to determine whether the correlation effect we have observed in the CMS case will carry over to a different payoff.

3.3.4 The correlation effect when pricing a CMS caplet

We work under the same assumptions as laid out in the previous section. For a given strike K, the time-zero value of the payoff of a caplet with reference swap rate y_T^N, and paid at time S_M is given by:

$$V_0^{\text{cap}} = \alpha D_{0T} \mathbb{E}_{\mathbb{F}}[(y_T^N - K)_+ D_{TS_M}].$$

In the numerical analysis below, we focus on the valuation of a single payment of the CMS caplet and we define the convexity correction to be:

$$\mathcal{C}_{N,M} = \frac{D_{0T}}{D_{0S_M}} \mathbb{E}_{\mathbb{F}}[(y_T^N - K)_+ D_{TS_M}] - \mathbb{E}_{\mathbb{S}^N}[(y_T^N - K)_+]. \tag{3.23}$$

Remark 17: In the numerical analysis, when computing the convexity correction, we fix the strike K at $0.07(7\%) \equiv y_0^N$.

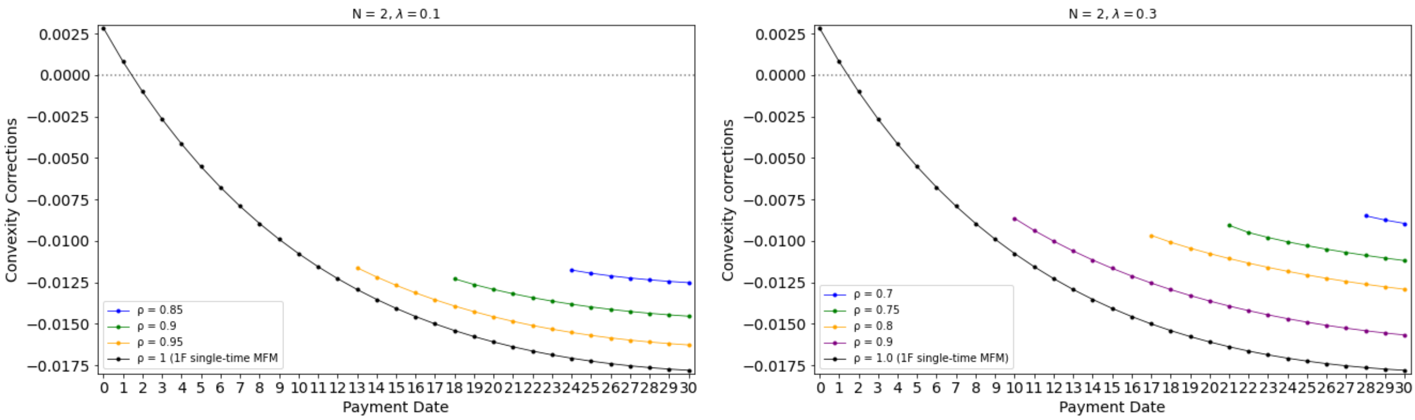

Figure 3.17: CMS Caplet: Convexity correction against payment date for fixed correlation $corr(\log(y_T^N), \log(y_T^M)) = \rho$
$N = 2, \tilde{M} = 30, y_0 = 0.07, D_{0T} = 1.0, \lambda_1 = 1$

We again observe a similar pattern as in the CMS case, whereby for large M, the correlation has a significant impact on convexity correction.

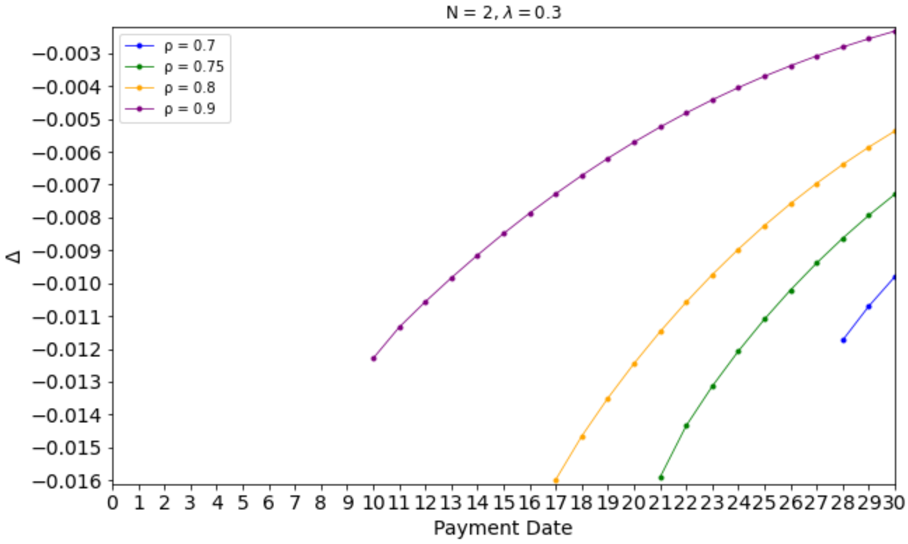

Figure 3.18: CMS Caplet: Significance of correlation
$N = 2, \tilde{M} = 30, y_0 = 0.07, D_{0T} = 1.0, \lambda_1 = 1$

As we have done above, in Figure 3.18, we look at the change in volatility used in a one-factor model to account for the correlation effect, and we again see that the volatility would have to modified by almost 1.7% for reflect the correlation effect on convexity corrections.

CHAPTER 4

A Markov-functional Approach to convexity corrections: The one-factor MF-Lite model

We propose in this chapter a computationally efficient Markov-functional approach at a single time for pricing European-type swap-based derivatives whose payoffs are of form $F(y_T^n)$, where $F : \mathbb{R} \to \mathbb{R}$ is a deterministic function and y_T^n is the underlying forward swap rate. We use the insight from the single-time MFM numerical analysis to guide the modelling approach. In particular, recall in section 3.2 of Chapter 3, we set out to identify the forward swap rates that matter most in the pricing of a CMS payment. We did so under three assumptions on the shape of the volatility curve: flat (assuming log-normal marginals on the swap rates in their respective swaption measures), skew and smile. In all three cases, at the payment dates S_M that are of practical interest (i.e $M = 0, 1$ or M far away from the reference swap rate maturity), we observed that the marginal distributions of the reference swap rate and payment swap rate have a significant impact on convexity corrections. We also found that the other swap rates have a non-negligible impact, but we point out firstly that the numerical analysis was carried out under fairly stressed market conditions, and secondly, the effect of the payment swap rate outweighs that of the other swap rates as M gets large. This points us to believe that if we want to price a single cashflow of a CMS for instance, one could set up a model such that the swap rates at the relevant maturities are modelled appropriately, and the information from the other swap rates could be incorporated in the modelling choices we make in an efficient way. We use this insight as our starting point.

We start the development of the modelling approach in the simpler one-factor case. We choose a maturity S_n and model y_T^n and P_T^n under the swaption measure $\mathbb{S}^n$. We proceed as follows: we postulate a functional form for the swap rate y_T^n in terms of a one-dimensional driver, whose distribution under $\mathbb{S}^n$ is yet to be determined. The known market-implied distribution of the swap rate under $\mathbb{S}^n$ and the postulated functional form then determine the distribution of the driver under $\mathbb{S}^n$.

Next we choose the functional form of P_T^n in terms of the driver. This constitutes a modelling choice. At this stage, we have also fixed the functional form of D_{TS_n} since $D_{TS_n} = 1 - y_T^n P_T^n$. Consequently, we know P_T^{n-1} as well since $P_T^{n-1} = P_T^n - \alpha_n D_{TS_n}$. This has implications on the distribution of the model driver under $\mathbb{S}^{n-1}$.

Note we have considerable freedom in the choice of the functional form of P_T^n but we are constrained by the martingale requirement of $\frac{D_{\cdot T}}{P^n}$ under $\mathbb{S}^n$. Note also that P^n is the variable that allows us to change measure from $\mathbb{S}^n$ to the forward measure $\mathbb{F}$. This modelling of y_T^n and P_T^n directly contrasts with the single-time MFM discussed in Chapter 2 where we must model y_T^1, then y_T^2, then y_T^3 etc until we reach y_T^n.

The idea described is illustrated in figure 4.1 below, whereby the modelling choices are given in green, and the functional forms that follow from these choices are illustrated in red. We refer to this setup as the *partial model*. We develop and discuss the partial model setup in Section 4.1.

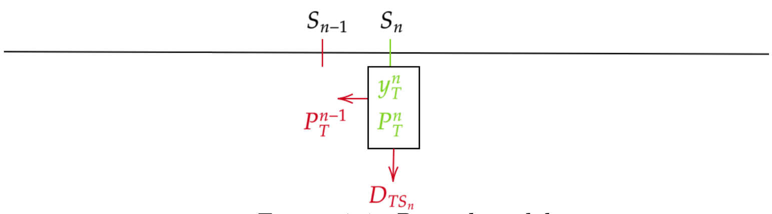

Figure 4.1: Partial model setup

The partial model described above specifies a model for y_T^n (and D_{TS_n}) under the forward measure $\mathbb{F}$. Similarly, we can set up a second partial model at S_m to specify a model for D_{TS_m} under $\mathbb{F}$. Therefore, combining the two, we are able to price the CMS payment y_T^n paid at S_m.

As we shall discuss in detail later, if we follow this approach, which we refer to as the *exact fit* for all payment dates S_m, the resulting CMS prices will not satisfy the consistency condition (3.4). This is not a problem for $m > n$ where there is no consistency condition to satisfy. To resolve the issue for $m < n$, we propose the *alternate fit* model. We elaborate on this in Section 4.2.2.

We have briefly discussed above how we could set the model up for a single maturity, and we mentioned the possibilities of doing the partial setup for two maturities. A careful reader will notice that in doing so, we end up with two model drivers. We make use of the *Inversion Principle*, which we will state in Section 4.1, to show that under a common measure - in this case, the forward measure - there exists a functional relationship between the two drivers, which bring us back to a one-factor setting. We discuss this in Section 4.2.

In Section 4.2.1, we show how we can extend the partial model setup to a complete, arbitrage-free model. Note that the method proposed follows very closely the construction of the single-time MFM and ends up being computationally intensive, hence not one that will be attractive in practice. But it shows that the partial model setup can be extended to a complete model. We also note that in this section, we have specified the model at a fixed time T. This is enough for the exotic products of European type that the model is designed to price. In principle, we could extend the model over the full term structure and consider the setup at any time $t < T$ using the martingale property of numéraire rebased assets to do so.

The model developed in this chapter is referred to as the one-factor swaption measure calibrated MF-Lite model, which we can abbreviate to 1F smcMFL model. We conclude this chapter with some numerical results comparing the performance of the 1F smcMFL model to the single-time MFM, and we investigate how sensitive the model is to the choices of functional forms that we make. We observe that the model proposed is numerically very close to the single-time MFM. There is an oscillating, zig-zag type behaviour appearing in the alternate fit setup, indicating that where we have no flexibility to calibrate the model to market information, the model's approximation for the swap rate distribution gets slightly worse (especially for payment dates well after the reference swap end date). We provide a detailed explanation in Chapter 5 as to why this oscillating behaviour occurs. But this clearly is not an issue, since for payment dates beyond the length of the reference swap rate, we can fall back onto the exact fit approach that yields results close to that of the single-time MFM, as desired. We nonetheless provide in Section 4.3.5 a numerical refinement to the alternate fit model that shrinks the zig-zags and brings the model closer to the single-time MFM.

Before moving on we make one final observation. Any one-dimensional term-structure model could be built using the approach described here. If, for each S_n, we fit y_T^n to its distribution as determined by the given term-structure model and if we choose the functional form P_T^n to be that given by the term-structure model (which would not be possible in practice as this would typically be a highly complex unknown functional form), then the resultant MF-Lite model is precisely the term-structure model. This is obvious but highlights the fact that any undesirable feature of a given MF-Lite model (such as the extreme zig-zag behaviour in the alternate fit) is the result of a poor modelling choice, not a weakness of the framework. Had we instead chosen the functional form consistent with a term-structure model, the undesirable feature would not be present.

4.1 Model setup for a single maturity

In the discussion above, we outlined at the high level how the 1F smcMFL model could be set up for any given maturity which we denoted by S_n.

We assume there exists some univariate random variable, denoted by η_n, summarising the state of the economy at time T. We could think of η_n as capturing the level of rates. Realistically, we would want the swap rates to be monotonic increasing in η_n, and pure discount bonds to be a decreasing function of the driver.

If we know the prices of swaptions for a set of strikes, we can work out the implied marginal distribution of y_T^n under its associated swaption measure $\mathbb{S}^n$. We want to appropriately model this distribution. In order to achieve this, we make use of the Inversion Principle, also known as the Inverse Transform method (Devroye (1986)) as a means to express y_T^n as a function of a standard normal random variable.

Theorem 4.1.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a given probability space. Let F be a continuous probability distribution function on $\mathbb{R}$ with F^{-1} defined by:*

$$F^{-1}(u) := \inf\{x : F(x) = u : 0 < u < 1\}.$$

If U is uniform [0,1] random variable, $F^{-1}(U)$ has distribution function F. Also, if X has distribution function F, $F(X)$ is uniformly distributed on [0,1].

If $X_n \sim \mathcal{N}(0,1)$ under $\mathbb{S}^n$, then following the second statement of the Theorem, we have that

$$\Phi(X_n) \sim U[0,1].$$

Using the first part of the Theorem, if y_T^n has a known marginal distribution, denoted by F_n^y under $\mathbb{S}^n$, then

$$X_n := \Phi^{-1}\big(F_n^y(y_T^n)\big)$$

has a $\mathcal{N}(0,1)$ distribution under $\mathbb{S}^n$. Thus,

$$y_T^n = (F_n^y)^{-1}(\Phi(X_n)). \tag{4.1}$$

We denote $g_n := (F_n^y)^{-1} \circ \Phi$.

As a first step in setting up the model, we postulate a functional form for y_T^n in terms of η_n, i.e,

$$y_T^n := f_n(\eta_n),$$

whereby we assume $f_n(\cdot)$ is monotonic increasing.

From these two assumptions, we can work out the distribution of η_n under $\mathbb{S}^n$.

$$\begin{aligned} y_T^n &= f_n(\eta_n) \\ \implies \eta_n &= f_n^{-1}(y_T^n) \\ &= f_n^{-1}(g_n(X_n)) =: q_n(X_n). \end{aligned} \tag{4.2}$$

By assumption of f_n being monotonic increasing, we know $f_n^{-1}(\cdot)$ exists, and we express η_n as a function $q_n : \mathbb{R} \to \mathbb{R}$, where $q_n = f_n^{-1} \circ g_n$, of a standard normal random variable. So, we have effectively determined the distribution of η_n under $\mathbb{S}^n$.

The second step consists of postulating a prior functional form for P_T^n which we denote by $\hat{P}_T^n$ in terms of the driver η_n. To find P_T^n, we will need to modify this prior in order to ensure $\frac{D_{\cdot T}}{P^n}$ is a martingale under $\mathbb{S}^n$. In order to specify $\hat{P}_T^n$, we recall from equation (2.4) in Chapter 2 that we can express the PVBP P_T^n as a function of a set of forward swap rates. So, we can come up with a set of priors for the swap rates, which we denote by $\{\hat{y}_T^i; i \in \{1, ..., n\}\}$ and use these to construct $\hat{P}_T^n$. We note that these prior swap rates are only used to set up the PVBP and are not used elsewhere in the final model. We have that:

$$\hat{y}_T^i(\eta_n) := \hat{f}_i(\eta_n), \tag{4.3}$$

whereby $\hat{f}_i : \mathbb{R} \to \mathbb{R}$ is chosen to reflect our belief about the marginal distributions of these swap rates in their own swaption measures. We shall provide a concrete choice for the functional form of the prior swap rates in 4.3.1 when we set up the 1F smcMFL model under the (shifted) log-normal assumption. We can then express:

$$\hat{P}_T^n(\eta_n) := \sum_{k=1}^{n} \alpha_k \left(\prod_{i=k}^{n} \frac{1}{1 + \alpha_i \hat{f}_i(\eta_n)} \right). \tag{4.4}$$

We assume that

$$P_T^n(\eta_n) := a_n \hat{P}_T^n(\eta_n),$$

for $a_n \in \mathbb{R}$. The parameter a_n is chosen such that the *martingale property*[1] holds. In

[1] Note that we are using this term loosely here; we could look at the expectation in equation (4.5) as the conditional expectation with respect to $\mathcal{F}_0$. We are only concerned with the model behaviour at a fixed time T. We could view this no-arbitrage condition as a consequence of the martingale property, which we henceforth shorten to the martingale property without ambiguity.

particular,

$$\begin{aligned}
\frac{D_{0T}}{P_0^n} &= \mathbb{E}_{\mathbb{S}^n}\left[\frac{D_{TT}}{P_T^n}\right] && (4.5)\\
&= \mathbb{E}_{\mathbb{S}^n}\left[\frac{1}{a_n \hat{P}_T^n(\eta_n)}\right]\\
\implies a_n &= \frac{P_0^n}{D_{0T}}\mathbb{E}_{\mathbb{S}^n}\left[\frac{1}{\hat{P}_T^n(\eta_n)}\right]. && (4.6)
\end{aligned}$$

Given the distribution of η_n under $\mathbb{S}^n$ in equation (4.2), we can compute a_n:

$$a_n = \frac{P_0^n}{D_{0T}}\int_{-\infty}^{\infty}\frac{1}{\hat{P}_T^n(q_n(x))}\phi(x)\mathrm{d}x,$$

where $\phi(.)$ is the standard normal density.

Remark 18*: Note that under reasonable modelling choices, we expect the parameter a_n to be roughly equal to 1. We could, for instance, choose to express y_T^n and P_T^n as a function of η_n as given by a short rate model, in which case we could take $\eta_n \equiv r_T$. We can make similar choices at each maturity in the tenor. If we now assume that the market implied distribution of each swap rate matches exactly what would be given by a short rate model, we can construct an arbitrage-free model that will be able to capture the market-implied distributions. The model, by construction, will be a short rate model calibrated to the full set of implied distributions, and the model driver will have the same distribution as the short rate from which we are choosing the implied distributions. Alternatively, instead of starting at S_n, we could use a Markov-functional approach whereby we start off by choosing the functional form y_T^1 by matching to the relevant implied distribution and unfolding the model by fixing only the swap rate functional form at each given maturity in the tenor, as described in Chapter 2. Owing to the fact that there is a unique arbitrage-free one-factor model that calibrates to the whole set of implied distributions, the short rate model will exactly match the Markov-functional model. In this particular case, the parameter $a_.$ corresponding to each maturity will be exactly unity. For any other choice of market implied distributions whilst keeping the choice of functional forms given by a short rate model, η_n will have a distribution different from that of the short-rate driver, and the parameter a_n will have to be chosen appropriately to satisfy the martingale property.*

From the modelling choices made above, we are able to recover the functional form of

D_{TS_n}. By definition of y_T^n, we have:

$$\begin{aligned} y_T^n &:= \frac{1 - D_{TS_n}}{P_T^n} \\ \implies D_{TS_n} &= 1 - y_T^n P_T^n \\ &= 1 - a_n f_n(\eta_n) \hat{P}_T^n(\eta_n). \end{aligned}$$

So far, we have been working under the swaption measure $\mathbb{S}^n$, whereby we have fixed the distribution of η_n. We now carry over the model setup to the the forward measure $\mathbb{F}$ and by a change of measure, we work out the distribution of η_n. Define $F_{\eta_n}(x) := \mathbb{E}_{\mathbb{F}}\Big[\mathbb{1}(\eta_n \leq x)\Big]$. For a given x^* in $\mathbb{R}$, we have that:

$$\begin{aligned} F_{\eta_n}(x^*) &:= \mathbb{E}_{\mathbb{F}}\Big[\mathbb{1}(\eta_n \leq x^*)\Big] \\ &= \frac{P_0^n}{D_{0T}} \mathbb{E}_{\mathbb{S}^n}\left[\mathbb{1}(\eta_n \leq x^*) \frac{D_{TT}}{P_T^n}\right] \\ &= \frac{P_0^n}{D_{0T}} \mathbb{E}_{\mathbb{S}^n}\left[\mathbb{1}(\eta_n \leq x^*) \frac{1}{a_n \hat{P}_T^n(\eta_n)}\right]. \end{aligned} \tag{4.7}$$

We point out at this stage that the distribution of the driver under the forward measure is determined by the choice of the PVBP functional form. We shall investigate numerically in later sections, the sensitivity of our approach to the choice of functional form for the PVBP.

Using the fact that $\eta_n = q_n(X_n)$, $X_n \sim \mathcal{N}(0,1)$ under $\mathbb{S}^n$, we have

$$\begin{aligned} F_{\eta_n}(x^*) &= \frac{P_0^n}{a_n D_{0T}} \mathbb{E}_{\mathbb{S}^n}\left[\mathbb{1}(q_n(X_n) \leq x^*) \frac{1}{\hat{P}_T^n(q_n(X_n))}\right] \\ &= \frac{P_0^n}{a_n D_{0T}} \mathbb{E}_{\mathbb{S}^n}\left[\mathbb{1}(X_n \leq q_n^{-1}(x^*)) \frac{1}{\hat{P}_T^n(q_n(X_n))}\right] \\ &= \frac{P_0^n}{a_n D_{0T}} \int_{-\infty}^{q_n^{-1}(x^*)} \frac{1}{\hat{P}_T^n(q_n(x))} \phi(x) \mathrm{d}x. \end{aligned}$$

By choosing to model y_T^n and P_T^n, we have equally fixed the functional form of P_T^{n-1}, since

$$P_T^{n-1} = P_T^n - \alpha_n D_{TS_n}. \tag{4.8}$$

Remark 19*: We make the following observation: we started with choosing the functional forms of y_T^n and P_T^n at the maturity date S_n. We assumed we know the marginal distribution of y_T^n under $\mathbb{S}^n$ and we moved over to the forward measure $\mathbb{F}$, where we are* *able to derive the distribution of η_n. Moving to the previous maturity, we already know the functional form of P_T^{n-1}, so we can derive the distribution of η_n under the swaption measure $\mathbb{S}^{n-1}$ using the fact that for $x \in \mathbb{R}$,*

$$\mathbb{E}_{\mathbb{S}^{n-1}}\left[\mathbb{1}(\eta_n < x)\right] = \frac{D_{0T}}{P_0^{n-1}}\mathbb{E}_{\mathbb{F}}\left[\mathbb{1}(\eta_n < x)P_T^{n-1}\right].$$

Given we know the marginal distribution of η_n under $\mathbb{F}$, we can compute the expectation on the RHS, and we are therefore able to derive the distribution of η_n under $\mathbb{S}^{n-1}$.

If we take $S_n = S_2$ for example, and we make some modelling choices at that maturity, it fixes up the functional form for P_T^1, and equivalently, we fix the functional form for y_T^1. This follows from the fact that $P_T^1 = (1 + \alpha_1 y_T^1)^{-1}$. So the modelling choices at S_2, fixes all the functional forms at S_1 as well as the distribution of η_n under $\mathbb{S}^1$. In this case, we have fixed the distribution of y_T^1 under $\mathbb{S}^1$ (and this distribution might be different from the market-implied distribution of y_T^1).

The partial model as defined is set up to be arbitrage-free. We can easily prove that this is the case by showing that the martingale property under the forward measure holds. In particular, we would want the following relationships to hold:

$$\mathbb{E}_{\mathbb{F}}[P_T^j] = \frac{P_0^j}{D_{0T}}, \qquad \text{for } j = n-1, n \tag{4.9}$$

$$\mathbb{E}_{\mathbb{F}}[P_T^n y_T^n] = \frac{y_0^n P_0^n}{D_{0T}}. \tag{4.10}$$

For $j = n$, equation (4.9) trivially holds by a change of measure to $\mathbb{S}^n$. Note that in the model, we equally fix the functional form for P_T^{n-1}. We show that the martingale property holds for this variable as well.

$$\begin{aligned}
\mathbb{E}_{\mathbb{F}}[P_T^{n-1}] &= \frac{P_0^n}{D_{0T}}\mathbb{E}_{\mathbb{S}^n}\left[\frac{P_T^{n-1}}{P_T^n}\right] \\
&= \frac{P_0^n}{D_{0T}}\mathbb{E}_{\mathbb{S}^n}\left[\frac{a_n\hat{P}_T^n(\eta_n)(1+\alpha_n f_n(\eta_n)) - \alpha_n}{a_n\hat{P}_T^n(\eta_n)}\right] \\
&= \frac{P_0^n}{D_{0T}}\mathbb{E}_{\mathbb{S}^n}\left[1 + \alpha_n f_n(\eta_n) - \frac{\alpha_n}{a_n\hat{P}_T^n(\eta_n)}\right] \\
&= \frac{P_0^n}{D_{0T}}\mathbb{E}_{\mathbb{S}^n}[1 + \alpha_n g_n(X_n)] - \alpha_n\frac{P_0^n}{D_{0T}}\mathbb{E}_{\mathbb{S}^n}\left[\frac{1}{a_n\hat{P}_T^n(\eta_n)}\right] \\
&= \frac{P_0^n}{D_{0T}}(1 + \alpha_n y_0^n) - \alpha_n.
\end{aligned} \tag{4.11}$$

Given that $g_n(X_n)$ is the market-implied distribution of y_T^n under $\mathbb{S}^n$, it immediately follows that $\mathbb{E}_{\mathbb{S}^n}[g_n(X_n)] = y_0^n$. The last term in the expression follows from the definition of the parameter a_n given in equation (4.6).

Equation (4.11) can be further simplified as follows:

$$
\begin{aligned}
\mathbb{E}_{\mathbb{F}}[P_T^{n-1}] &= \frac{P_0^n}{D_{0T}}\left(1+\alpha_n\frac{D_{0T}-D_{0S_n}}{p_0^n}\right)-\alpha_n \\
&= \frac{P_0^n+\alpha_n D_{0T}-\alpha_n D_{0S_n}-\alpha_n D_{0T}}{D_{0T}} \\
&= \frac{P_0^{n-1}}{D_{0T}},
\end{aligned}
$$

and we have our desired result.

To show that equation (4.10) holds in our model, we have:

$$
\begin{aligned}
\mathbb{E}_{\mathbb{F}}[P_T^n y_T^n] &= \frac{P_0^n}{D_{0T}}\mathbb{E}_{\mathbb{S}^n}\left[P_T^n y_T^n \frac{D_{TT}}{P_T^n}\right] \\
&= \frac{P_0^n}{D_{0T}}\mathbb{E}_{\mathbb{S}^n}[y_T^n] \\
&= \frac{P_0^n}{D_{0T}}\mathbb{E}_{\mathbb{S}^n}[f_n(\eta_n)] \\
&= \frac{P_0^n}{D_{0T}}\mathbb{E}_{\mathbb{S}^n}[g_n(X_n)] \\
&= \frac{P_0^n y_0^n}{D_{0T}}.
\end{aligned}
$$

4.2 Model setup for two maturities

So far, we have shown how we could set up the model for a single maturity date. However, when pricing a CMS, two maturities are of particular importance: the reference swap rate maturity and the payment date. Ideally, we would want to be able to use the partial model setup described above for the two maturities of interest. To generalise the discussion, we denote these two times as S_n and S_m. If we apply the setup above to these two maturities, we would end up with two drivers η_n and η_m with known distributions under $\mathbb{F}$. There exists however a functional relationship between the two drivers. This enables us to unify the two partial models under the forward measure $\mathbb{F}$ using the Inversion Principle given in Theorem 4.1.1.

For $k = n, m$, we know the distribution of η_k under $\mathbb{F}$ which we denoted by F_{η_k}. With this information, we can construct the model under $\mathbb{F}$. Let $Y \sim \mathcal{N}(0,1)$ under $\mathbb{F}$ and define for $k = n, m$:

$$\eta_k := h_k(Y), \tag{4.12}$$

where $h_k := F_{\eta_k}^{-1} \circ \Phi$. The variable η_k thus defined has the required distribution F_{η_k} under $\mathbb{F}$. We can now establish a functional relationship between η_n and η_m as follows:

$$\begin{aligned}\eta_m &= h_m(Y)\\ &= h_m(h_n^{-1}(\eta_n))\\ &= h_m((F_{\eta_n}^{-1} \circ \Phi)^{-1}(\eta_n))\\ &= F_{\eta_m}^{-1}(\Phi(\Phi^{-1}(F_{\eta_n}(\eta_n))))\\ &= F_{\eta_m}^{-1}(F_{\eta_n}(\eta_n)).\end{aligned}$$

Building on remark 19, since the modelling choices at one given maturity fixes the functional form for the PVBP at the previous maturity, it is important to consider the model when the maturities are one step away from each other. When $m = n+1$, we start off by setting up the partial model at S_m, i.e, we postulate a functional form for y_T^{n+1} and P_T^{n+1} as a function of the driver which we simply denote here by η. This will force upon us a functional form for P_T^n (hence, P_T^n is no longer a modelling choice) and the distribution of η under $\mathbb{S}^{n+1}$ and $\mathbb{F}$. As we have discussed before, through the knowledge of the functional form of P_T^n and the distribution of η under $\mathbb{F}$, we equally know the distribution of η under $\mathbb{S}^n$. The question therefore is how do we capture the implied distribution of the swap rate y_T^n under $\mathbb{S}^n$.

One way to tackle this is through the use of the Inversion Theorem 4.1.1 to find a functional form for y_T^n that would give us the distribution of the swap rate under the swaption measure $\mathbb{S}^n$ we want to capture. We know the distribution of η under $\mathbb{S}^n$, which we denote by $\bar{F}_{\eta_n}$. We also know our target distribution for y_T^n under $\mathbb{S}^n$, which we denoted by F_n^y and which we can express as a function of a standard normal random variable, $X_n \sim \mathcal{N}(0,1)$. By Theorem 4.1.1, we have that:

$$\begin{aligned}\eta &= \bar{F}_{\eta_n}^{-1}(\Phi(X_n))\\ y_T^n &= (F_n^y)^{-1}(\Phi(X_n))\end{aligned}$$

We can thus express y_T^n as a function of η as follows:

$$y_T^n = (F_n^y)^{-1}(\bar{F}_{\eta_n}(\eta)). \tag{4.13}$$

By property of probability distributions, we know $(F_n^y)^{-1}$ is non-decreasing, so we can set $f_n := (F_n^y)^{-1} \circ \bar{F}_{\eta_n}$. This particular choice of functional form for the swap rate at time $\mathbb{S}^n$ ensure the model captures the market implied distribution of y_T^n under $\mathbb{S}^n$. Equation (4.13) is not an obvious choice of functional form and it requires some computational effort.

4.2.1 Extending to a complete arbitrage-free model

We have shown above that the *partial* model is arbitrage-free for a given maturity. For practical purposes, this is usually enough. We still want to check that the partial model can be extended to a complete arbitrage-free model. Below, we describe one way we could extend the model over the whole tenor structure. The extension of the partial model follows a similar reasoning as the single-time MFM setting discussed in Chapter 2. WLOG, we assume $n_2 > n_1$. We do the partial model setup for maturities S_{n_1} and S_{n_2}. Starting at S_{n_2}, we unfold the model backwards stepping one maturity down, and calibrating the model to the known market-implied distribution of the swap rate at each maturity until we reach S_{n_1+2}. Since we have already chosen the functional forms at S_{n_1}, we set the functional form of the swap rate $y_T^{n_1+1}$ so that the functional relationship between the swap rate and the PVBP (we refer ahead to equation (4.14)) is satisfied. We provide below a simple illustration of the extended model. We repeat the same steps until we reach the maturity date S_2. The model choice for the swap rate at S_2, together with the PVBP P_T^2 determined by the model choices at S_3, fully determines the model at S_1. We extended the model in such a way that at most maturities, we are able to calibrate the model to the market-implied distribution of the swap rates. The extended model is theoretically sound, but one that may not be attractive in practice. In the next section, we shall discuss how we could set up the partial models such as the resulting model is arbitrage free. We lose some flexibility over the calibrating aspect of the model, but the numerical analysis that follows indicate that under judicious modelling choices, we could set up the 1F smcMFL model to closely reproduce the single-time MFM results.

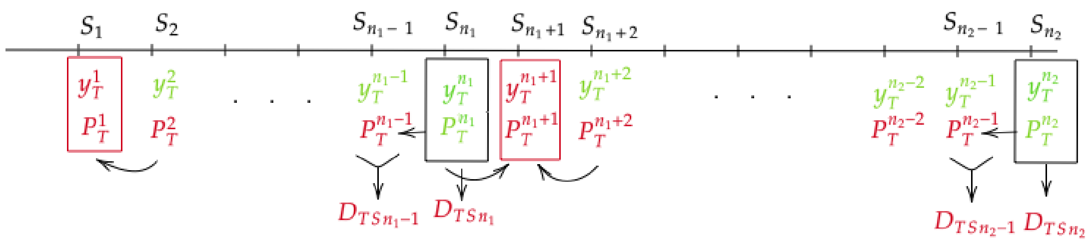

Figure 4.2: A complete arbitrage-free model

We explain the complete model construction below:

(1) Starting at time S_{n_2}, we make modelling choices for $y_T^{n_2}$ and $P_T^{n_2}$. We showed above that this fixes the functional form for $P_T^{n_2-1}$. Working down from S_{n_2-1} to time S_{n_1+2}, at each time step, we choose a functional form for y_T^j, and we make the following choice:

$$f_j(\eta) := (F_j^y)^{-1}(\bar{F}_{\eta_j}(\eta)),$$

for $j \in \{n_1+2, \ldots, n_2-1\}$, where F_j^y is the known implied distribution function of y_T^j under $\mathbb{S}^j$ and $\bar{F}_{\eta_j}(\cdot)$ is the distribution of η under $\mathbb{S}^j$, fixed from the previous time step, as can be easily deduced from the discussion in remark 19.

(2) With the choice of functional form for $y_T^{n_1+2}$ made and $P_T^{n_1+2}$ fixed from previous time step, we have fixed the functional form of $P_T^{n_1+1}$. At time S_{n_1}, we have already made the modelling choices for $P_T^{n_1}$ and $y_T^{n_1}$. We therefore have to set up the functional form of $y_T^{n_1+1}$ in order to be consistent with the PVBP choices at time S_{n_1} and S_{n_1+1}. The consistent choice would be:

$$\begin{aligned} P_T^{n_1+1} + \alpha_{n_1+1} y_T^{n_1+1} P_T^{n_1+1} &= P_T^{n_1} + \alpha_{n_1+1} \\ \implies y_T^{n_1+1} &= \frac{P_T^{n_1} + \alpha_{n_1+1} - P_T^{n_1+1}}{\alpha_{n_1+1} P_T^{n_1+1}}. \end{aligned} \tag{4.14}$$

Remark 20*: Note that by making the above choice, we lose monotonicity of the functional form for $y_T^{n_1+1}$ and the distribution of $y_T^{n_1+1}$ is fixed under $\mathbb{S}^{n_1+1}$, and has not been matched to the implied distribution. But with appropriate choices of functional forms, we expect the approximate model distribution for the swap rate to be relatively close to the market implied distribution.*

(3) Working from the modelling choices at S_{n_1}, down to maturity S_2, similar to the above, we are free to make choices for the functional forms of y_T^j, $j \in \{2, \ldots, n_1 - 1\}$, and we can use similar choice as in step (1) above. By the time we reach S_1, P_T^1 is fixed from the choices at S_2 and the functional form of y_T^1 automatically follows since:

$$P_T^1 = \frac{\alpha_1}{1 + \alpha_1 y_T^1}.$$

(4) At each time step j for $j \in \{1, \ldots, n_2\}$, we can recover the functional form for the pure discount bond using the fact that:

$$D_{TS_j} = 1 - y_T^j P_T^j.$$

and we fixed the functional form of the PVBP at the previous time step using:

$$P_T^{j-1} = P_T^j - \alpha_j D_{TS_j}, \text{ for } j > 2. \tag{4.15}$$

The model, thus extended, will be arbitrage-free, which is verified below, and is also consistent by construction.

At times S_{n_i}, for $i \in \{1, 2\}$, we have shown that the model is arbitrage-free. We have equally proved that at time S_{n_i-1}, the required martingale property holds. For $j \in \{1, \ldots, n_2-2\} \backslash \{n_1-1, n_1\}$, we use a backward inductive argument to prove the martingale property holds.

Assume the martingale property holds at time S_{j+1}, for some $j \in \{1, \ldots, n_2-2\} \backslash \{n_1-1, n_1\}$. We want to prove it equally holds at S_j:

$$
\begin{aligned}
\mathbb{E}_{\mathbb{F}}[P_T^j] &= \frac{P_0^{j+1}}{D_{0T}} \mathbb{E}_{\mathbb{S}^{j+1}}\left[\frac{P_T^j}{P_T^{j+1}}\right] \\
&= \frac{P_0^{j+1}}{D_{0T}} \mathbb{E}_{\mathbb{S}^{j+1}}\left[\frac{P_T^{j+1} - \alpha_{j+1} D_{TS_j+1}}{P_T^{j+1}}\right] \\
&= \frac{P_0^{j+1}}{D_{0T}} \mathbb{E}_{\mathbb{S}^{j+1}}\left[1 - \frac{\alpha_{j+1}(1 - y_T^{j+1} P_T^{j+1})}{P_T^{j+1}}\right] \\
&= \frac{P_0^{j+1}}{D_{0T}} \left(1 - \alpha_{j+1} \mathbb{E}_{\mathbb{S}^{j+1}}\left[\frac{1}{P_T^{j+1}}\right] + \alpha_{j+1} \mathbb{E}_{\mathbb{S}^{j+1}}[y_T^{j+1}]\right). \qquad (4.16)
\end{aligned}
$$

Given we have chosen to model y_T^{j+1} such that its functional form $f_{j+1}(\cdot)$ captures the market implied distribution, it follows that

$$
\mathbb{E}_{\mathbb{S}^{j+1}}[y_T^{j+1}] = \mathbb{E}_{\mathbb{S}^{j+1}}[f_{j+1}(\eta)] = y_0^{j+1}.
$$

Remark 21: Note that even if we made a different choice for y_T^j at S_1 and S_{N+1}, for the martingale property to hold at these two time steps, we need y_T^2 and y_T^{N+2} respectively to satisfy the above equations and this follows from the choice of functional forms.

We can easily see that:

$$
\mathbb{E}_{\mathbb{S}^{j+1}}\left[\frac{1}{P_T^{j+1}}\right] = \mathbb{E}_{\mathbb{S}^{j+1}}\left[\frac{D_{TT}}{P_T^{j+1}}\right] = \frac{D_{0T}}{P_0^{j+1}}.
$$

Equation (4.16) then simplifies to:

$$
\begin{aligned}
\mathbb{E}_{\mathbb{F}}[P_T^j] &= \frac{P_0^{j+1} - \alpha_{j+1} D_{0T} + \alpha_{j+1}(D_{0T} - D_{0S_{j+1}})}{D_{0T}} \qquad (4.17) \\
&= \frac{P_0^j}{D_{0T}}. \qquad (4.18)
\end{aligned}
$$

In particular, using the result in equation (4.18) from above, we have:

$$\begin{aligned}\mathbb{E}_{\mathbb{F}}[D_{TS_j}] = \frac{1}{\alpha_j}\mathbb{E}_{\mathbb{N}}[P_T^j - P_T^{j-1}] &= \frac{1}{\alpha_j}\left[\frac{P_0^j - P_0^{j-1}}{D_{0T}}\right] \\ &= \frac{D_{0S_j}}{D_{0T}}.\end{aligned}$$

For maturities where we have free choice for the functional forms of y_T^j, and we chose them to match the implied distribution under their respective swap rates, equation (4.10) will hold in the extended model, and this can be shown by following similar steps as we did in the partial model, but instead we apply a change of measure from $\mathbb{F}$ to $\mathbb{S}^j$.

We look at the case $j \in \{1, n_1 + 1\}$ in particular to check if it satisfies the martingale property.

$$\begin{aligned}\mathbb{E}_{\mathbb{N}}[y_T^1 P_T^1] &= \frac{P_0^1}{D_{0T}}\mathbb{E}_{\mathbb{S}^1}[y_T^1] \\ &= \frac{P_0^1}{D_{0T}}\left(\mathbb{E}_{\mathbb{S}^1}\left[\frac{1}{P_T^1}\right] - \frac{1}{\alpha_1}\right) \\ &= \frac{P_0^1}{D_{0T}}\left(\frac{D_{0T}}{P_0^1} - \frac{1}{\alpha_1}\right) \\ &= \frac{P_0^1}{D_{0T}}\left(\frac{D_{0T} - D_{0S_1}}{\alpha_1 D_{0S_1}}\right) \\ &= \frac{P_0^1 y_0^1}{D_{0T}}\end{aligned}$$

For $j = n_1 + 1$, we use the fact that both $P_T^{n_1}$ and $P_T^{n_1+1}$ satisfy the martingale property to show that equation (4.10) holds for that particular choice of $y_T^{n_1+1}$.

$$\begin{aligned}\mathbb{E}_{\mathbb{F}}[y_T^{n_1+1} P_T^{n_1+1}] &= \mathbb{E}_{\mathbb{F}}\left[\frac{P_T^{n_1} + \alpha_{n_1+1} - P_T^{n_1+1}}{\alpha_{n_1+1} P_T^{n_1+1}} P_T^{n_1+1}\right] \\ &= \mathbb{E}_{\mathbb{F}}\left[\frac{P_T^{n_1} + \alpha_{n_1+1} - P_T^{n_1+1}}{\alpha_{n_1+1}}\right] \\ &= \frac{1}{\alpha_{n_1+1}}\mathbb{E}_{\mathbb{F}}[P_T^{n_1}] + 1 - \frac{1}{\alpha_{n_1+1}}\mathbb{E}_{\mathbb{F}}[P_T^{n_1+1}] \\ &= \frac{P_0^{n_1} + \alpha_{n_1+1} D_{0T} - P_0^{n_1+1}}{\alpha_{n_1+1} D_{0T}} \\ &= \frac{(1 + \alpha_{n_1+1} y_0^{n_1+1}) P_0^{n_1+1} - P_0^{n_1+1}}{\alpha_{n_1+1} D_{0T}} \\ &= \frac{y_0^{n_1+1} P_0^{n_1+1}}{D_{0T}}.\end{aligned}$$

Note that the second to last equality holds by the fact that:

$$\begin{aligned} D_{0S_{n_1+1}} &= D_{0T} - y_0^{n_1+1} P_0^{n_1+1}, \\ P_0^{n_1+1} &= P_0^{n_1} + \alpha_{n_1+1} D_{0S_{n_1+1}}. \end{aligned}$$

Plugging the first expression into the second one and rearranging will yield the required identity. Note that the extended model satisfies the consistency condition (3.4).

4.2.2 The Alternate Fit and the Exact Fit

If we were to compute the price of a CMS payoff at different payment dates, we could use the 1F smcMFL approach to set up different partial models keeping the modelling choices at the reference swap rate maturity the same, but making different modelling choices at each of the payment dates. No model will price all the vanilla instruments in line with the market. Indeed, a model can lose explanatory value if its sole aim is to fit to many vanilla instruments. The approach of using different partial models is therefore in line with market practice.

However, in pricing a CMS, it is desirable not only to price the closely related vanilla instruments i.e. swaptions correctly but also to correctly price the fixed leg of the reference swap. If this does not hold for the modelling approach used, then the prices produced will lead to an arbitrage opportunity.

To see this, let the reference index be N, and let the payment dates be $S_1, S_2, \ldots, S_N$. For each $i = 1, \ldots, N$, let V_0^i denote the value assigned at time-zero to the payoff $y_T^N D_{TS_i}$ at time T by the i^{th} partial model and let V_0^{FIX} denote the value assigned to the fixed leg of the swap. Note that value V_0^{FIX} can be obtained in a model-independent way from the reference swaption prices. Now if $V_0^{FIX} > \sum_{i=1}^{N} \alpha_i V_0^i$ for example, we could buy the N individual CMS cashflows and sell the fixed leg of a forward starting swap, length N, starting at T to create an arbitrage. If we have $V_0^{FIX} = \sum_{i=1}^{N} \alpha_i V_0^i$, then we shall say that the set of partial models is *consistent*.

To achieve consistency, we could start off by making our modelling choices at time S_N. As we have seen above, this would fix the functional form of P_T^{N-1}. We jump two maturities before to S_{N-2} and we can make our modelling choices for y_T^{N-2} and P_T^{N-2}. We first observe that the functional relationship

$$D_{TS_{N-1}} = \frac{P_T^{N-1} - P_T^{N-2}}{\alpha_{N-1}}.$$

gives us a functional form for $D_{TS_{N-1}}$ consistent with the choices made at maturities S_N and S_{N-2}. We can fix the functional form of y_T^{N-1} to ensure the model is consistent (but in practice, we never need to determine this explicitly). We do so by defining y_T^{N-1}

as:

$$y_T^{N-1} := \frac{P_T^{N-2} + \alpha_{N-1} - P_T^{N-1}}{\alpha_{N-1} P_T^{N-1}}.$$

So, starting from S_N, we make modelling choices at each alternate time S_{N-2}, S_{N-4}, $\ldots$, to S_2, if N is even or S_1 if N is odd. We fix up the functional form of the swap rates at the times S_j,$(j \in \{N-3, N-5, \ldots\})$ where we do not have the flexibility to make free choices, using the relationship:

$$(1 + \alpha_j y_T^j) P_T^j = P_T^{j-1} + \alpha_j. \tag{4.19}$$

We shall refer to the framework designed to meet the consistency condition as described above as the *alternate fit* model. Observe that we have specified a single arbitrage-free model which models all the pure discount bonds D_{TS_i}, $i = 1, \ldots, N$. We could extend the model out to some payment date S_M when $M > N$ by fitting S_{N+2} and so on until $N + 2i \geq M$. To use the model to value a CMS payoff at S_M, we only need to fit the model at two times, S_N and S_M, if $|M - N|$ is even, or at three times, S_N, S_{M-1} and S_{M+1} if $|M - N|$ is odd. This contrasts with the single-time MFM developed in Chapter 2 where to model D_{TS_i}, we have to formulate the model from maturity S_1.

When the payment date exceeds the reference swap rate maturity, we no longer need the consistency condition to hold. We are free to do the partial model setup at any maturity, a setup we refer to as the *exact fit* approach. Note that this approach is arbitrage-free for any particular payment date. We shall observe however in the numerical analysis that the convexity corrections obtained from the exact fit setup are very close to that of the single-time MFM. From the alternate fit setup, at payment dates whereby we have no control over the model choices, we are forced to work with the distribution ascribed by the model, giving us prices that differ from that of the exact fit approach (and the single-time MFM). We shall show in Section 4.3.5 how we can refine the alternate fit approach to bring the model closer to the single-time MFM. In practice however, the exact fit approach will be sufficient.

4.3 Numerical Results

4.3.1 Model choices and functional forms

We use the same tenor structure setup as described in Section 3.2.1 of Chapter 3, and the initial conditions as given in table 1 in Appendix A.

For the numerical analysis that follows, we consider two possibilities for the market-implied marginal distributions of the swap rates under their respective swaption measures: log-normal and shifted log-normal. For the assumptions made, we now define explicit choices for the functional forms when setting the partial model up (i.e the model at a single maturity).

For a given maturity S_n, $n \in \{1, \ldots, \tilde{M}\}$, we postulate the following functional form for the swap rate in terms of the model driver η_n:

$$f_n(\eta_n) := \theta_n + \gamma_n \exp(\delta_n \eta_n), \tag{4.20}$$

where

$\theta_n \geq 0$ is the shift parameter

$\gamma_n = \exp(\log(y_0^n - \theta_n) - \frac{1}{2}\hat{\sigma}_n^2 T)$

$\delta_n = \hat{\sigma}_n \sqrt{T}$

Remark 22: Under the log-normal assumption, we take $\theta_n = 0$ and $\hat{\sigma}_n$ is the log-normal volatility as given in Appendix A. Under the shifted log-normal assumption, we take $\hat{\sigma}_n$ to be the implied volatility that yields the same ATM swaption prices under log-normal distribution. We tabulate the values used for the implied volatilities in this case in table 2 in Appendix C.

We recall that we chose to fix the distribution of η_n such that the model is calibrated to the market-implied distribution of the swap rate in its own measure. We can easily deduce from the given choice of functional form for y_T^n that η_n is standard Gaussian under $\mathbb{S}^n$ in either case.

The second choice involves postulating a functional form for $\hat{P}_T^n$ and we provided a general form in equation (4.4). We therefore need to specify $\hat{y}_T^i$, for $i \in \{1, \ldots, n\}$. We assumed the market-implied marginal distribution of the set of forward swap rates are either log-normal or shifted log-normal in their own swaption measures. Under $\mathbb{S}^n$, the distribution of the swap rates $\hat{y}_T^i$, $i \in \{1, \ldots, n-1\}$ will no longer be (shifted) log-normal, but it can be reasonably argued that it will still be close. That said, we choose to express $\hat{y}_T^i$ as follows:

$$\hat{y}_T^i(\eta_n) := \theta_i + (\hat{y}_0^i - \theta_i) \exp(\hat{\sigma}_i \sqrt{T} \eta_n - \frac{1}{2}\hat{\sigma}_i^2 T), \tag{4.21}$$

where $\hat{y}_0^i$ is the (possibly convexity-adjusted) forward swap rate.

A crude approximation for $\hat{y}_0^i$ is given in appendix C and we state the result below:

$$\hat{y}_0^i \approx \theta_i + (y_0^i - \theta_i)\exp\left(\hat{\sigma}_i\left[\left(\frac{D_{0S_n}}{P_0^n}\sum_{l=1}^{n}\frac{\alpha_l}{1+\alpha_l y_0^n} - 1\right)\frac{\hat{\sigma}_n(y_0^n - \theta_n)}{y_0^n} - \right.\right.$$
$$\left.\left.\left(\frac{D_{0S_i}}{P_0^i}\sum_{l=1}^{i}\frac{\alpha_l}{1+\alpha_l y_0^i} - 1\right)\frac{\hat{\sigma}_i(y_0^i - \theta_i)}{y_0^i}\right]T\right). \tag{4.22}$$

We observe from equation (4.22), that for $i = n$, $\hat{y}_0^n$ collapses to y_0^n and the choice of functional form for $\hat{y}_T^n$ matches the choice of functional form for y_T^n made above in equation (4.20). The functional form for P_T^n is thus given by:

$$P_T^n(\eta_n) := a_n \sum_{k=1}^{n} \alpha_k \left(\prod_{i=k}^{n} \frac{1}{1+\alpha_i \hat{y}_T^i(\eta_n)}\right),$$

where $a_n \in \mathbb{R}$ is appropriately chosen.

In the next section, we implement the model firstly under the assumption that the swap rates are log-normal distributed under their respective swaption measures. We also carry out a sensitivity test to see how the model fares if we omit the adjustment $\hat{y}_0^i$, and use y_0^i instead, for $i = \{1, \ldots, n\}$ when constructing the PVBP. We repeat the analysis with the shifted log-normal assumption.

4.3.2 Log-normal market-implied distributions

Following the steps outlined above, we set the models up under the assumption that the swap rates are log-normally distributed under their respective swaption measures (i.e. $\theta_j = 0$, for $j \in \{1, \ldots, \tilde{M}\}$). We compare the convexity corrections obtained from two model setups to the results from the single-time MFM and we observe how they perform as a function of the payment date. We carry out the analysis for $N = 2$ and $N = 10$.

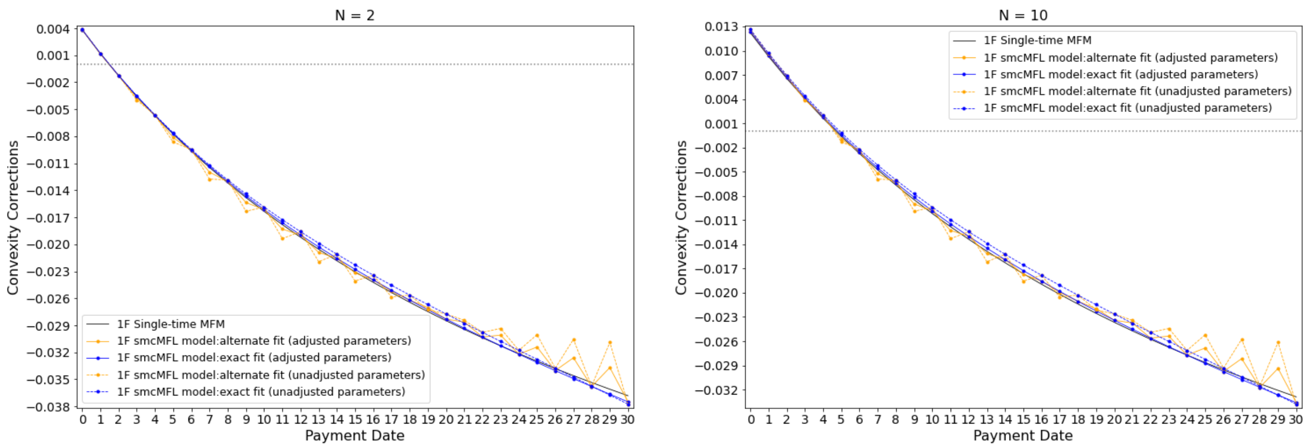

Figure 4.3: Convexity Corrections against payment date under log-normal assumptions

From Figure 4.3, we can observe that by doing the partial model setup at both the reference maturity and the payment date (i.e the exact fit), the results match closely to those from the single-time MFM. In that approach, we capture both market-implied marginal distribution of the swap rates that we have shown to be most significant for the pricing of the CMS. However, as we have discussed before, this model does not satisfy the consistency condition. On the other hand, for the alternate fit approach, we have a consistent, arbitrage-free model, but we can observe that at every odd payment date, we do not capture the distribution of the payment swap rate under its associated swaption measure, but rather, in order to achieve consistency, we set up the functional form for the swap rate based on modelling choices made one maturity step before and after. The distribution of the payment swap rate ascribed by the model differs from the log-normal assumption. This translates to the convexity corrections given by the alternate fit model differing relatively significantly, as the payment date increases, from the other models (exact fit and single-time MFM). In Chapter 5, we revisit the *zig-zag* behaviour we observe in the alternate fit approach and provide an explanation for its occurrence. In Section 4.3.5, we propose a method of refining the results from the alternate fit model to achieve a closer match to the single-time MFM. We observe that using the convexity-adjusted parameters improves the results slightly - we can see that the size of the oscillations decreases for large payment dates.

Remark 23*: Note that the oscillations of the alternate fit model in the convexity correction increase with the payment date index M. The corresponding oscillations of the convexity-adjusted value of a CMS payment would be much less (because the value is obtained from the convexity adjusted forward by multiplying by the discount factor D_{0S_M} - which decreases with M).*

4.3.3 Shifted log-normal market-implied distributions

We repeat the previous analysis but this time we assume that the market implied distribution of the swap rates are shifted log-normally distributed under their respective swaption measures. We assume that the shift parameter is -0.05, unless otherwise specified. Note that we choose the volatility parameter in the shifted log-normal case such that the prices of ATM swaptions match those obtained under the assumption of log-normal distribution. We assume that the prior swap rates we use to construct the functional form $\hat{P}^{\cdot}_{T}$ have shifted log-normal forms.

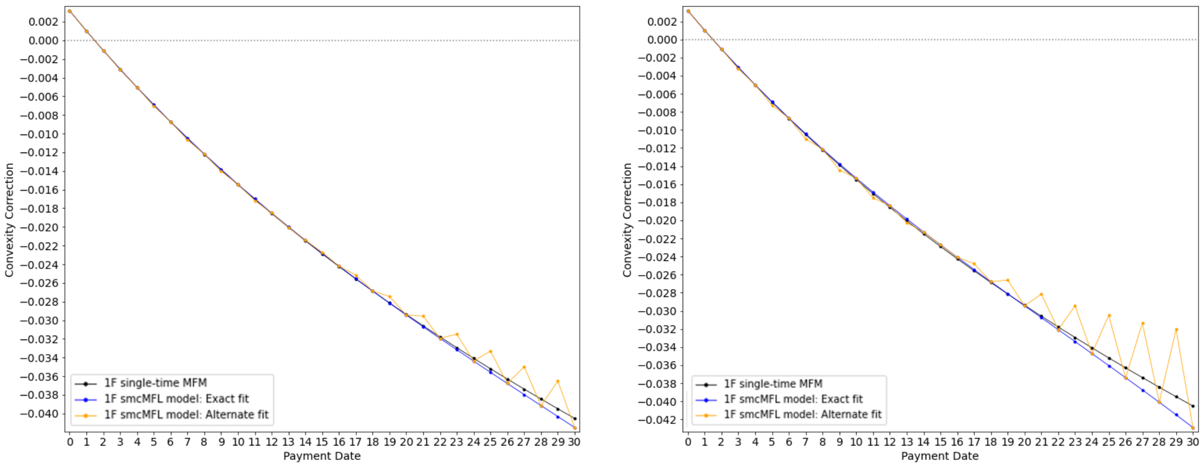

Figure 4.4: Convexity Corrections against payment date under shifted log-normal assumption; N = 2
Left: using convexity-adjusted parameters as given in equation (4.22)
Right: using original parameters y_0^k *in prior swap rates specification for the PVBP*

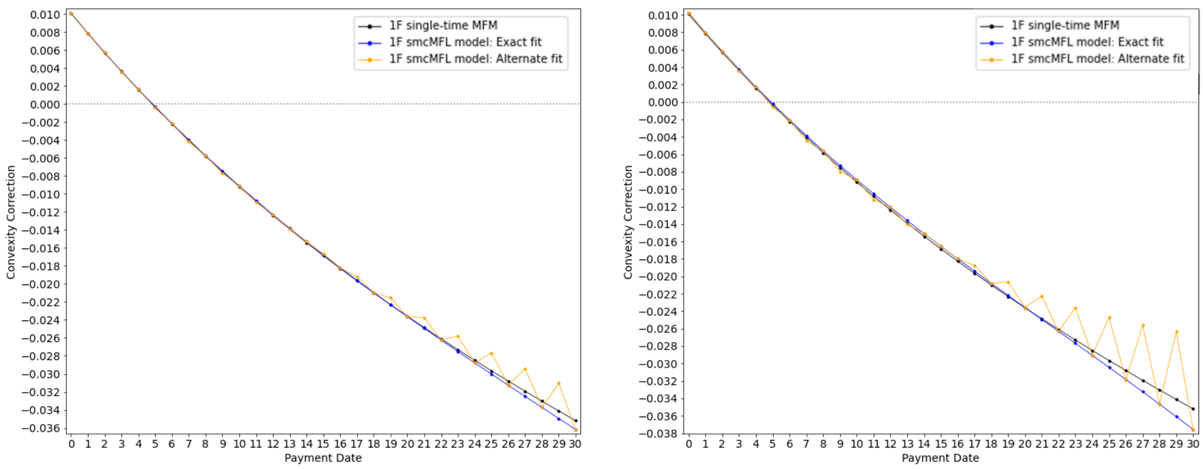

Figure 4.5: Convexity Corrections against payment date under shifted log-normal assumption; N = 10
Left: using convexity-adjusted parameters as given in equation (4.22)
Right: using original parameters y_0^k *in prior swap rates specification for the PVBP*

We can observe that both the *alternate fit* model and the *exact fit* model perform remarkably well up to payment date S_{16} in both cases. As the payment date then increases, we can see that the *exact fit* model gives results close to those obtained from the single-time MFM, while the discrepancy in *alternate fit* model at odd payment dates increases. It should be stressed however than when $S_M > S_N$, we can focus on the exact fit approach as we do not need the consistency conditions to hold.

4.3.4 Sensitivity of model to choice of functional forms

In the previous sections, we have set up functional forms for the prior swap rates $\hat{y}_T^l$ that match closely what we believe their market implied distributions are. In particular, recall we constructed the functional form $\hat{P}_T^n$, for $n \in \{1, \dots, \tilde{M}\}$, as follows:

$$\hat{P}_T^n := \sum_{k=1}^{n} \alpha_k \Big(\prod_{l=k}^{n} \frac{1}{1 + \alpha_l \hat{y}_T^l} \Big).$$

We chose to define $\hat{y}_T^l$ based on our assumption for the market implied distributions.

In order to determine how sensitive the models are to the choice of functional forms, we now choose functional forms for the prior swap rates such that they are modelled differently from their actual market-implied distributions. In the numerical analysis, we assume that the market implied distributions of swap rates under their respective swaption measures are shifted log-normal, but we define $\hat{y}_T^l$ using the log-normal form. We henceforth refer to this setting as the *SLN-LN* case. Note that we are using same model input as in the previous section.

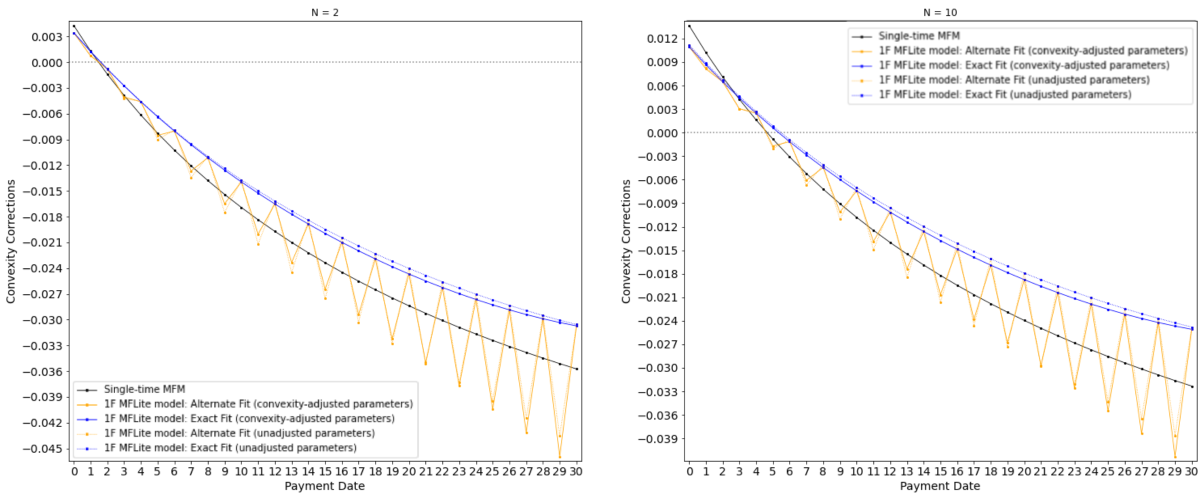

Figure 4.6: Convexity Corrections against payment date: Sensitivity of model to choice of functional forms

In Figure 4.6, we observe firstly that there is minimal difference between the results obtained using the convexity-adjusted parameters $\hat{y}_0^l$ and those obtained using the unadjusted y_0^l, for $l \in \{1, \dots, n\}$, in the specification of the prior swap rates for the PVBP P_T^n. Secondly, as the payment date increases, we notice a divergence between the exact fit model and the single-time MFM. This points us to believe that the PVBP plays a crucial role in the pricing problem, a point we will come to again when discussing the two-factor model setup. This analysis serves as a preliminary indication of the importance of modelling the PVBP appropriately. This idea is further reinforced when we look at the performance of the alternate fit model. Compared to the previous cases (Figures 4.3 - 4.5), we observe that even at early payment dates, the discrepancy between the convexity corrections at even and odd payment dates vary significantly, resulting in a *zig-zag* behaviour. As mentioned before, we shall explore a method later on to refine the alternate fit approach, and we will discuss in depth the cause of such behaviour in the next chapter.

We end this section with the observation that in the SLN-LN case, we modelled the prior swap rates differently from their actual market-implied distributions. In particular, for a given maturity S_n, the choice of functional forms for $\hat{y}^l_T$ (used in setting up the PVBP and y^n_T (the swap rate we aim to model appropriately) do not coincide. We discuss below a possible adjustment that could be made in order to account for this mismatch.

We note that a limitation of making such choices is that the functional forms used for $\hat{y}^l_T$ do not allow for negative rates (by the assumption of shifted log-normal distribution for the swap rates, there is a possibility of negative rates if we choose to set the shift parameter to be negative). So in the numerical analysis that follows, we use a positive shift parameter in order to observe how sensitive the models are to a mismatch of distribution in the choice of functional forms.

We want to reflect the knowledge of the market-implied distribution of y^n_T in the construction of the PVBP. To take this into account, we introduce a dummy variable $\hat{\eta}$ and we fix its distribution under $\mathbb{S}^n$ as follows:

$$\hat{\eta} := (\hat{y}^n_T)^{-1}(g_n(X)),$$

where $X \sim \mathcal{N}(0,1)$ under $\mathbb{S}^n$,

$$\hat{y}^n_T(x) = y^n_0 \exp(\sigma_n \sqrt{T} x - \sigma^2_n T),$$

and

$$g_n(X) = \theta_n + (y^n_0 - \theta_n) \exp(\hat{\sigma}_n \sqrt{T} X - \frac{1}{2} \hat{\sigma}^2_n T).$$

We note that for $\hat{y}^n_T$, we use the log-normal volatility given in Appendix A, and for the specification of g_n, we use the implied volatility as given in Appendix C.

Remark 24: In setting up the functional form for $\hat{P}^n_T$, we set the functional forms for the prior swap rates as follows:

- *For $x \in \mathbb{R}$, $\{\hat{y}^1_T(x), \dots, \hat{y}^n_T(x)\}$ are chosen to be of log-normal form*
- *Define $\hat{x} := (\hat{y}^n_T)^{-1}(g_n(x))$*
- *Define a set of modified functional forms for the prior swap rates $\hat{y}^{n:i}_T(x), i \in \{1, \dots, n\}$, where $\hat{y}^{n:i}_T(x) := \hat{y}^i_T(\hat{x}) = \hat{y}^i_T\big((\hat{y}^n_T)^{-1}(g_n(x))\big)$. Then $\hat{y}^{n:n}_T$ has the right market implied distribution under $\mathbb{S}^n$. We use these modified functional forms for the swap rates to construct $\hat{P}^n_T$, hence the PVBP P^n_T.*

For the numerical results below, we assume the shift parameter takes value 0.02. We look at how convexity corrections behave as a function of payment time and how the models compare.

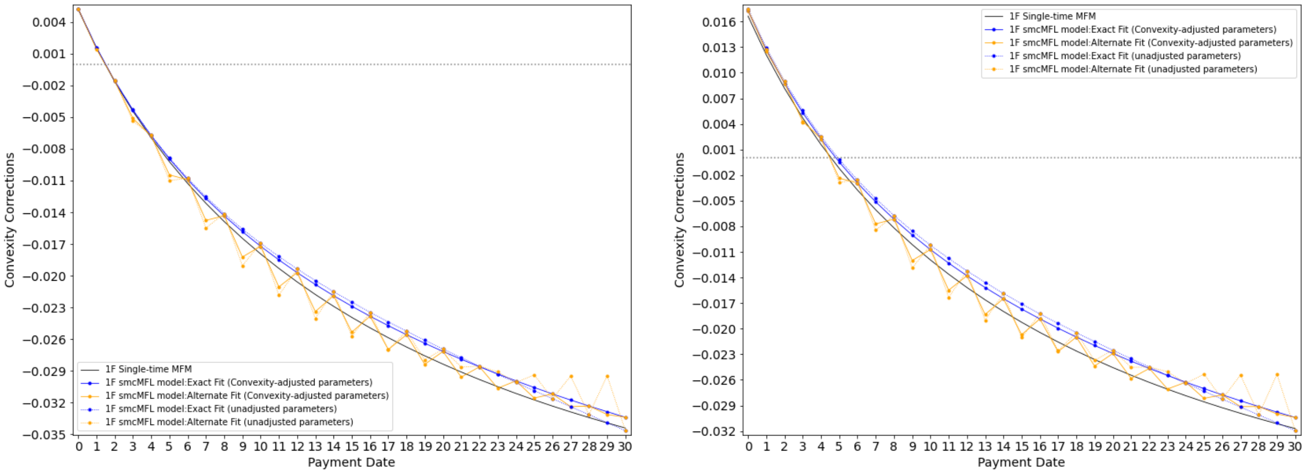

Figure 4.7: Convexity Corrections against payment date: SLN-LN analysis using refined functional forms

We observe from Figure 4.7, that the refinement introduced when specifying the PVBP has reduced the oscillatory behaviour of the alternate fit model significantly compared to the analogous results in Figure 4.6. We further observe that the convexity corrections from the exact fit approach are now closer to that of the single-time MFM. Similar to Figure 4.6, we observe that using the convexity-adjusted values $\hat{y}_0^l$ does not have much of an impact on the model results. There is a slight deterioration in the alternate fit results at odd payment dates very far from the reference swap rate maturity, but as we have argued before, for these payment dates, we can use the exact fit approach instead.

4.3.5 Refining the alternate fit setup of the 1F smcMFL Model

The numerical analysis of the alternate fit setup of the MF-Lite model revealed a 'zig-zag' behaviour - at even payment dates where we can calibrate the model to the desired marginal distribution (we choose the reference swap rate maturity to be even), the convexity correction is close to that of the single-time MFM; at alternate (odd) payment dates, whereby the functional forms are fixed, the convexity correction is either too high or too low compared to that of the single-time MFM. We provide below a method to refine the alternate fit setup. Using numerically observed information on the swap rates from the single-time MFM under the forward measure $\mathbb{F}$, we fine-tune the parameters used in setting up the PVBP and update the functional form in the 1F smcMFL approach to closely reproduce the single-time MFM results. The refined alternate fit setup described here is only meant to demonstrate that we can construct an arbitrage-free consistent MF-Lite approach that can achieve a close match to an arbitrage-free model.

For the numerical results in this section, we focus on the assumption that the swap rates are log-normally distributed under their respective swaption measures, but the approach should carry over to a broader range of marginal distributions. In Chapter 5, we revisit the zig-zag behaviour in more depth and provide an explanation as to why we observe this in the alternate fit setup. In the next section, we discuss how we can use numerically observed information on the swap rates under the forward measure to alter the PVBP in the 1F smcMFL approach to reproduce the single-time MFM results. We stress that in practice, firstly the consistency condition is only relevant when the payment date is less than or equal to the maturity date of the reference swap. As we have seen in the numerical results above, the prices from the alternate fit setup are very close to those of the exact fit, that in practice, we could use the exact setup instead.

Refining the volatilities

We have assumed that the swap rates are log-normally distributed under their associated swaption measures. Consider the single-time MFM under $\mathbb{F}$. While the swap rates will no longer be log-normal, we will still assume them to be roughly so. We can express:

$$\log(y_T^i) \approx \Sigma_i Y + C_i, \tag{4.23}$$

where $Y \sim \mathcal{N}(0,1)$ under $\mathbb{F}$, for $i \in \{1,\ldots,\tilde{M}\}$. We define, for given $i,j \in \{1,\ldots,\tilde{M}\}$:

$$\begin{aligned} r_{j,i} &:= \frac{\mathrm{var}(\log(y_T^j))}{\mathrm{var}(\log(y_T^i))} \\ &= \frac{\Sigma_j^2 \mathrm{var}(Y)}{\Sigma_i^2 \mathrm{var}(Y)} \\ &= \frac{\Sigma_j^2}{\Sigma_i^2}. \end{aligned} \tag{4.24}$$

Note that if we change to a different measure, the distribution of Y will change but the functional forms (4.23) will not. In particular, the ratio $r_{j,i}$ is measure-independent. Fix $n \in \{1,\ldots,\tilde{M}\}$. Under the swaption measure $\mathbb{S}^n$, y_T^n is exactly log-normal with volatility σ_n. For the current alternate fit setup, we have used the log-normal volatilities for the prior swap rates $\hat{y}_T^i$ for $i < n$ in forming the PVBP P_T^n. Instead, we now adjust the volatilities so that the ratio (4.24) holds. For $i \in \{1,\ldots,n\}$, we define:

$$\sigma_{n:i} := \sigma_n \sqrt{r_{n,i}}. \tag{4.25}$$

Note that we can get hold of the ratios numerically from the single-time MFM. Going back to the 1F smcMFL model, we now have:

$$y_T^n(\eta) := y_0^n \exp(\sigma_n \sqrt{T}\eta - \frac{1}{2}\sigma_n^2 T),$$

$$\hat{P}_T^n(\eta) := \sum_{k=1}^{n} \alpha_k \prod_{i=k}^{n} (1 + \alpha_i (\bar{y}_T^{n:i}(\eta))^{-1},$$

where

$$\bar{y}_T^{n:i}(\eta) := y_0^i \exp(\sigma_{n:i} \sqrt{T}\eta - \frac{1}{2}\sigma_{n:i}^2 T).$$

We now look at the convexity correction against payment time, using the refined PVBP functional form.

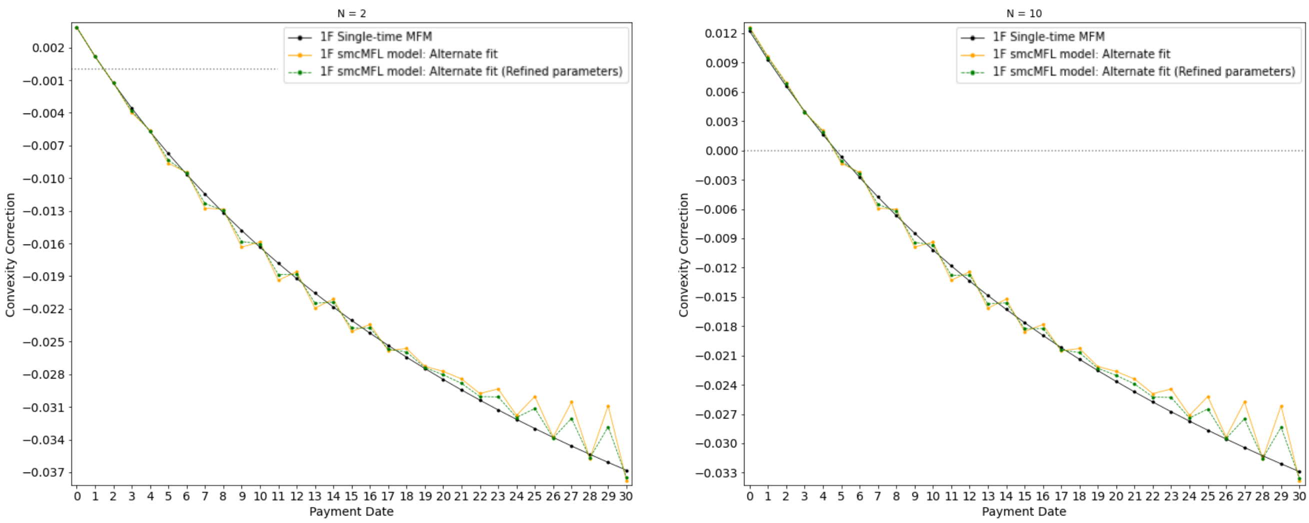

Figure 4.8: Convexity correction against payment date: Refining the volatilities of the prior swap rates

We observe a slight improvement in the alternate fit setup, but the 'zig-zag' behaviour is still present. We provide in the next section a further modification to the functional form of the PVBP in order to refine the model results.

Refining the initial value of the swap rates

In the previous section, for a fixed $n \in \{1, \ldots, \tilde{M}\}$ we refined the volatilities of the (prior) swap rates $\{\hat{y}_T^i : i \in \{1, \ldots, n-1\}\}$. There has been an improvement in the results obtained from the alternate fit setup, but we can still further refine the parameters y_0^i to bring the model closer to the single-time MFM.

We know that:

$$\begin{aligned} y_T^n &= y_0^n \exp(\sigma_n \sqrt{T}\eta - \frac{1}{2}\sigma_n^2 T) =: \exp(\sigma_n \sqrt{T}\eta + b_n) \\ &= \exp(\sigma_{n:n}\sqrt{T}Y + \hat{b}_n), \end{aligned}$$

where $\eta \sim \mathcal{N}(0,1)$ under $\mathbb{S}^n$, $Y \sim \mathcal{N}(0,1)$ under $\mathbb{F}$ (roughly). Note that the $\hat{b}_n \in \mathbb{R}$ can be (numerically) worked out from the single-time MFM.

From the above, we can find an expression for the variable Y in terms of η:

$$Y = \frac{\sigma_n}{\sigma_{n:n}}\eta + \frac{b_n - \hat{b}_n}{\sigma_{n:n}\sqrt{T}}. \tag{4.26}$$

For $i < n$, we also have that:

$$\bar{y}_T^{n:i}(Y) = y_0^i \exp(\sigma_{n:i}\sqrt{T}Y - \frac{1}{2}\sigma_{n:i}^2 T) =: \exp(\sigma_{n:i}\sqrt{T}Y + \hat{b}_i). \tag{4.27}$$

Substituting the expression for Y from (4.26) in equation (4.27) yields

$$\begin{aligned} \bar{y}_T^{n:i} &= \exp(\sigma_{n:i}\sqrt{T}Y + \hat{b}_i) \\ &= \exp\left(\sigma_{n:i}\sqrt{T}\left[\frac{\sigma_n}{\sigma_{n:n}}\eta + \frac{b_n - \hat{b}_n}{\sigma_{n:n}\sqrt{T}}\right] + \hat{b}_i\right) \\ &= \exp\left(\sigma_n \frac{\sigma_{n:i}}{\sigma_{n:n}}\sqrt{T}\eta + (b_n - \hat{b}_n)\frac{\sigma_{n:i}}{\sigma_{n:n}} + \hat{b}_i\right) \\ &=: \exp(\bar{\sigma}_{n:i}\sqrt{T}\eta + \bar{b}_i), \end{aligned}$$

where

$$\begin{aligned} \bar{\sigma}_{n:i} &= \sigma_n \frac{\sigma_{n:i}}{\sigma_{n:n}}, \\ \bar{b}_i &= (b_n - \hat{b}_n)\frac{\hat{\sigma}_i}{\hat{\sigma}_n} + \hat{b}_i. \end{aligned}$$

We have expressed $\bar{y}_T^{n:i}$ in terms of a standard normal random variable under $\mathbb{S}^n$ with updated parameters. We use these functional forms to define the PVBP.

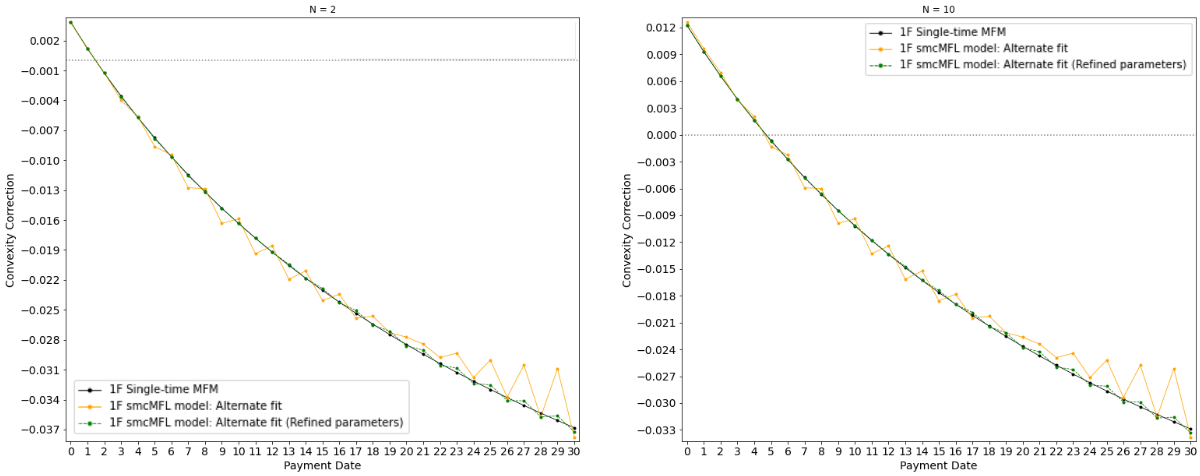

Figure 4.9: Convexity correction against payment date: using refined priors for the PVBP

As we noted earlier, the purpose of the above analysis is to demonstrate that a more careful choice of PVBP improves the oscillatory behaviour we observe in the alternate fit model, addressing the undesirable feature of the setup.

CHAPTER 5

A Markov-functional approach to convexity corrections: The Two-factor MF-Lite model

In Chapter 3, we used the single-time MFM developed in the earlier chapter to investigate the impact of the (joint) distribution of the swap rates on convexity corrections when pricing a CMS payment (or related options). We observed that, as the payment date gets further away from the reference swap rate maturity, the marginal distribution of the payment swap rate has a material effect on convexity corrections, in addition to the distribution of the reference swap rate. In Chapter 4, we set out to construct a practical one-factor MF-Lite approach to the pricing problem, that would take into account this observation. Going back to the numerical analysis of Chapter 3, under log-normal assumptions on the swap rates in their own swaption measures, and imposing a Gaussian copula on the model drivers, we investigated the impact of the correlation between the log of the swap rates on convexity corrections, pointing us to believe that a two-factor model would be appropriate for the pricing of a CMS (in particular when the payment date is far from the reference swap rate maturity). To this end, the goal of this chapter is to incorporate a second factor in the MF-Lite approach.

We would naturally want to extend the approach developed in Chapter 4 to the two-factor case. We recall then we started off by postulating functional forms for the swap rate and the PVBP in terms of some (one-dimensional) model driver under the swaption measure. We then carried the model over to the forward measure. We could repeat this procedure for a different maturity. We then unified the two partial model setups by observing that under the common forward measure $\mathbb{F}$, under the one-factor assumption, there exists a functional relationship between the two drivers. For the two-factor setup, we could potentially take a similar approach. However, it proved to be difficult to set up a simple, yet efficient two-factor model that is structurally similar to the 2F single-time MFM when starting the modelling setup under the swaption measure. We discuss this point further in the later Section 5.3.

Instead, for the development of the two-factor MF-Lite approach, we start the modelling process under the forward measure. In Section 5.1, we show how we construct a two-factor forward measure calibrated MF-lite model, which we shall abbreviate to 2F fmcMFL model. The danger with this approach is that we have no direct market data that would inform the choices we make under the forward measure. To elaborate, recall we construct a functional form for the PVBP by choosing a set of priors for the forward swap rates. Assume the swap rates are log-normally distributed with respect to their own swaption measures, with some known log-normal volatilities. Under the forward measure, the distribution will not be log-normal, but we could still make the naive choice of log-normal functionals with the given volatility (known only under the swaption measure) for the swap rates under $\mathbb{F}$. This idea is reflected in the choices that Cedervall and Piterbarg (2012) make when setting their model up under the forward measure, but they do not examine the consequences of such choices on their model performance. However, as we shall see in Section 5.1.1, an analysis of the fmcMFL model in its simpler one-factor setup reveals that the model fails to accurately replicate the single-time MFM results. The PVBP plays an important role in bridging the model between the forward measure and the swaption measure, and that variable should be modelled appropriately. In Section 5.2, we propose a refinement on the functional form of the PVBP (hence a 2F *refined* fmcMFL model) that will bring the 2F MF-Lite approach closer to the 2F single-time MFM.

In Section 5.3, we propose a simplistic/naive 2F MF-Lite model set up under the swaption measure. As we have touched upon above, setting up a two-factor MF-Lite approach in the swaption measure, whilst retaining the structure of the single-time MFM is particularly challenging. The approach we propose in Section 5.3 strips away much of the model complexity of the single-time MFM. The downside is that the choices of functional forms are suboptimal, which could result in the model being numerically far from the single-time MFM. The numerical results under Gaussian assumptions however show that if the model is properly calibrated, it can still closely replicate the single-time MFM. The model does not satisfy the consistency conditions we discussed in Section 4.2.2 of Chapter 4, but this is not an issue if the payment date is after the reference swap rate expiry.

5.1 A two-factor forward measure calibrated MF-Lite model

In this section, we consider a 2F MF-Lite model set up under the forward measure $\mathbb{F}$, which for brevity we refer to as the 2F fmcMFL model. The approach provides a two-factor computationally fast model, but for which prices are not close to the single-time MFM. It is used to highlight the potential issues that can arise when one is not careful enough in the modelling choices when setting the model up with respect to a measure under which we have no market data to inform the distributional assumptions we make on the forward swap rates and therefore the PVBP. We shall explore this point in more depth in the numerical analysis in Section 5.1.1 by observing the performance of the 1F fmcMFL model under the assumption that the swap rates are log-normally distributed under their respective swaption measures. Note that we go back to the simpler one-factor setup of the model for this analysis. Our goal is to examine how the proposed approach fares in a simplified context. We propose a refinement to the functional form of the PVBP in Section 5.2 to account for this model deficiency.

We describe the 2F fmcMFL model setup for a single maturity which we denote by S_n. As for the 1F smcMFL model developed in Chapter 4, we set up partial models by specifying the functional forms for the swap rates and the PVBPs at two maturities, and we unify the partial setups under the common measure $\mathbb{F}$. Here we are starting the modelling process in the forward measure so these functional forms will already be in terms of a common two-dimensional driver which we denote by $(x_T^{(1)}, x_T^{(2)})$. Although we are describing the model for a given maturity, the same steps can be followed for a second maturity.

Our first task is to set up a model for the PVBP P_T^n. In contrast to the swaption MF-Lite model of Chapter 4, we need the PVBP before we are able to carry out the calibration step to match the distribution of the swap rate y_T^n to its market-implied distribution under $\mathbb{S}^n$. Assume under the forward measure $\mathbb{F}$, associated with taking $D_{\cdot T}$ as numéraire,

$$\begin{pmatrix} x_T^{(1)} \\ x_T^{(2)} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \right). \tag{5.1}$$

We note the similarity in the choice of driver between the 2F fmcMFL model and the 2F single-time MFM. We aim to bring the two models as close as possible, so we borrow the structure of the single-time MFM to inform the choices we make in the 2F fmcMFL model. For $i \in \{1, \dots, n\}$ and for $\beta_i^{(1)}, \beta_i^{(2)} \in \mathbb{R}$, we define:

$$z_T^i := \frac{\beta_i^{(1)} x_T^{(1)} + \beta_i^{(2)} x_T^{(2)}}{\sqrt{(\beta_i^{(1)})^2 \lambda_1 + (\beta_i^{(2)})^2 \lambda_2}}. \tag{5.2}$$

We postulate a prior model for the swap rates, which we denote by $\hat{y}_T^i$, in terms of z_T^i. We define:

$$\hat{y}_T^i := \hat{f}_i(z_T^i). \tag{5.3}$$

The function $\hat{f}_i : \mathbb{R} \to \mathbb{R}$ is a known deterministic function that reflects our belief about the marginal distribution of the i^{th} swap rate in its own swaption measure if z_T^i were $\mathcal{N}(0,1)$ in that measure. For example, if swaption prices based on y_T^i were given by Black's formula with implied volatility σ_i, then we would postulate:

$$\hat{f}_i(z_T^i) := y_0^i \exp(\sigma_i \sqrt{T} z_T^i - \frac{1}{2}\sigma_i^2 T).$$

Using the standard functional relationship as we have done in Section 4.1 of Chapter 4, we can construct a prior model for the PVBP $\hat{P}_T^n$ as follows:

$$\begin{aligned} \hat{P}_T^n(x_T^{(1)}, x_T^{(2)}) &= \hat{P}_T^n(z_T^1, z_T^2, ..., z_T^n) \\ &:= \sum_{k=1}^{n} \alpha_k \left(\prod_{i=k}^{n} \frac{1}{1 + \alpha_i \hat{f}_i(z_T^i)} \right). \end{aligned} \tag{5.4}$$

We need to ensure that the no-arbitrage condition holds under $\mathbb{F}$. In particular, since P^n is a martingale under $\mathbb{F}$, we would want the equality:

$$\mathbb{E}_{\mathbb{F}}[P_T^n] = \frac{P_0^n}{D_{0T}}$$

to hold. Hence, we adjust the prior by defining $P_T^n := a_n \hat{P}_T^n$, where:

$$a_n := \frac{P_0^n}{D_{0T} \mathbb{E}_{\mathbb{F}}[\hat{P}_T^n(x_T^{(1)}, x_T^{(2)})]}. \tag{5.5}$$

The constant term $a_n \in \mathbb{R}$ fixes up the no-arbitrage condition.

It remains at this stage to construct a functional form for the swap rate. We do so by calibrating the model to the known marginal distribution of the swap rate y_T^n under its associated swaption measure using the *Markov functional sweep* technique. We assume we can model the swap rate y_T^n as a monotonic increasing function f_n of z_T^n. Below, we show how we recover f_n from swaption/digital swaption prices.

For a given strike $K \in \mathbb{R}$, denote the market price of a payer's swaption by $V_0^n(K)$.

Within the model, this is given by:

$$\begin{aligned} V_0^n(K) &= D_{0T} \mathbb{E}_{\mathbb{F}}[P_T^n (y_T^n - K)_+] \\ &= D_{0T} \mathbb{E}_{\mathbb{F}}[P_T^n (y_T^n - K) \mathbb{1}\{y_T^n > K\}]. \end{aligned}$$

Differentiating with respect to K, we obtain:

$$\begin{aligned}\mathcal{D}_0^n(K) &:= -(V_0^n(K))' \\ &= D_{0T}\mathbb{E}_{\mathbb{F}}[P_T^n \mathbb{1}\{y_T^n > K\}] \\ &= D_{0T}\mathbb{E}_{\mathbb{F}}[P_T^n \mathbb{1}\{f_n(z_T^n) > K\}]. \end{aligned} \tag{5.6}$$

We define and calculate for a given $z^* \in \mathbb{R}$:

$$\tilde{J}_0^n(z^*) := D_{0T}\mathbb{E}_{\mathbb{F}}[P_T^n \mathbb{1}\{z_T^n > z^*\}].$$

We observe that we can get hold of the function $\tilde{J}_0^n$ using the model assumptions on the drivers and the postulated functional form of the PVBP. By the monotonicity assumption of f_n, we can find a unique $K^* \in \mathbb{R}$, such that the set identity holds:

$$\{z_T^n > z^*\} = \{y_T^n > K^*\}.$$

Given z^*, we compute:

$$\begin{aligned}\tilde{J}_0^n(z^*) &= D_{0T}\mathbb{E}_{\mathbb{F}}[P_T^n \mathbb{1}\{z_T^n > z^*\}] \\ &= D_{0T}\mathbb{E}_{\mathbb{F}}[P_T^n \mathbb{1}\{f_n(z_T^n) > f_n(z^*)\}]. \end{aligned} \tag{5.7}$$

For a given z^*, comparing equation (5.6) to equation (5.7), we note that finding that unique K^* such that the above set identity holds is equivalent to knowing $f_n(z^*)$ - so we know the function f_n numerically.

We have now calibrated the model to the known marginal distribution of y_T^n under the swaption measure $\mathbb{S}^n$. We note that we have used the functional form postulated for the PVBP P_T^n in the calibration step and we also have a functional form for the swap rate y_T^n. This completes the model specification, and as we have seen before, we can derive the functional form of the pure discount bond D_{TS_n}:

$$\begin{aligned}D_{TS_n}(x_T^{(1)}, x_T^{(2)}) &= 1 - y_T^n(x_T^{(1)}, x_T^{(2)}) P_T^n(x_T^{(1)}, x_T^{(2)}) \\ &= 1 - a_n f_n(z_T^n)\hat{P}_T^n(x_T^{(1)}, x_T^{(2)}).\end{aligned}$$

We recall from Section 4.1 of Chapter 4, that making two model decisions for a given maturity in the tenor (recall we choose the functional form of y_T^n and P_T^n) limits our flexibility to make free choices one maturity earlier. Note that the same issue will arise here. When doing the partial model setup to another maturity, we proposed the *alternate fit* model, discussed in Section 4.2.2, that takes into account the consistency conditions that the model has to satisfy for no-arbitrage. For maturities beyond the reference swap rate maturity, we proposed the *exact fit* approach. We can use the same methodology in the two-factor case.

Remark 25 (Comparison of the 1F smcMFL model and the 1F fmcMFL model)*: If we take $\beta_i^{(2)} = 0$, then $z_T^i = \frac{x_T^{(1)}}{\sqrt{\lambda_1}} =: x$ for all i and the 2F fmcMFL model collapses to a one-factor model. The functional form for the PVBP in both the swaption and the forward MF-Lite models is the same, but in the forward model, it is a function of x, which has a $\mathcal{N}(0,1)$ distribution under $\mathbb{F}$, and in the swaption model, it is a function of the driver η_n, whereby $\eta_n = h_n(\tilde{x})$, where $\tilde{x} \sim \mathcal{N}(0,1)$ under $\mathbb{F}$. As we shall see in the next section, the function h_n will not be the identity function and this leads to problems with the approach outlined above.*

5.1.1 Numerical investigation of a 1F fmcMFL model versus a 1F smcMFL model

We assume that the forward swap rates are log-normally distributed under their respective swaption measures. We are using the same tenor structure setup, initial conditions and model input as laid out in Section 3.2.1 and related appendix A. The analysis in this section is carried out using the less complex one-factor model setup, taking $\beta_i^{(1)} = 1$ and $\beta_i^{(2)} = 0$, for $i \in \{1, ..., \tilde{M}\}$. We take $\lambda_1 = 1$ and $\lambda_2 = 0$. We use the 1F smcMFL model, developed in Chapter 4 as a benchmark model to which we compare the 1F fmcMFL model. We shall henceforth refer to the former simply as the *benchmark model.* The choices we make for the partial model setup for a given maturity S_n of the 1F fmcMFL model is as follows (Note that based on the above parameter choices, we have a one-dimensional driver, which we simply denote by $x_T \sim \mathcal{N}(0,1)$ under $\mathbb{F}$):

$$
\begin{aligned}
y_T^n(x_T) &:= y_0^n \exp(\sigma_n \sqrt{T} x_T - \frac{1}{2}\sigma_n^2 T) \\
P_T^n(x_T) &:= a_n \sum_{k=1}^{n} \alpha_k \left(\prod_{j=k}^{n} (1 + \alpha_j \hat{y}_T^j(x_T))^{-1} \right), \qquad a_n \in \mathbb{R}
\end{aligned}
$$

where $\hat{y}_T^j(x_T) = y_0^j \exp(\sigma_j \sqrt{T} x_T - \frac{1}{2}\sigma_j^2 T)$.

We look at the convexity correction when pricing single payment of the CMS as a function of payment date.

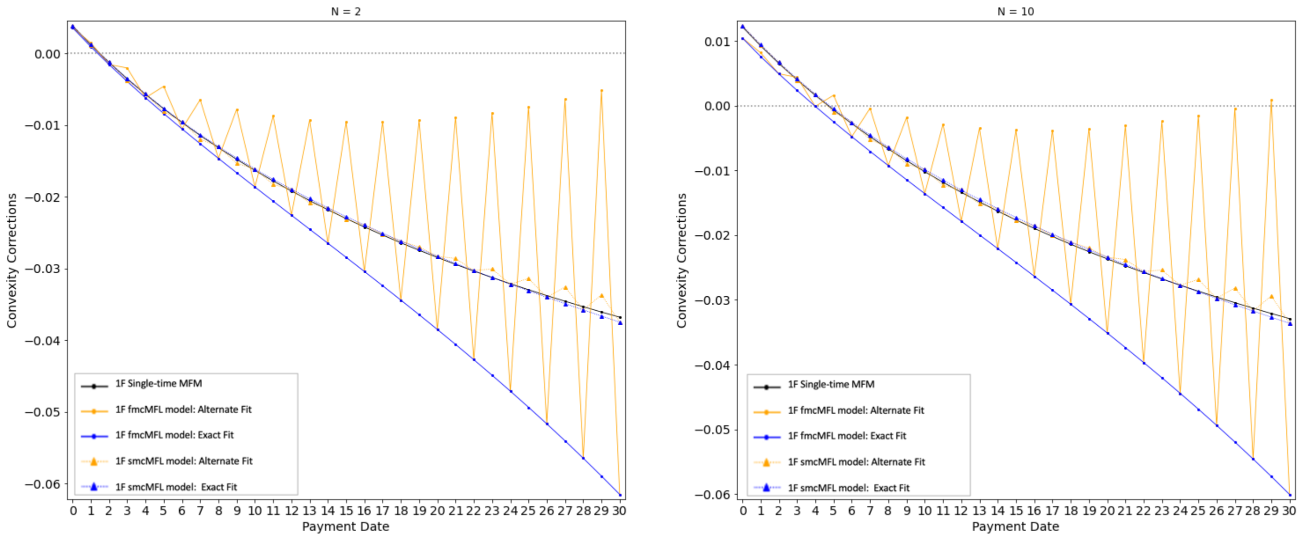

Figure 5.1: Convexity Correction against Payment Date; Left: N = 2, Right: N = 10

Looking at Figure 5.1, we observe that the exact fit setup of the 1F fmcMFL model consistently underestimates the convexity corrections (the results are too low compared to that obtained from the 1F single-time MFM). On the other hand, for the alternate fit setup, we observe a striking jaggedness. There are two components: one is the convexity correction whereby we have control over the model choices, i.e at even payment dates, hence the results coincide with the exact fit setup, and is therefore too low; at odd payment maturities, whereby we have no control over the model choices, we observe that the convexity correction is too high relative to the single-time MFM. In the next Section, we investigate this particular behaviour. We note that the (much less pronounced) zig-zag behaviour also appeared in the numerical results in the 1F smcMFL model. We showed how we could refine the model to decrease the size of the oscillations by tuning in the parameters used in setting up the PVBP. We will discuss below how to modify the PVBP to address the more extreme behaviour we observe here.

Remark 26: We refer back to remark 23 and we note that the observation pointed out on the oscillatory behaviour for the alternate fit model still applies here, but it does not fully explain the wild oscillations. We shall investigate this below.

Investigating the 1F fmcMFL model

We recall in the benchmark model developed in Section 4.1, we started off by postulating functional forms for the swap rate $y^n_T(\eta_n)$ and the PVBP $P^n_T(\eta_n)$ under log-normal assumption as follows:

$$y^n_T(\eta_n) = y^n_0 \exp(\sigma_n \sqrt{T}\eta_n - \frac{1}{2}\sigma^2_n T),$$

$$P_T^n(\eta_n) = a_n \sum_{k=1}^{n} \alpha_k \left(\prod_{j=k}^{n} \frac{1}{1+\alpha_j \hat{y}_T^j(\eta_n)} \right), \quad \text{where} \tag{5.8a}$$

$$\hat{y}_T^j(\eta_n) = y_0^j \exp(\sigma_j \sqrt{T}\eta_n - \frac{1}{2}\sigma_j^2 T). \tag{5.8b}$$

The modelling choice we make under the assumption that the swap rate y_T^n is log-normally distributed in its own swaption measure results in the driver η_n being standard Gaussian under $\mathbb{S}^n$. We then carry the model over to the forward measure. We determine the distribution of η_n under $\mathbb{F}$ as follows; for $x \in \mathbb{R}$:

$$F_{\eta_n}(x) := \mathbb{E}_{\mathbb{F}}[\mathbb{1}(\eta_n \leq x)] = \frac{P_0^n}{D_{0T}} \mathbb{E}_{\mathbb{S}^n}\left[\mathbb{1}(\eta_n \leq x)\frac{1}{P_T^n}\right].$$

It can be checked via a qq plot that the distribution of η_n stays roughly normal under $\mathbb{F}$ but its variance increases. Further, the variance of η_n under $\mathbb{F}$ increases with n.

As we know the distribution of η_n under the forward measure $\mathbb{F}$, we can express η_n as a function of a standard Normal random variable as we have previously seen in equation (4.12). We therefore have that:

$$y_T^n(Y) := y_0^n \exp\left(\sigma_n \sqrt{T} h_n(Y) - \frac{1}{2}\sigma_n^2 T\right) = y_0^n \exp\left(\sigma_n \sqrt{T} F_{\eta_n}^{-1}(\Phi(Y)) - \frac{1}{2}\sigma_n^2 T\right).$$

Observe that is the same functional form as $\hat{y}_T^n$ used in the construction of the PVBP.

For the 1F fmcMFL model, the PVBP is given by (5.8a) with two differences; (i) $\hat{y}_T^j$ is given by (5.8b) with η_n replaced by $x_T \sim \mathcal{N}(0,1)$ under $\mathbb{F}$ and (ii) the constant a_n is consequently different to ensure the martingale property for P_T^n under $\mathbb{F}$ holds. Figure 5.2 below compares the functional form $\hat{y}_T^n(x)$ for the swaption version and forward version of the model. Note that we plot the log-transform of the functional forms. It shows that the function $h_n(\cdot)$ is not the identity function (If it was, the two functional forms would have coincided). Instead, we observe that the log-transform of the functional form for $\hat{y}_T^n$ from the benchmark model has a steeper slope than the one from the 1F fmcMFL model, indicating an increase in the variance of η_n as we move to the forward measure. This measure-change effect is not reflected in the construction of the PVBP. This results in the variance of the PVBP being too small when the model is set up under the forward measure. As noted above, in this log-normal scenario, in the 1F fmcMFL model, we started off by constructing the PVBP $\hat{P}_T^n$, by postulating a set of priors $\{\hat{y}_T^j : j \in \{1, ..., n\}\}$ using the same set of parameters as in the benchmark model. Figure 5.2 indicates that when forming the 1F fmcMFL model, one should adjust the parameters y_0^j and σ_j for $j = 1, ..., n$ to reflect the distributions of the swap rates in the forward measure when forming the PVBP before the calibration step is done.

The wild oscillations of the MF-Lite model when set up under the forward measure that we observe in Figure 5.1 can be explained by focusing on the effect on the PVBP of the increased variance of η_n when moving to the forward measure.

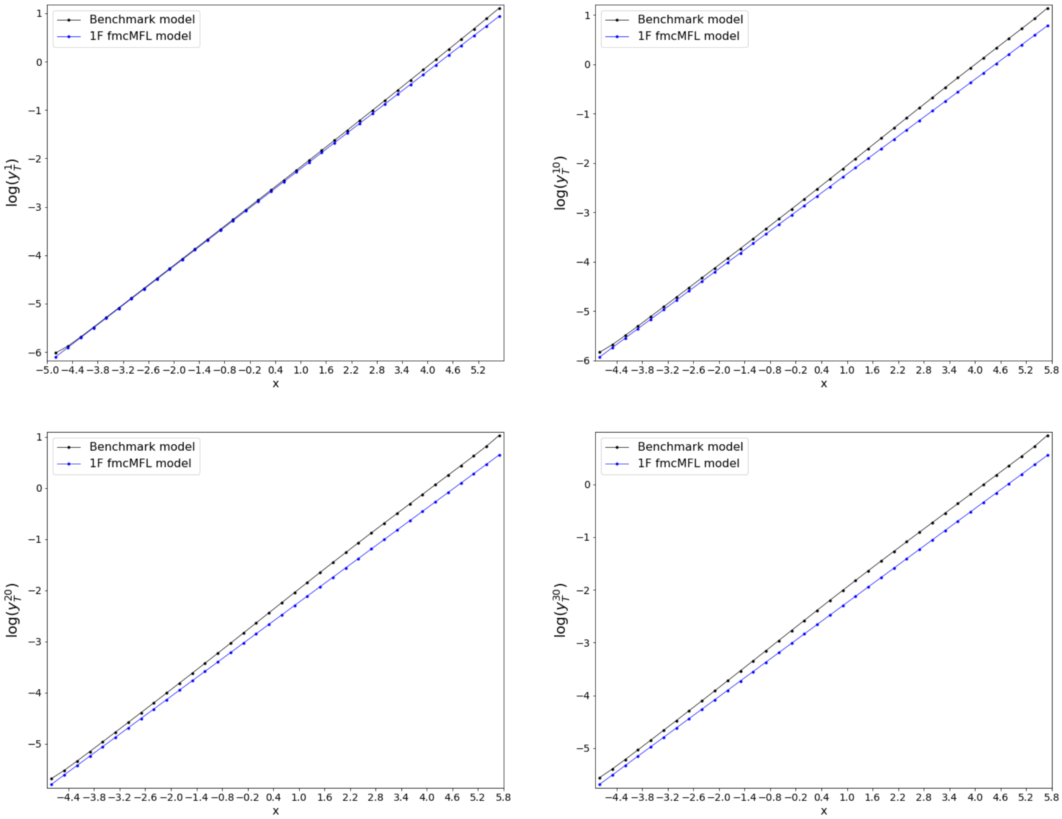

Figure 5.2: Functional form of $\log(y_T^n)$ *for* $n = \{1, 10, 20, 30\}$*: Benchmark model v 1F fmcMFL model*

Remark 27 (Explaining the zig-zag effect of the alternate fit model)*: We can do an explicit calculation to demonstrate the effect of lower variance by considering the extreme case when some of the PVBPs are deterministic (i.e variance zero).*

Example: consider the following unrealistic alternate fit model. Let m *and* n *be even. We will assume the PVBPs* P_T^n, P_T^m *and* P_T^{m-2} *are deterministic. We will now consider a payment of the reference swap rate* y_T^n *paid at either* S_m *or* S_{m-1}.

Suppose payment is made at S_m*. We have:*

$$y_T^n D_{TS_m} = y_T^n - y_T^n y_T^m P_T^m.$$

In the case where the PVBPs P_T^n *and* P_T^m *are deterministic, the swaption measures* $\mathbb{S}^n$, $\mathbb{S}^m$ *and the forward measure* $\mathbb{F}$ *coincide. In this case, under the assumption that the* *marginal distributions of the swap rates are log-normal*

$$\begin{aligned}
\mathbb{E}_{\mathbb{F}}[y_T^n D_{TS_m}] &= \mathbb{E}_{\mathbb{F}}[y_T^n] - \mathbb{E}_{\mathbb{F}}[y_T^n y_T^m P_T^m] \\
&= y_0^n - y_0^n y_0^m \exp(\sigma_n \sigma_m T) P_0^m \\
&= y_0^n - y_0^n y_0^m P_0^m - y_0^n y_0^m P_0^m (\exp(\sigma_n \sigma_m T) - 1) \\
&= y_0^n D_{0S_m} - y_0^n (1 - D_{0S_m})(\exp(\sigma_n \sigma_m T) - 1).
\end{aligned}$$

The convexity adjustment to the forward is then

$$\mathcal{C}_{n,m} = -y_0^n \frac{1 - D_{0S_m}}{D_{0S_m}} (\exp(\sigma_n \sigma_m T) - 1).$$

So for $m > 0$, m even, this gives a negative convexity correction.

Now let the payment date be S_{m-1} and take the accrual factor to be 1. We have:

$$\begin{aligned}
D_{TS_{m-1}} &= P_T^{m-1} - P_T^{m-2} \\
&= P_T^m - D_{TS_m} - P_T^{m-2} \\
&= (P_T^m - P_T^{m-2}) - (1 - y_T^m P_T^m) \\
y_T^n D_{TS_{m-1}} &= y_T^n (P_T^m - P_T^{m-2}) - y_T^n (1 - y_T^m P_T^m). \\
\mathbb{E}_{\mathbb{F}}[y_T^n D_{TS_{m-1}}] &= \mathbb{E}_{\mathbb{F}}[y_T^n (P_T^m - P_T^{m-2})] - \mathbb{E}_{\mathbb{F}}[y_T^n (1 - y_T^m P_T^m)] \\
&= y_0^n (P_0^m - P_0^{m-2}) - (y_0^n - y_0^n y_0^m \exp(\sigma_n \sigma_m T) P_0^m) \\
&= y_0^n (P_0^m - P_0^{m-2}) - (y_0^n - y_0^n y_0^m P_0^m) + y_0^n y_0^m P_0^m \exp(\sigma_n \sigma_m T) - 1) \\
&= y_0^n D_{0S_{m-1}} + y_0^n (1 - D_{0S_m})(\exp(\sigma_n \sigma_m T) - 1).
\end{aligned}$$

So for an odd payment date, this gives a positive convexity correction given by:

$$\mathcal{C}_{n,m-1} = y_0^n \frac{(1 - D_{0S_m})}{D_{0S_{m-1}}} (\exp(\sigma_n \sigma_m T) - 1).$$

Observe that the total value of the convexity correction at $m - 1$ and m (appropriately discounted) add up to zero.

The reason for this is that the convexity correction for the payoff $y_T^n P_T^{m-2}$ and $y_T^n P_T^m$ for even n and m will both be zero as in this case the swaptions measures $\mathbb{S}^n$, $\mathbb{S}^m$, $\mathbb{S}^{m-2}$ and the forward measure $\mathbb{F}$ all agree. This demonstrates the source of the oscillatory behaviour.

5.2 A refined two-factor forward measure calibrated MF-Lite model

As done before, we set the 2F *refined* fmcMFL model up under the forward measure for a given maturity denoted by S_n. We begin by forming the PVBP P_T^n. Let $(x_T^{(1)}, x_T^{(2)})$ be a bivariate random variable summarising the state of the economy at time T. We make the same assumption on the model driver as in equation (5.1) and we construct a set of variables $(z_T^i : i \in \{1, ..., n\})$ as in equation (5.2). We have learnt from the analysis of the one-factor models in the last section that when forming the priors of the swap rates, the assumption that z_T^i are normal in the forward measure leads to problems. We rectify this by refining the choice of priors used to set up the PVBP.

In the 2F fmcMFL model we have discussed in the previous section, we recall that under the forward measure, we postulated priors for the swap rates as defined in equation (5.3). To correct for the mismatch highlighted in the last numerical analysis, we propose to transform the variable z_T^i, $i \in \{1, ..., n\}$, in an informed way, and express the prior swap rates in terms of the transformed variable.

We introduce some deterministic function $h_n : \mathbb{R} \to \mathbb{R}$, the derivation of which is explained in Section 5.2.1. Note that we emphasise the subscript on the function, pointing out that we construct it using information known only under the swaption measure $\mathbb{S}^n$, and we transform the set of variables $\{z_T^i : i \in \{1, ..., n\}\}$ using the single function. We then postulate a set of prior swap rates as follows:

$$\hat{y}_T^{n:i}(z_T^i) := \hat{f}_i(h_n(z_T^i)), \tag{5.9}$$

where $\hat{f}_i : \mathbb{R} \to \mathbb{R}$ is a known deterministic function, chosen in a way to reflect our prior belief on the marginal distribution of the swap rates.

Remark 28: Recall in the 1F smcMFL model, developed in Chapter 4, when setting the model up at S_n, the priors for setting up the PVBP were $\hat{y}_T^i = \hat{f}_i(\eta_n)$. The intuition here is to aim to find a function h_n so that the random variable $h_n(z_T^i)$ has the same distribution as η_n under the forward measure. Note that z_T^i is $\mathcal{N}(0,1)$ under $\mathbb{F}$ for all i, so the function h_n does not depend on i. Then each of the priors $\hat{y}_T^{n:i}$ will have the same distribution under $\mathbb{F}$ as those in the swaption version of the model.

In the swaption version of the model, we suppressed the dependence on the index n of the priors as the dependence was clear through the dependence on η_n.

We can now postulate a prior form $\hat{P}_T^n(x_T^{(1)}, x_T^{(2)})$ as we did before in equation (5.4), but using the refined priors defined in equation (5.9).

We have now shown how to model the PVBP for a single maturity. Suppose we do the partial model setup for another maturity which we denote by S_m. If we follow similar steps as above and define the prior forms as:

$$\hat{y}_T^{m:i}(z_T^i) := \hat{f}_i(h_m(z_T^i)).$$

We are able to construct a prior form for the PVBP $\hat{P}_T^m$. We note at this stage that we have chosen different priors to construct the PVBP terms. It may seem natural to make a unified choice for the priors under the forward measure. For instance, WLOG, assume $m < n$; we could set:

$$h_i = \begin{cases} h_m, & \text{for } i \leq m \\ \frac{n-i}{n-m} h_m + \frac{i-m}{n-m} h_n, & \text{for } m < i < n \end{cases}$$

and use the above to form a single set of priors. However, we only have information about the distribution of the swap rates in their respective swaption measures and we use this to inform our choices for h_n and h_m. We are aiming for the correct behaviour on average for the PVBPs and it is not clear a hybrid choice would improve the model setup.

Using the fact that P^n is a martingale under $\mathbb{F}$, we adjust the postulated functional form for the PVBP by defining $P_T^n := a_n \hat{P}_T^n$ and we set a_n as given in equation (5.5) with the expectation evaluated using the refined functional form for $\hat{P}_T^n$.

We have seen above how we model the PVBP P_T^n under the forward measure. It still remains to decide how to construct the functional form for the swap rate y_T^n so that the model is calibrated to its known marginal distribution under its associated swaption measure. To do so, we perform a *Markov-functional sweep* to fit the model to swaption prices. Expressing y_T^n as some monotonic increasing function, f^n of z_T^n, we can recover f^n following the same calibration steps as we did for the 2F fmcMFL model. We can follow the same steps at maturity S_m and this completes the model.

5.2.1 Forward swap rate priors construction

As touched upon above, we lean on market information available to us under the swaption measure in order to construct the function h_n. We recall the role of the function h_n is to modify the set of variables z_T^i, for $i \in \{1, \ldots, n\}$ so that the PVBP takes into account the measure-change effect (from swaption measure to forward measure) on the distribution of the driver. We outline below how we construct such a function. The procedure for h_m is similar. We note that in order to find the function h_n, we set up a prior model starting in the swaption measure as detailed below. The purpose of this prior model is purely to find the function h_n and plays no further role in the final model.

Under the swaption measure $\mathbb{S}^n$, let

$$\begin{pmatrix} \tilde{x}_T^{(1)} \\ \tilde{x}_T^{(2)} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \right).$$

Define

$$\tilde{z}_T^i := \frac{\beta_i^{(1)} \tilde{x}_T^{(1)} + \beta_i^{(2)} \tilde{x}_T^{(2)}}{\sqrt{(\beta_i^{(1)})^2 \lambda_1 + (\beta_i^{(2)})^2 \lambda_2}}.$$

we define the prior forms:

$$\tilde{y}_T^{n:i}(\tilde{z}_T^i) := \hat{f}_i(q_n(\tilde{z}_T^i)),$$

where q_n is chosen such that the distribution of $\tilde{y}_T^{n:n}$ matches the implied market distribution under $\mathbb{S}^n$ given by swaption prices on the n^{th} swap rate y_T^n.

We can form a prior $\tilde{P}_T^n$ using $\tilde{y}_T^{n:i}$, for $i \in \{1, ..., n\}$ as we have done before and we find a constant term $\tilde{a}_n$ such that

$$\mathbb{E}_{\mathbb{S}^n}\left[\frac{1}{\tilde{a}_n \tilde{P}_T^n}\right] = \frac{D_{0T}}{P_0^n}.$$

We can now move on to the construction of the function h_n. To do so, we focus on the 1-dimensional variable $\tilde{z}_T^n$. Our aim is to find h_n with:

$$\eta_n := q_n(\tilde{z}_T^n) = h_n(Y),$$

where $Y \sim \mathcal{N}(0, 1)$ under $\mathbb{F}$.

Observe that η_n is a function of $\tilde{x}_T^{(1)}$ and $\tilde{x}_T^{(2)}$. We can find the distribution of η_n under the forward measure $\mathbb{F}$:

$$\begin{aligned} F_{\eta_n}(x) &:= \mathbb{E}_{\mathbb{F}}\Big[\mathbb{1}(\eta_n \le x)\Big] \\ &= \frac{P_0^n}{D_{0T}} \mathbb{E}_{\mathbb{S}^n}\left[\mathbb{1}(q_n(\tilde{z}_T^n) \le x) \frac{1}{\tilde{a}_n \tilde{P}_T^n(\tilde{x}_T^{(1)}, \tilde{x}_T^{(2)})}\right]. \end{aligned}$$

The function h_n can be recovered once we know the distribution of η_n under $\mathbb{F}$. This is the function we use in constructing the prior forms for the final MF-Lite model in the forward measure.

Remark 29: We make the observation that if we take $\beta_i^{(2)} = 0$, hence going back to a one-factor model, the PVBP would be a function of a single variable $x_T^{(1)}$, and the postulated *prior model for the PVBP would be the same as that obtained if we were to set the model up in the swaption measure and move to the forward measure by performing a measure change. We recall in remark 25 we made the observation that the variables with respect to which we formulate the functional form of the PVBP have different distributions under $\mathbb{F}$. The refinement proposed above rectifies this mismatch in distribution.*

5.2.2 Numerical results

We use the same tenor structure as we have described in Section 3.2.1. We assume that the swap rates are log-normally distributed under their respective swaption measures with log-normal volatilities as given in Table 1 in Appendix A. We want to be able to compare the 2F refined fmcMFL model to the 2F single-time MFM. We therefore set the model parameters up in line with the choices we have made in Section 3.3.2 of Chapter 3, and the model intitial conditions as given in Appendix A.

We recall that based on the prior model setup in Section 3.3.1, under log-normal assumption, we derived an approximate formula for the correlation between the log of the swap rates, which we restate here:

$$\begin{aligned}\operatorname{corr}(\log(y_T^j), \log(y_T^m)) &\approx \frac{\mathbb{E}_{\mathbb{F}}[(\beta_j^{(1)} x_T^{(1)} + \beta_j^{(2)} x_T^{(2)})(\beta_m^{(1)} x_T^{(1)} + \beta_m^{(2)} x_T^{(2)})]}{\sqrt{\operatorname{var}(\log y_T^j)\operatorname{var}(\log y_T^m)}} \\ &= \frac{\beta_j^{(1)}\beta_m^{(1)}\lambda_1 + \beta_j^{(2)}\beta_m^{(2)}\lambda_2}{\sqrt{(\beta_j^{(1)})^2\lambda_1 + (\beta_j^{(2)})^2\lambda_2}\sqrt{(\beta_m^{(1)})^2\lambda_1 + (\beta_m^{(2)})^2\lambda_2}},\end{aligned}$$

for $j, m \in \{1, \ldots, \tilde{M}\}$.

For the numerical analysis below, we make the following choices: we set $\lambda_1 = 1$ and we vary the correlation structure via λ_2; we note that there is an upper-bound on λ_2 as explained in Section 3.3.2. We recall that we choose to set the $\beta^{(i)}$ parameters, for $i \in \{1, 2\}$ by linking them to the PCA decomposition of the covariance matrix of the log of the LIBORs, as we have seen in equation (3.18). This requires us to specify the eigenvectors $(a_{ki})_{k=1}^{\tilde{M}}$, $i \in \{1, 2\}$, and for this numerical analysis, we use the same choice as in equation (3.21). We recall that the parameter λ controls the shape of the second eigenvector. We take $\lambda = 0.1, 0.3$. The correlation values used in the numerical analysis are given in Appendix D.

We now look at the convexity correction as a function of payment date and compare the results from the 2F refined fmcMFL model to that of the 2F single-time MFM.

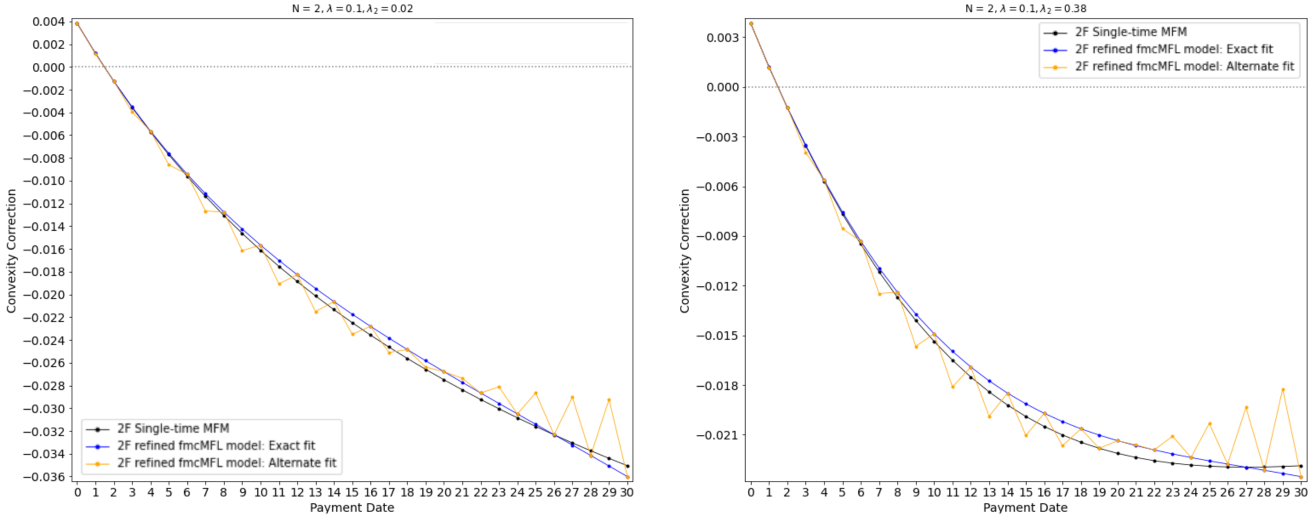

Figure 5.3: $N = 2, \tilde{M} = 30, D_{0T} = 1.0, y_0 = 0.07, \sigma_2 \approx 0.15$;
LHS: $\lambda = 0.1, \lambda_2 = 0.02$
RHS: $\lambda = 0.1, \lambda_2 = 0.38$
Convexity correction against payment date

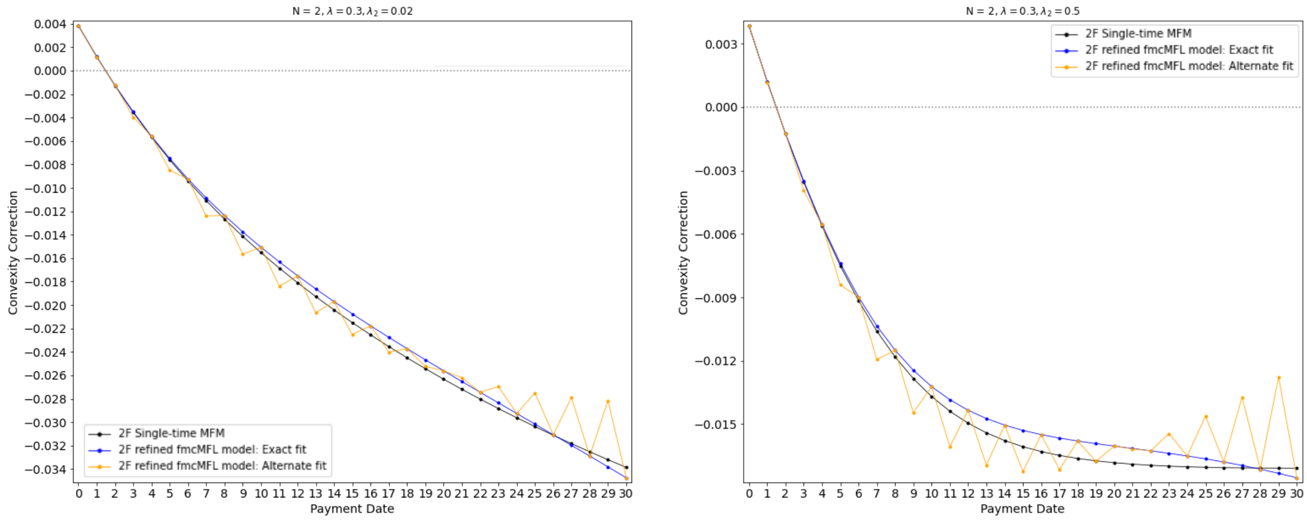

Figure 5.4: $N = 2, \tilde{M} = 30, D_{0T} = 1.0, y_0 = 0.07, \sigma_2 \approx 0.15$;
LHS: $\lambda = 0.3, \lambda_2 = 0.02$
RHS: $\lambda = 0.3, \lambda_2 = 0.5$
Convexity correction against payment date

We repeat the same analysis as above with N = 10.

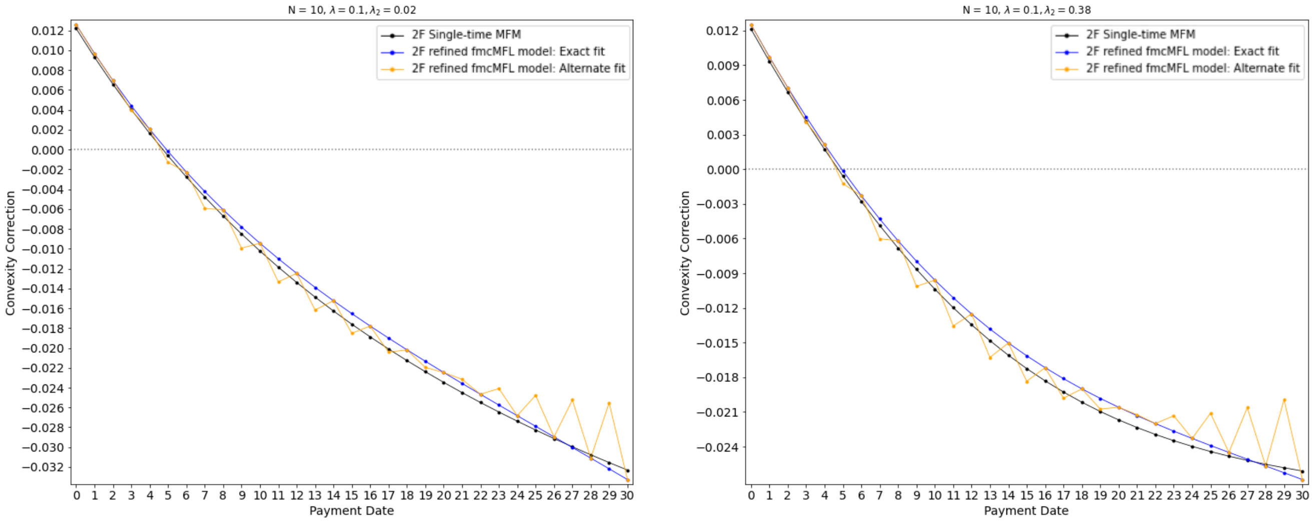

Figure 5.5: $N = 10, \tilde{M} = 30, D_{0T} = 1.0, y_0 = 0.07, \sigma_2 \approx 0.15$;
LHS: $\lambda = 0.1, \lambda_2 = 0.02$
RHS: $\lambda = 0.1, \lambda_2 = 0.38$
Convexity correction against payment date

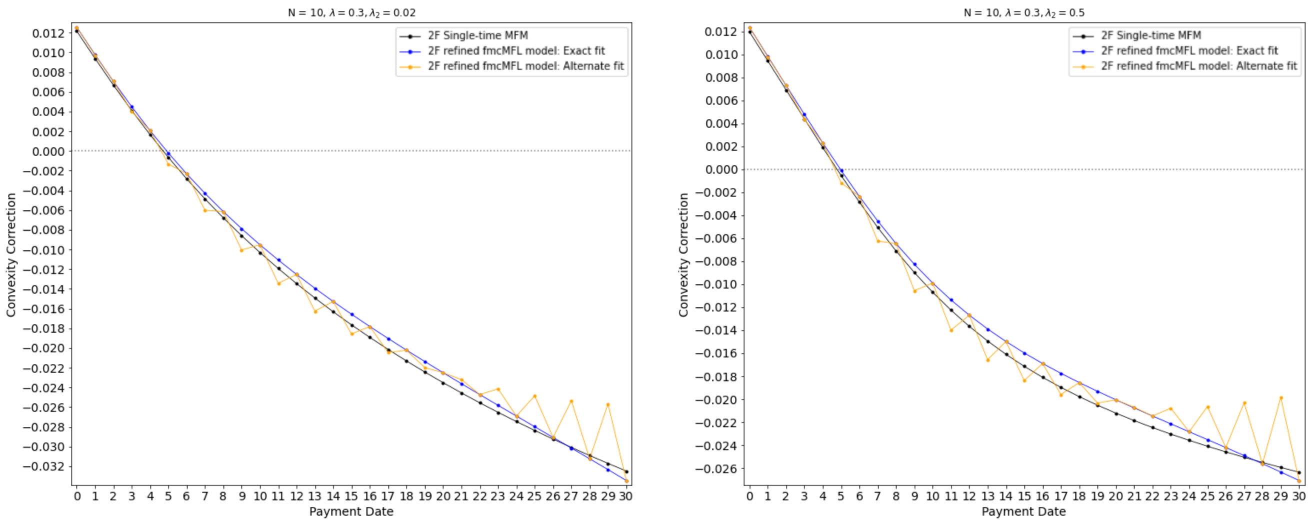

Figure 5.6: $N = 10, \tilde{M} = 30, D_{0T} = 1.0, y_0 = 0.07, \sigma_2 \approx 0.15$;
LHS: $\lambda = 0.3, \lambda_2 = 0.02$
RHS: $\lambda = 0.3, \lambda_2 = 0.5$
Convexity correction against payment date

We observe from the figures above that the results from the exact fit model setup is close to the results from the 2F single-time MFM. The variation/oscillations in convexity corrections from the alternate fit setup is much lower in the 2F refined fmcMFL model compared to the 2F fmcMFL model developed in Section 5.1, thanks to the careful modelling of the PVBP. We note however that as the payment date increases beyond the reference swap rate end date, we do not need the consistency condition to hold, so we can use the exact fit setup, which in the refined model, is relatively close to the results from the single-time MFM, as desired. Comparing the results for fixed N and λ, but varying λ_2 (i.e we decrease the correlation), we observe that the convexity correction is close across the models for payment dates within the reference index, but as the payment date increases beyond, we observe that the correlation does have an effect.

We have proposed a two-factor Markov-functional approach to pricing constant maturity swaps. The model is calibrated to the known marginal distributions of the reference swap rate and the payment swap rate while taking into account the correlation between the two rates. The results obtained from the model is comparable to the 2F single-time MFM approach. In setting up the model, careful choices have been made when setting up the prior forms so that the two-factor MF-Lite approach mirrors closely the model structure of the single-time MFM. In the next section, we discuss a much simpler two-factor setup in the swaption measure. The model still performs well when the correlation is high, but as the correlation decreases, the model fails to capture its effect on convexity corrections. We would argue that the setup is still worth discussing due to its simplicity and efficiency. We name this model the 2F naive smcMFL model.

5.3 A two-factor naive MF-Lite approach in the swaption measure

In contrast to the previous 2F MF-Lite models introduced, we explore in this section a simple two-factor approach to pricing a CMS, whereby we start off by making model choices under the swaption measure. It would seem natural to do so since we have market information about the swap rates under the swaption measure. However, setting up the model as such, whilst keeping the fundamental model structure of the 2F single-time MFM is challenging. We discuss this point below.

At a given maturity, say S_n, starting under the swaption measure $\mathbb{S}^n$, assume we have a 2-dimensional driver $\mathbf{x}_T := (x_T^{(1)}, x_T^{(2)})$. We construct a set of variables $\mathbf{z} := (z_T^1, z_T^2, ..., z_T^n)$ by taking linear combinations of the two components of the driver using the same weights as in the single-time MFM. We assume we can express y_T^n as a monotonic increasing function of z_T^n - we are only considering a particular linear combination - the functional form will be chosen such that the model is calibrated to the known market-implied marginal distribution of y_T^n. Notice the contrast with the 1F smcMFL model. In the latter, we proposed a functional form for the swap rate in terms of the one-dimensional driver η_n. We then chose the distribution of η_n such that the model is calibrated to the appropriate market-implied distribution for the swap rate. Here, since we have a 2-dimensional driver, we cannot determine the joint distribution of $\mathbf{x}_T$ solely by calibrating the model. Some choice has to be made for the joint distribution of $x_T^{(1)}$ and $x_T^{(2)}$; we note that this will inform the marginal distribution of z_T^i, for $i = 1, ..., n$. Hence, the only flexibility left is to choose a suitable functional form for y_T^n. We then postulate a functional form for P_T^n as some function of $\mathbf{z}$ using the standard functional relationship as we have previously seen in equation (5.4) (Note that this would require us to formulate priors for the earlier swap rates). Having made these choices under $\mathbb{S}^n$, we can now carry the model over to the forward measure $\mathbb{F}$ in a similar fashion as we did for the one-factor case - by a change of measure to $\mathbb{F}$, we can compute the joint distribution of $x_T^{(1)}$ and $x_T^{(2)}$ under the forward measure. We stress the importance of appropriately modelling the PVBP here, for this variable is crucial in carrying the model over from $\mathbb{S}^n$ to the forward measure. Now suppose we were to follow the same procedure for the other maturity of interest, denoted by S_m. We would have to start with a two-dimensional driver $\hat{\mathbf{x}}_T := (\hat{x}_T^{(1)}, \hat{x}_T^{(2)})$ under $\mathbb{S}^m$, and carry the partial model setup over to the forward measure. It still remains to decide how to unify the two separate parts of the model, more precisely, the two 2-dimensional drivers $\mathbf{x}_T$ and $\hat{\mathbf{x}}_T$ under the forward measure. The inversion principle will not work here. The model setup, as described ends up becoming too complicated to be viable in practice.

The approach proposed here, which we shall refer to as the *2F naive smcMFL model* is a simplified, yet practical version and an extension to the 1F smcMFL model. We discuss below the model construction.

5.3.1 Model setup for a single maturity

The 2F naive smcMFL model setup for a single maturity follows vertabim the construction we have seen in Section 4.1 of Chapter 4. At a given maturity S_n, starting off under the swaption measure $\mathbb{S}^n$, we assume there exists a random variable, denoted by η_n, and we can postulate a functional form, which is assumed to be monotonic increasing, for the swap rate y_T^n in terms of η_n:

$$y_T^n(\eta_n) := f_n(\eta_n),$$

for $f_n : \mathbb{R} \to \mathbb{R}$. We further assume that we know the distribution of y_T^n under $\mathbb{S}^n$, (which can be derived from swaption prices). We can choose the distribution of η_n under $\mathbb{S}^n$ such that the model is calibrated to the known marginal distribution of y_T^n. We next postulate a functional form for the PVBP P_T^n as follows:

$$P_T^n(\eta_n) := a_n \hat{P}_T^n(\eta_n),$$

where $a_n \in \mathbb{R}$ is chosen such that the no-arbitrage equation $D_{0T} = P_0^n \mathbb{E}_{\mathbb{S}^n}[1/P_T^n]$ holds. Moving over to the forward measure, choosing the functional form for the PVBP P_T^n

determines the distribution of η_n under $\mathbb{F}$, which we denote by F_{η_n}. The knowledge of the distribution function of η_n allows us to construct a functional form for η_n in terms of a standard Normal random variable under $\mathbb{F}$. More precisely, as we have seen in Chapter 4, if we define $Y := \Phi^{-1}\big(F_{\eta_n}(\eta_n)\big)$, then $Y \sim \mathcal{N}(0,1)$ under $\mathbb{F}$. Rearranging, we get

$$\eta_n(Y) := (F_{\eta_n})^{-1}(\Phi(Y)).$$

We recall that from the two model choices (the functional forms of y_T^n and P_T^n in terms of η_n), we can recover the functional forms of D_{TS_n} and P_T^{n-1} since:

$$\begin{aligned} D_{TS_n} &= 1 - y_T^n P_T^n, \\ P_T^{n-1} &= P_T^n - \alpha_n D_{TS_n}. \end{aligned}$$

We can therefore specify a new set of functional forms from the postulated ones as follows:

$$\begin{aligned} y_T^n(\eta_n(Y)) &=: h_n^y(Y) \\ P_T^n(\eta_n(Y)) &=: h_n^P(Y) \\ D_{TS_n}(\eta_n(Y)) &=: h_n^D(Y) \end{aligned} \tag{5.10}$$

If we were to set up the model to price a single payment of a CMS based on the forward swap rate y_T^n and payment made at time S_m, the valuation of which is given by $D_{0T}\mathbb{E}_{\mathbb{F}}[y_T^n D_{TS_m}]$, we would have to set up the partial models for the two different maturities S_n and S_m. The question therefore arises as to how do we *knit* the two parts together? Unlike the one-factor case, there is no straightforward method to do so (Indeed, in the one-factor setup, when we do the partial models at two maturities, we end up with two model drivers and there exists a unique functional relationship between them under $\mathbb{F}$. We lose that uniqueness in the two-factor case). We discuss in the next section a modelling choice we can make to bring the two partial models together.

5.3.2 The exact fit approach revisited

We recall in Section 4.2 of Chapter 4, we proposed an alternate fit approach as a consistent model that takes into account the functional relationships between the variables we are modelling. We lost some flexibility in calibrating the model, but the numerical analysis showed that we still end up with a model close to the 2F single-time MFM. In this naive approach, coming up with the functional forms is done under a one-factor assumption. We could potentially adapt the alternate fit model in this case; in order to set up a consistent model, at any given maturity S_m that lies within an odd time step from the reference swap rate maturity S_n, i.e $|m-n|$ is odd, we would have to do the single-maturity setup at three different maturities (S_n, S_{m-1} and S_{m+1}). The pure discount bond at maturity S_m, and consequently, the swap rate functional forms are expressed as functions of two variables η_{m-1} and η_{m+1}, in order to satisfy the functional relationships that exist between the variables as we have seen in Section 4.2. Note that, at this point, we have lost the ability to calibrate the model to a desired correlation between the swap rates y_T^n and y_T^m. We would achieve a consistent model, but one that is not well-behaved at alternate maturities, so we do not include an analysis here.

We focus rather on the exact fit approach. We describe the model construction in two steps:

Step 1 For each maturity S_i in the tenor structure, from the setup described in Section 5.3.1, we can get hold of the functional forms $y_T^i(\eta_i)$ and $P_T^i(\eta_i)$, with the distribution of η_i known under $\mathbb{F}$. We have equally shown in equation (5.10) that we can construct a new set of functional forms for the variables in terms of a standard Gaussian random variable under $\mathbb{F}$.

Step 2 Now consider two maturities S_n and S_m, where n corresponds to the reference index and m corresponds to the payment index. The second step consists of specifying the joint distribution of y_T^n and D_{TS_m}. Note we may have $n = m$. Let Y_1 and Y_2 have a joint Gaussian distribution, unit variance, mean zero, correlation denoted by $\hat{\rho}$. Take $y_T^n = h_n^y(Y_1)$ and $D_{TS_m} = h_m^D(Y_2)$.

Remark 30*: Observe that from the way the model is set up, when $n = m$, we can choose the correlation $\hat{\rho}$ as desired; we are not forced to assign $\hat{\rho} = 1$.*

Secondly, we made a specific assumption for the joint distribution of (Y_1, Y_2). Note that it is a specialised case of using a Gaussian copula. More generally, we could have specified the joint dependence between Y_1 and Y_2 using any copula function.

5.3.3 Numerical results: Log-normal market-implied distributions

The term structure setup, initial conditions and model inputs used are the same as described in Section 5.2.2. We assume that the swap rates are log-normally distributed in their own swaption measures.

We first consider convexity correction as a function of payment date, taking $N = 2$ and we allow the payment date S_M to vary between the setting date up to a maximum payment date $S_{\tilde{M}}$. We discussed previously in Section 5.2.2 the choices made to get hold of the correlation between the log of the reference swap rate and the payment swap rate with respect to the forward measure $\mathbb{F}$ using the single-time MFM under log-normal assumptions. In order to be able to compare the 2F smcMFL model to the single-time MFM, we will use the same correlation between the log of the swap rates as input in the model. We assumed that, for a fixed reference index N and payment index M, we can express y_T^N as a function of Y_1 and y_T^M as a function of Y_2, where Y_1 and Y_2 are jointly Gaussian with correlation $\hat{\rho}$. For the given log-normal distribution assumption on the swap rates, for a given index N and payment index M, $\hat{\rho}$ can be chosen to match the correlation between $\log(y_T^N)$ and $\log(y_T^M)$ obtained from the single-time MFM. (since the log of the swap rates are linear functions of Y_1 and Y_2). See Appendix D for the correlation parameters between the log of the swap rates used in specifying the model. Note that in the plots below, we specify the correlation between $\log(y_T^N)$ and $\log(y_T^{30})$, which we denote by $\rho_{N,30}$. This indicates a lower bound on the correlation parameters used to generate the results (all the other correlations will be higher than $\rho_{N,30}$).

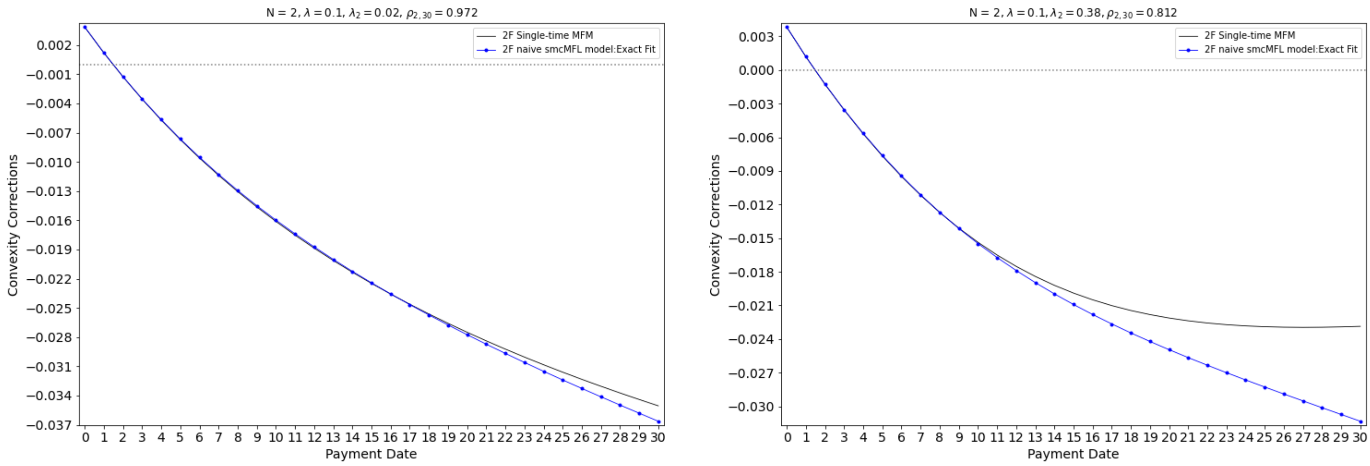

Figure 5.7: $N = 2, T = 20, D_{0T} = 1.0, y_0 = 0.07$;
Convexity correction against payment date with correlation between the log of the swap rates as given in Appendix D

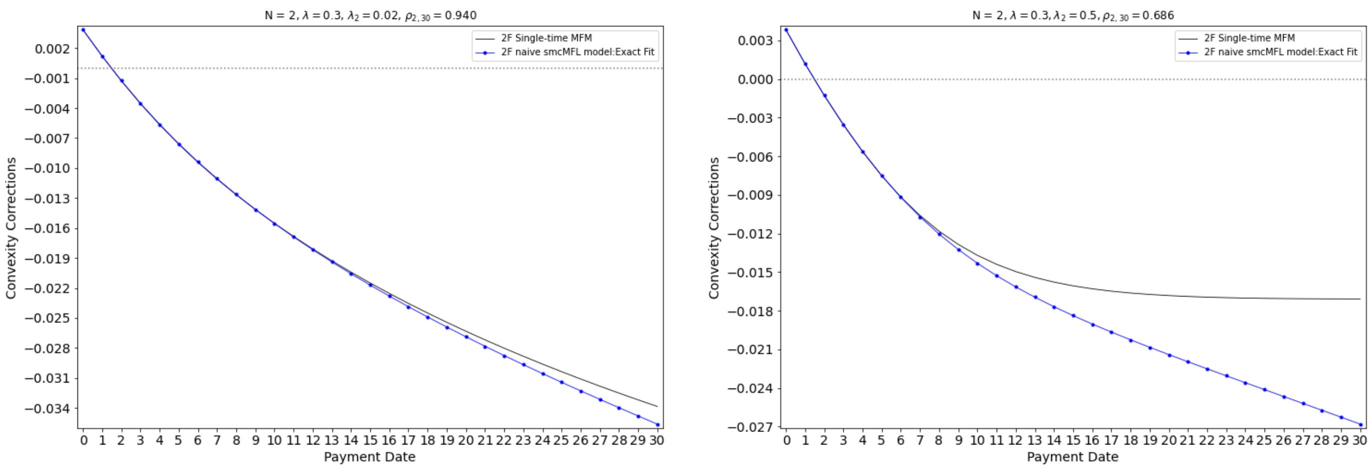

Figure 5.8: $N = 2, \tilde{M} = 30, D_{0T} = 1.0, y_0 = 0.07$;
Convexity correction against payment date with correlation between the log of the swap rates as given in Appendix D

We repeat the same analysis as above with N = 10.

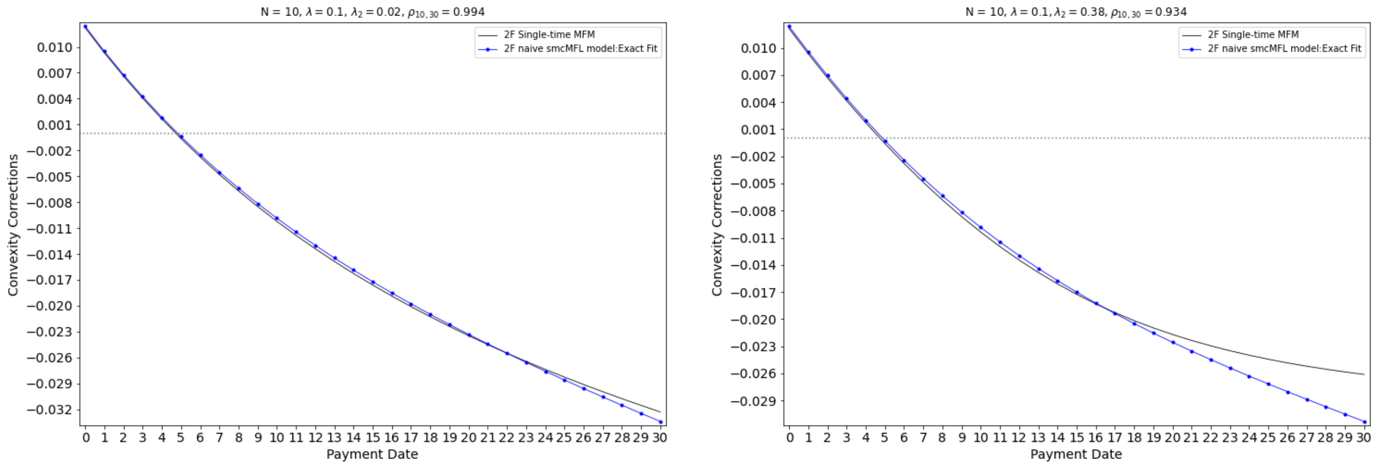

Figure 5.9: $N = 10, \tilde{M} = 30, D_{0T} = 1.0, y_0 = 0.07$;
Convexity correction against payment date with correlation between the log of the swap rates as given in Appendix D

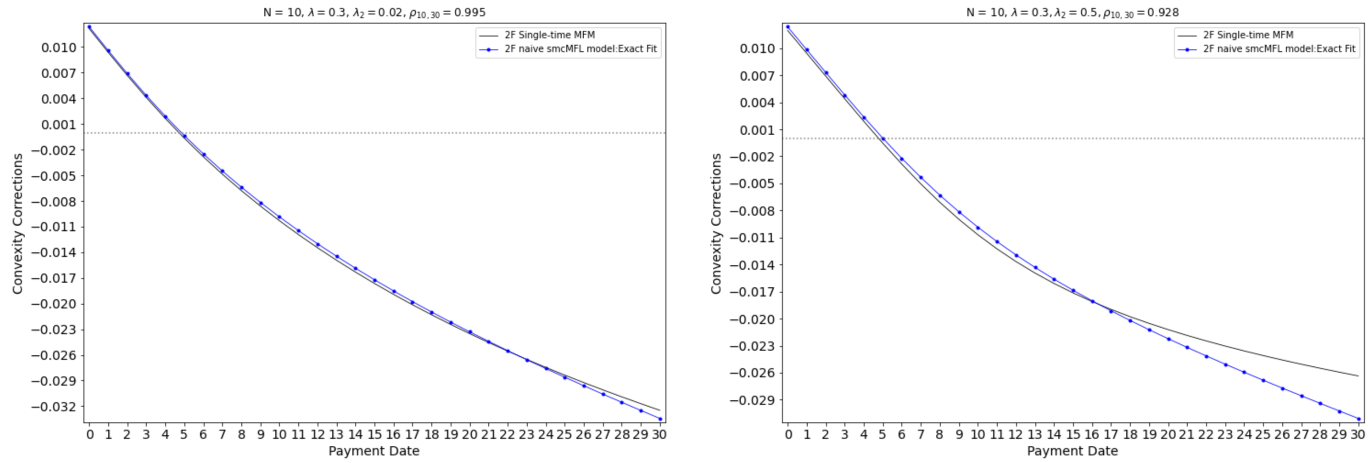

Figure 5.10: $N = 10, \tilde{M} = 30, D_{0T} = 1.0, y_0 = 0.07$;
Convexity correction against payment date with correlation between the log of the swap rates as given in Appendix D

We observe from the above figures that when the correlation is high (high enough that the model is almost indistinguishable from the one-factor setup), the two models are close to each other, as we would expect. But we can see that (de)correlation is certainly having an impact. The functional forms specified in this 2F naive smcMFL approach are suboptimal and this is why calibration to the correlation of the log of the reference and payment swap rates does not work well. Getting the correlation between the log of the reference swap rate and the log of the payment PDB correct in the calibration step largely overcomes the deficiencies in the choice of functional forms and gives a remarkably good match to the convexity corrections of the single time MFM, as we shall see in the next section. From a practical point of view this model could be considered for use in pricing a CMS.

5.3.4 Correlation between $\log(y_T^N)$ and $\log(D_{TS_M})$

As we have argued above, using the correlation between the log of the swap rates as input in the naive approach developed here is not enough to make up for the lack of structure in the model due to the simplistic choices we have made on the functional forms. Considering step 2 of the exact fit approach in Section 5.3.2, we can argue that if we take the correlation between the log of the swap rate and the log of the pure discount bond instead as input in the MF-Lite approach, we would get a closer match to the single-time MFM. In order to investigate this, we first go back to the single-time MFM setup to find an approximate formula for the correlation between $\log(y_T^N)$ and $\log(D_{TS_M})$.

Using the results from Section 2.2 of Chapter 2, for $j \in \{1, ..., \tilde{M}\}$, we set up the prior model for the swap rates in the log-normal context as follows:

$$\log(y_T^j) = \log(y_0^j) + \beta_j^{(1)} x_T^{(1)} + \beta_j^{(2)} x_T^{(1)} + C,$$

for some constant C.

We begin by finding an expression for the pure discount bond D_{TS_M}, for $M \in \{1, ..., \tilde{M}\}$ in terms of LIBORs.

$$D_{TS_M} = \frac{D_{TS_M}}{D_{TS_{M-1}}} . \frac{D_{TS_{M-1}}}{D_{TS_{M-2}}} . \dots . \frac{D_{TS_1}}{D_{TT}} . D_{TT} = \prod_{k=1}^{M} (1 + \alpha_k L_T^k)^{-1}.$$

Taking the log transformation of both sides, we have that:

$$\log(D_{TS_M}) = -\sum_{k=1}^{M} \log(1 + \alpha_k L_T^k).$$

Apply a first order Taylor Expansion to $\log(1 + \alpha_k L_T^k)$ about L_0^k to obtain:

$$\log(1 + \alpha_k L_T^k) \approx \log(1 + \alpha_k L_0^k) + \frac{\alpha_k}{1 + \alpha_k L_0^k}(L_T^k - L_0^k) + \dots \tag{5.11}$$

We can find an expression for $(L_T^k - L_0^k)$ in terms of the log of the LIBORS using equation (3.12) and we have that:

$$L_0^k(\log(L_T^k) - \log(L_0^k)) \approx L_T^k - L_0^k.$$

Plugging the above expression back into equation (5.11), and substituting the approximate expression for $\log(1+\alpha_k L_T^k)$ into the expression for $\log(D_{TS_M})$, we have:

$$\begin{aligned}\log(D_{TS_M}) &\approx -\sum_{k=1}^{M}\left(\log(1+\alpha_k L_0^k) + \frac{\alpha_k L_0^k}{1+\alpha_k L_0^k}(\log(L_T^k)-\log(L_0^k))\right)\\ &= \log(D_{0S_M}) - \sum_{k=1}^{M}\frac{\alpha_k L_0^k}{1+\alpha_k L_0^k}(\log(L_T^k)-\log(L_0^k)). \qquad (5.12)\end{aligned}$$

From the 2-factor separable LMM, we recall we model $\log(L_T^k)$ as follows:

$$\log(L_T^k) = \log(L_0^k) + a_{k1}x_T^{(1)} + a_{k2}x_T^{(2)} + \tilde{C},$$

where $\tilde{C}$ is some constant term - note that the MF sweep will remove the arbitrage introduced by the drift approximation. Going back to equation (5.12), we can express $\log(D_{TS_M})$ as:

$$\log(D_{TS_M}) \approx \log(D_{0S_M}) - \sum_{k=1}^{M}\frac{\alpha_k L_0^k}{1+\alpha_k L_0^k}(a_{k1}x_T^{(1)} + a_{k2}x_T^{(2)}).$$

Define

$$\zeta_M^{(1)} := -\sum_{k=1}^{M}\frac{\alpha_k L_0^k}{1+\alpha_k L_0^k}a_{k1} \qquad\qquad \zeta_M^{(2)} := -\sum_{k=1}^{M}\frac{\alpha_k L_0^k}{1+\alpha_k L_0^k}a_{k2}$$

We have:

$$\log(D_{TS_M}) \approx \log(D_{0S_M}) + \zeta_M^{(1)}x_T^{(1)} + \zeta_M^{(2)}x_T^{(2)}.$$

We can now find an approximate formula for the correlation between $\log(y_T^N)$ and $\log(D_{TS_M})$:

$$\operatorname{corr}(\log(y_T^N), \log(D_{TS_M})) \approx \frac{\beta_N^{(1)}\zeta_M^{(1)}\lambda_1 + \beta_N^{(2)}\zeta_M^{(2)}\lambda_2}{\sqrt{(\beta_N^{(1)})^2\lambda_1 + (\beta_N^{(2)})^2\lambda_2}\sqrt{(\zeta_M^{(1)})^2\lambda_1 + (\zeta_M^{(2)})^2\lambda_2}},$$

where we have assumed

$$\begin{pmatrix} x_T^{(1)} \\ x_T^{(2)} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\right),$$

under the forward measure $\mathbb{F}$ and λ_i corresponds to the eigenvalue obtained from the decomposition of the correlation matrix. We now look at convexity corrections against payment date. As opposed to before, in the 2F naive smcMFL model, we use the correlation between $\log(y_T^N)$ and $\log(D_{TS_M})$ obtained from the 2F single-time MFM, as tabulated in Appendix D.1, as input for the parameter $\hat{\rho}$ for a given reference index N

and payment index M. Similar to the previous section, in the figures below, we report the correlation between $\log(y_T^N)$ and $\log(D_{TS_{30}})$, giving us an indication of the lowest correlation parameter used in the model. (To avoid notational ambiguity, we still denote the correlation between $\log(y_T^N)$ and $\log(D_{TS_{30}})$ as $\rho_{N,30}$)

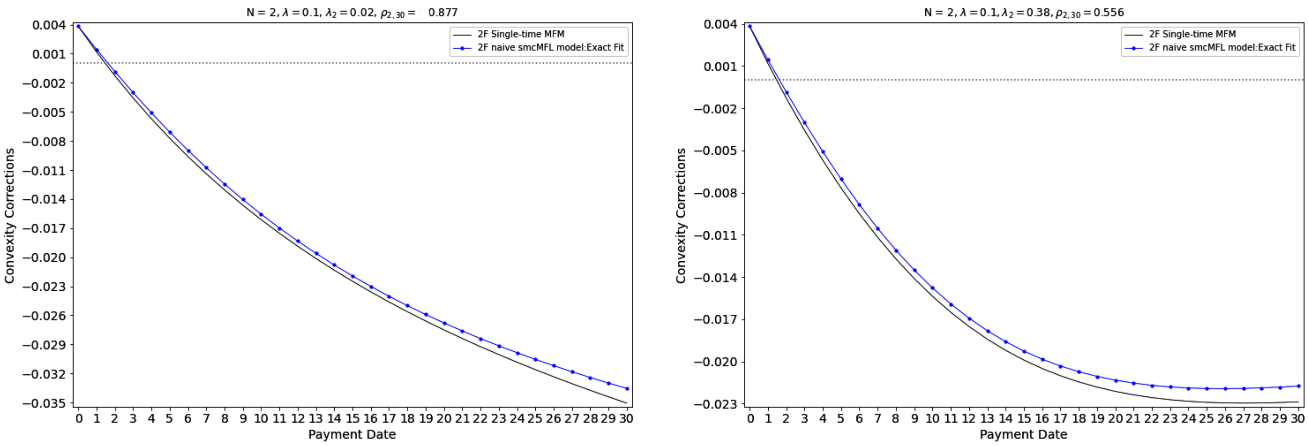

Figure 5.11: $N = 2, \tilde{M} = 30, D_{0T} = 1.0, y_0 = 0.07$
Convexity correction against payment date; using correlation between the $\log(y_T^N)$ *and* $\log(D_{TS_M})$ *as input in 2F naive smcMFL model*

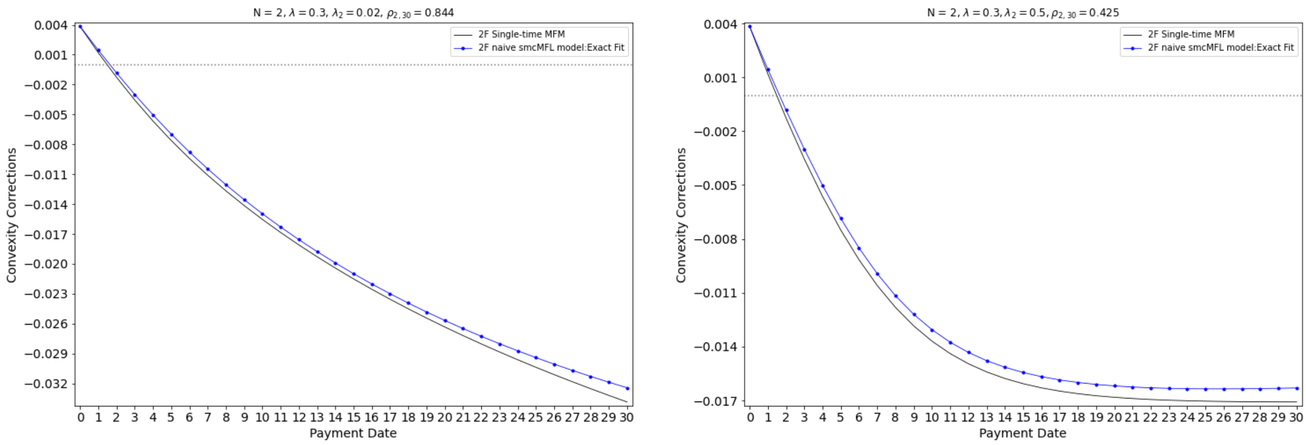

Figure 5.12: $N = 2, \tilde{M} = 30, D_{0T} = 1.0, y_0 = 0.07$
Convexity correction against payment date; using correlation between the $\log(y_T^N)$ *and* $\log(D_{TS_M})$ *as input in 2F naive smcMFL model*

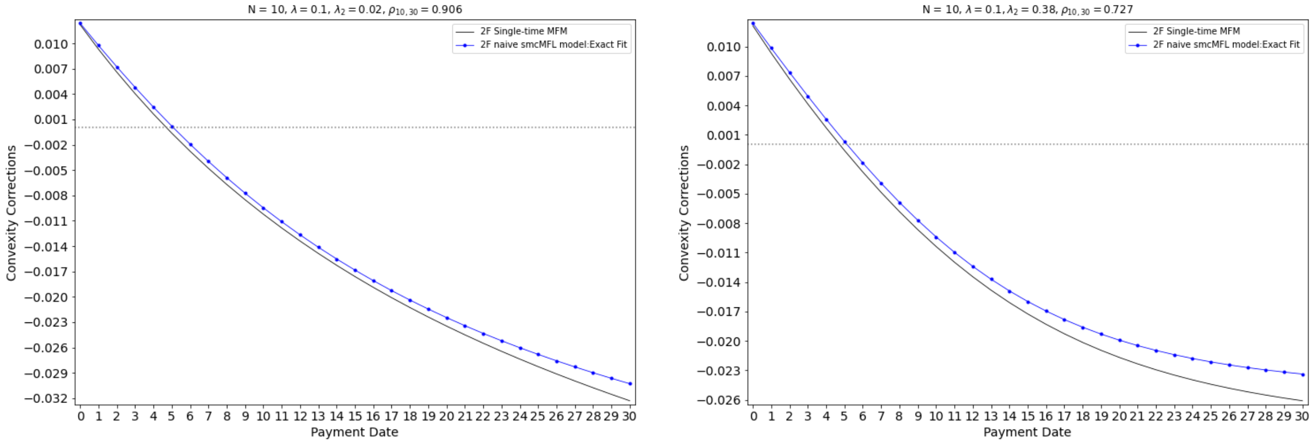

Figure 5.13: $N = 10, \tilde{M} = 30, D_{0T} = 1.0, y_0 = 0.07$
Convexity correction against payment date; using correlation between the $\log(y_T^N)$ *and* $\log(D_{TS_M})$ *as input in 2F naive smcMFL model*

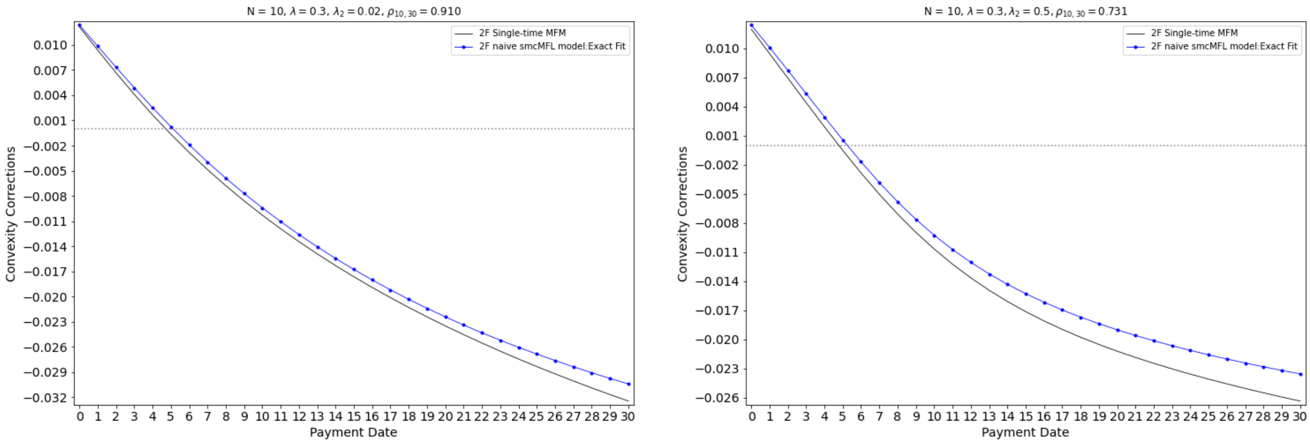

Figure 5.14: $N = 10, \tilde{M} = 30, D_{0T} = 1.0, y_0 = 0.07$
Convexity correction against payment date; using correlation between the $\log(y_T^N)$ *and* $\log(D_{TS_M})$ *as input in 2F naive smcMFL model*

There is an improvement in the convexity correction results, with the swaption 2F MF-lite model giving convexity corrections closer to the results from the 2F single-time MFM, hence achieving the goal of constructing a practical, computationally-fast model that can be used to price CMS. This 2F naive approach further highlights the importance of modelling the PVBP and the swap rate in an appropriate way, especially when the modelling is done under a measure with respect to which we have no direct market information available, hence justifying the methodical approach taken in the 2F fmcMFL approaches discussed in the previous sections.

Appendix

A Initial conditions and model input

j	y_0^j	σ_j	D_{0S_j}	P_0^j
1	0.07	0.15	0.934579439	0.934579439
2	0.07	0.148889	0.873438728	1.808018168
3	0.07	0.148035	0.816297877	2.624316044
4	0.07	0.147253	0.762895212	3.387211256
5	0.07	0.146505	0.712986179	4.100197436
6	0.07	0.14578	0.666342224	4.76653966
7	0.07	0.145073	0.622749742	5.389289402
8	0.07	0.144382	0.582009105	5.971298506
9	0.07	0.143705	0.543933743	6.515232249
10	0.07	0.143043	0.508349292	7.023581541
11	0.07	0.142394	0.475092796	7.498674337
12	0.07	0.141759	0.444011959	7.942686297
13	0.07	0.141138	0.414964448	8.357650744
14	0.07	0.140531	0.387817241	8.745467985
15	0.07	0.139938	0.36244602	9.107914005
16	0.07	0.139359	0.338734598	9.446648603
17	0.07	0.138794	0.31657439	9.763222993
18	0.07	0.138244	0.295863916	10.05908691
19	0.07	0.137709	0.276508333	10.33559524
20	0.07	0.137188	0.258419003	10.59401425
21	0.07	0.136681	0.241513087	10.83552733
22	0.07	0.136189	0.225713165	11.0612405
23	0.07	0.135712	0.210946883	11.27218738
24	0.07	0.135249	0.19714662	11.469334
25	0.07	0.1348	0.184249178	11.65358318
26	0.07	0.134366	0.172195493	11.82577867
27	0.07	0.133946	0.160930367	11.98670904
28	0.07	0.13354	0.150402212	12.13711125
29	0.07	0.133148	0.140562815	12.27767407
30	0.07	0.13277	0.131367117	12.40904118

Table 1: Initial Condition and log-normal volatility

B Parameterising the marginal distributions under Normal-Log Normal assumption

In order to choose the parameters for the mixed distribution, we compute the price of an ATM swaption under the lognormal assumption with chosen volatility, σ_i. We can find the closed form formula for the calibrating prices we feed into the model under the assumption that $y_T^i \sim \mathcal{N}(y_0^i, \eta_i^2)$. For each given i, we then choose an η_i value that matches the ATM swaption value under the Gaussian assumption to that of the lognormal case. Under the Gaussian assumption, it follows that:

$$
\begin{aligned}
(V_0^i(K))^N &:= P_0^i \mathbb{E}_{\mathbb{S}^i}[(y_T^i - K)_+] = P_0^i \Big(\mathbb{E}_{\mathbb{S}^i}[y_T^i \mathbb{1}(y_T^i > K)] - K \mathbb{E}_{\mathbb{S}^i}[\mathbb{1}(y_T^i > K)] \Big) \\
&= P_0^i \Big(\int_k^\infty \frac{1}{\sqrt{2\pi}\eta_i} y \exp\Big(-\frac{1}{2\eta_i^2}(y - y_0^i)^2 \Big) \mathrm{d}y - K\Big[1 - \Phi\Big(\frac{K - y_0^i}{\eta_i}\Big)\Big]\Big).
\end{aligned}
$$

Denoting by K^* the expression $\frac{K - y_0^i}{\eta_i}$, and applying a change of variables from y to $z := \frac{y - y_0^i}{\eta_i}$ in the integral term, we have that:

$$
\begin{aligned}
(V_0^i(K))^N &= P_0^i \Big(\int_{k^*}^\infty \frac{1}{\sqrt{2\pi}} (\eta_i z + y_0^i) \exp\Big(-\frac{1}{2} z^2 \Big) \mathrm{d}z - K(1 - \Phi(K^*)) \Big) \\
&= P_0^i \Big(\frac{\eta_i}{\sqrt{2\pi}} \exp\Big(-\frac{1}{2}(K^*)^2 \Big) + (y_0^i - K)[(1 - \Phi(K^*))] \Big). \qquad \text{(B.1)}
\end{aligned}
$$

We find the η_i parameter value by equating the ATM swaption value obtained under log-normal assumption to the one obtained under Gaussian distribution assumption. Assuming $K = y_0^i$, under the Gaussian assumption, following equation (B.1), the ATM swaption value is given by:

$$
(V_0^i(y_0^i))^N = P_0^i \frac{\eta_i}{\sqrt{2\pi}}. \qquad \text{(B.2)}
$$

Under the log-normality assumption, with swaption implied volatility given by σ_i, the ATM swaption value is given by:

$$
(V_0^i(y_0^i))^{\mathrm{LN}} = P_0^i y_0^i \Big(2\Phi\Big[\frac{1}{2}\sigma_i \sqrt{T}\Big] - 1 \Big). \qquad \text{(B.3)}
$$

Equating (B.2) and (B.3), it follows that

$$
\eta_i = \sqrt{2\pi} y_0^i \Big(2\Phi\Big[\frac{1}{2}\sigma_i \sqrt{T}\Big] - 1 \Big).
$$

This choice of η_i will give the same ATM swaption price for any value of γ_i chosen.

B.1 Skew effect on convexity corrections

In section 3.2.3, we introduced a skew in the implied volatility curve by assuming that the the marginal distribution of y_T^i under its associated swaption measure $\mathbb{S}^i$ is given by a mixture of normal and log-normal distributions. The mixture is controlled by the weight parameter γ. We investigate the skew effect on convexity corrections for three values of γ : $0.1, 0.5$ and 0.9. Only the results for $\gamma = 0.5$ was reported in the earlier sections. We include here the results for $\gamma = 0.1$ and $\gamma = 0.9$.

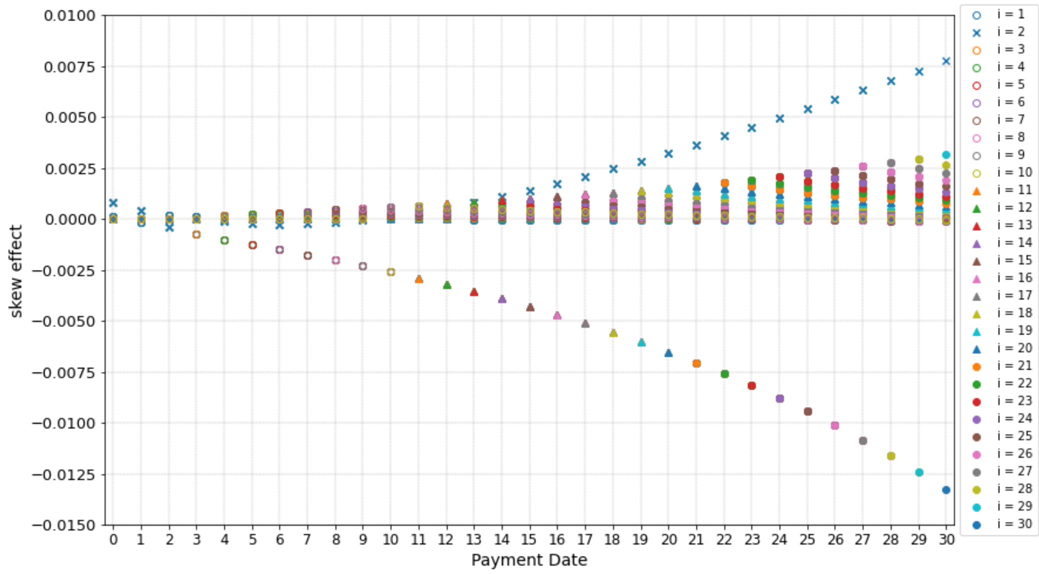

Figure 15: Skew effect
$\gamma = 0.1, N = 2, \tilde{M} = 30, y_0 = 0.07, D_{0T} = 1.0$

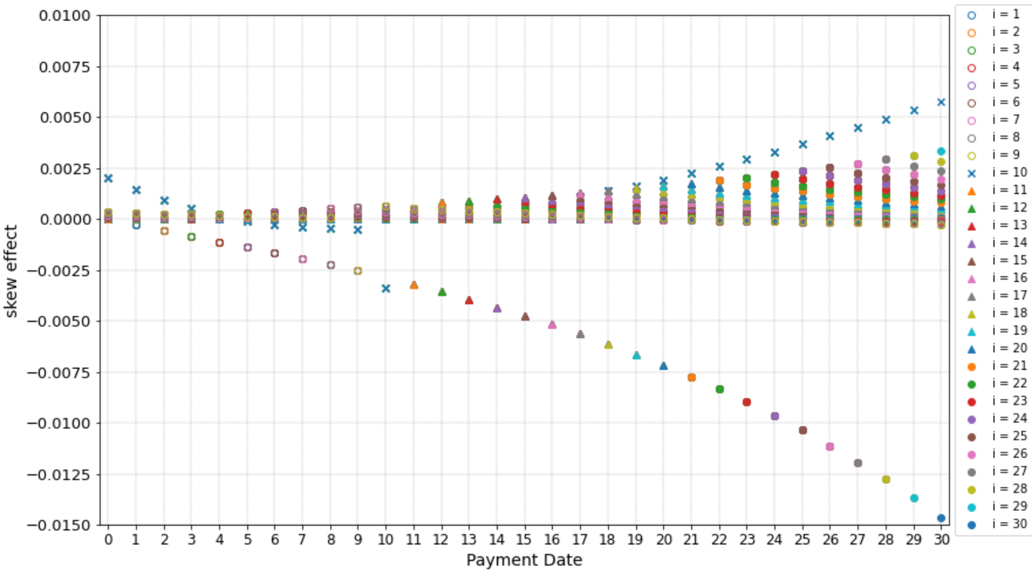

Figure 16: Skew effect
$\gamma = 0.1, N = 10, \tilde{M} = 30, y_0 = 0.07, D_{0T} = 1.0$

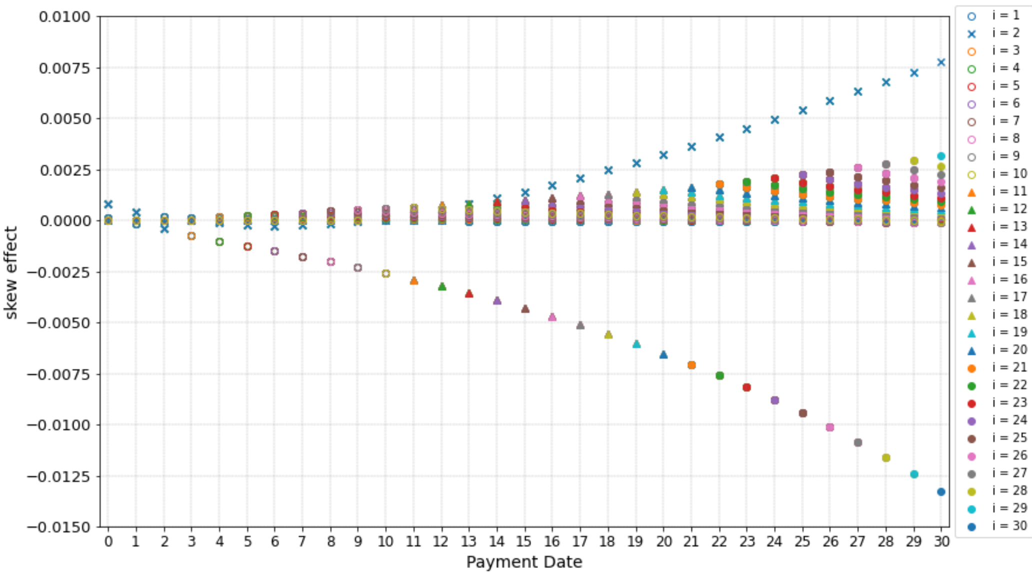

Figure 17: Skew effect
$\gamma = 0.9, N = 2, \tilde{M} = 30, y_0 = 0.07, D_{0T} = 1.0$

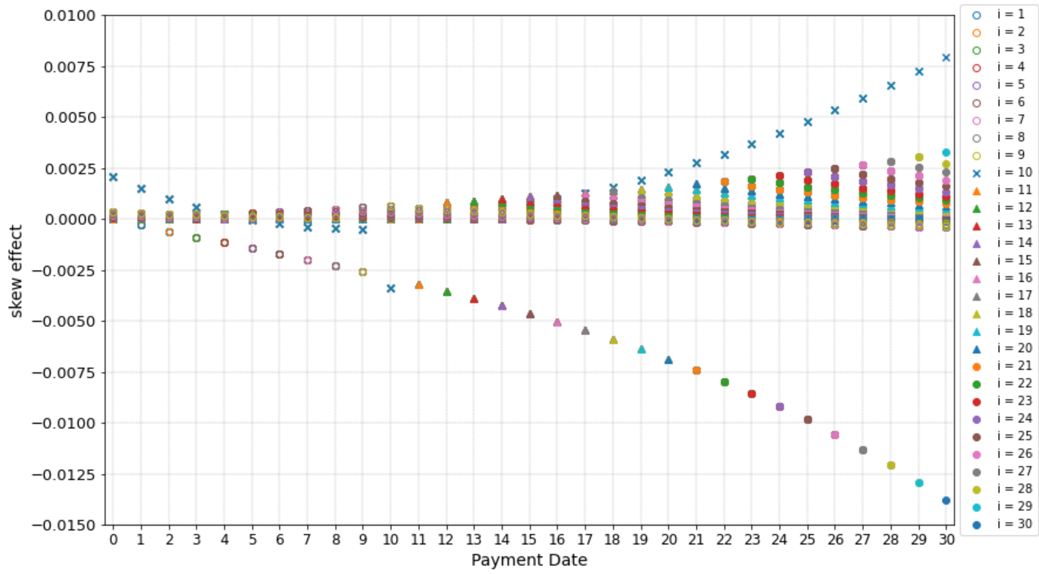

Figure 18: Skew effect
$\gamma = 0.9, N = 10, \tilde{M} = 30, y_0 = 0.07, D_{0T} = 1.0$

We highlighted in section 3.2.3 that at small payment dates (M = 0,1), the marginal distribution of the reference swap rate dominates the valuation since the skew effect $\Gamma_{M,N}$ is of bigger magnitude than the rest. On the other hand, when $M = 30$, the skew effect associated with the payment swap rate far outweighs the skew effect associated with the reference swap rate, but $|\Gamma_{30,N}| > |\Gamma_{30,i}|$ for $i \in \{1, \dots 29\} \backslash \{N\}$. This shows that as the payment date is taken further away from the reference swap rate maturity, both the reference swap rate and the payment swap rate are significant to the valuation of the CMS.

Upon closer observation, we further note the following:

1. The effect $\Gamma_{M,M}$ (i.e the skew effect associated with the payment swap rate) is always negative and increases in magnitude as the payment date increases

2. The effect $\Gamma_{M,N}$ (i.e the skew effect associated with the reference swap rate) starts off positive, decreases to some negative value, then increases back to some positive value for large payment dates

3. The skew effect for when the payment date coincides with the reference swap rate maturity is negative and larger in magnitude than $\Gamma_{M,N}$ for payment dates in the neighbourhood of S_N

We aim to provide an explanation for the above.

Numerical observation: The skew effect on the functional forms of the swap rates

In the single-time MFM, we chose to express the swap rates as a monotonic increasing function of the driver x taken to be a standard Gaussian random variable. When $\gamma = 0$, we are essentially assuming that the swap rates are normally distributed under their respective swaption measures. Under the forward measure $\mathbb{F}$, they will have a different distribution (We point out that the measure change involves the PVBP term, which is itself a complicated function of the swap rates), but they can roughly be assumed to be Gaussian. We therefore expect the function form of the swap rate to be (roughly) linear in x. When $\gamma = 1$, i.e, the swap rates are log-normally distributed under their respective swaption measures, we expect the functional forms to be convex (roughly an exponential function of the driver x). Hence as γ increases, the functional forms of the swap rates becomes more convex. We illustrate the functional forms for y_T^2 below. (The plot for y_T^{10} will be similar)

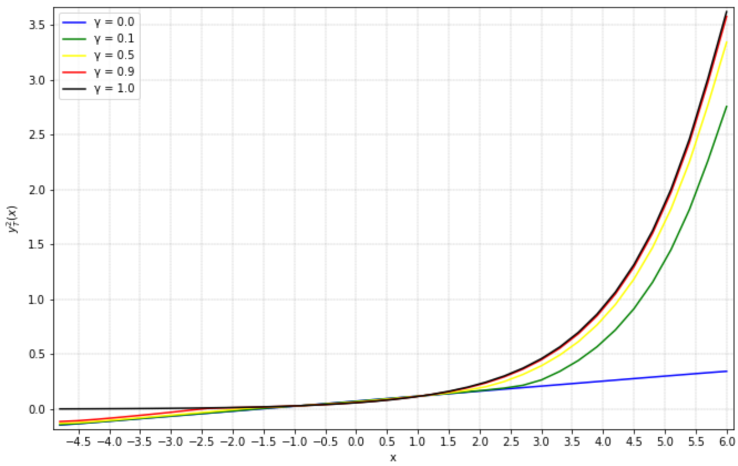

Figure 19: Functional forms of y_T^2 for a given γ
$y_0^2 = 0.07, \sigma_2 = 0.149$

We recall that the valuation of a single payment of the CMS under the forward measure $\mathbb{F}$ involves an expectation of the product of the reference swap rate and the pure discount bond with maturity given by the payment date of the CMS. In particular,

$$V_0^{CMS} = D_{0T}\mathbb{E}_{\mathbb{F}}[y_T^N D_{TS_M}] = D_{0T}\mathbb{E}_{\mathbb{F}}[y_T^N D_{TS_M}(y_T^1, y_T^2, \ldots, y_T^M)]$$

Assume we are varying the marginal distribution of y_T^N via γ^N. When $M < N$, the change in the valuation of the CMS will only come from the reference swap rate and not D_{TS_M}. As we observe from figure 19, an increase in γ causes the functional form to become more convex for $x > 0$. For $x < 0$, the function values are negative and they decrease in magnitude as γ increases. The impact of this on the expectation is that increasing γ^N pushes up the expected value, explaining the initial positive effect we see for $\Gamma_{M,N}$.

To further understand the trend in the effect, especially for $M > N$, we first work out the dependence between the swap rates and the pure discount bond. We recall the following relationship between D_{TS_M} and the swap rates:

$$D_{TS_M} := D_{TS_M}(y_T^1, \ldots, y_T^M) = 1 - \frac{y_T^M}{1 + \alpha_M y_T^M} \sum_{k=1}^{M-1} \alpha_k \prod_{l=k}^{M-1} \frac{1}{1 + \alpha_l y_T^l}.$$

If we want to understand how the pure discount bond D_{TS_M} changes as the payment swap rate y_T^M varies, we consider the first order derivative of D_{TS_M} with respect to y_T^M, keeping all the other variables y_T^i, for $i \in \{1, 2, \ldots, M-1\}$ constant. Note that the forward swap rates are related to each other through the model driver x, and we assume that they are all monotonically increasing in x. In the one-factor model, we could choose to express any swap rate in terms of a chosen forward swap rate, in the sense that:

$$y_T^i := \hat{f}^i(y_T^M), \qquad \text{for } i \in \{1, \ldots, \tilde{M}\},$$

where $\hat{f}^i : \mathbb{R} \to \mathbb{R}$ is some deterministic function. This is essentially the view we are taking when computing the derivative of D_{TS_M} with respect to y_T^M. We omit the dependence between the swap rates in order to isolate the effect of the the payment swap rate on D_{TS_M}. The analysis below only gives us a rough idea of the relationship between the pure discount bond and the forward swap rates.

$$\begin{aligned}
\frac{\partial D_{TS_M}}{\partial y_T^M} &= -\left(\sum_{k=1}^{M-1} \alpha_k \prod_{l=k}^{M-1} \frac{1}{1 + \alpha_l y_T^l} \right) \frac{\mathrm{d}}{\mathrm{d}x} \left(\frac{y_T^M(x)}{1 + \alpha_M y_T^M(x)} \right) \times \frac{1}{\frac{\mathrm{d}y_T^M}{\mathrm{d}x}} \\
&= -\left(\sum_{k=1}^{M-1} \alpha_k \prod_{l=k}^{M-1} \frac{1}{1 + \alpha_l y_T^l} \right) \times \frac{1}{(1 + \alpha_M y_T^M(x))^2} \times \frac{1}{\frac{\mathrm{d}y_T^M}{\mathrm{d}x}} \qquad < 0
\end{aligned}$$

It therefore follows that there is an inverse relationship between D_{TS_M} and y_T^M.

Fix $i \in \{1, \ldots, M-1\}$. We now view D_{TS_M} as a function of y_T^i. Keeping all other swap rates constant, we have that:

$$\begin{aligned}
\frac{\partial D_{TS_M}}{\partial y_T^i} &= -\frac{y_T^M}{1+\alpha_M y_T^M} \times \frac{\partial}{\partial y_T^i}\left(\sum_{k=1}^{i} \alpha_k \prod_{l=k}^{M-1} \frac{1}{1+\alpha_l y_T^l} + \sum_{k=i+1}^{M-1} \alpha_k \prod_{l=k}^{M-1} \frac{1}{1+\alpha_l y_T^l}\right) \\
&= -\frac{y_T^M}{1+\alpha_M y_T^M} \times \left(\sum_{k=1}^{i} \alpha_k \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{1+\alpha_i y_T^i(x)}\right) \times \frac{1}{\frac{\mathrm{d}y_T^i}{\mathrm{d}x}} \prod_{l=k,l\neq i}^{M-1} \frac{1}{1+\alpha_l y_T^l}\right) \\
&= \frac{y_T^M}{1+\alpha_M y_T^M} \times \left(\sum_{k=1}^{i} \alpha_k \frac{\alpha_i}{(1+\alpha_i y_T^i)^2} \times \frac{1}{\frac{\mathrm{d}y_T^i}{\mathrm{d}x}} \prod_{l=k,l\neq i}^{M-1} \frac{1}{1+\alpha_l y_T^l}\right) \\
&= \frac{\alpha_i y_T^M}{(1+\alpha_M y_T^M)(1+\alpha_i y_T^i)} \times \frac{1}{\frac{\mathrm{d}y_T^i}{\mathrm{d}x}} \times \left(\sum_{k=1}^{i} \alpha_k \prod_{l=k}^{M-1} \frac{1}{1+\alpha_l y_T^l}\right) \qquad > 0
\end{aligned}$$

The pure discount bond D_{TS_M} is monotonically increasing in y_T^i, for $i \in \{1, \ldots, M-1\}$.

When $M > N$, a change in γ^N will affect both y_T^N and D_{TS_M}, and we know from the earlier analysis, that increasing γ^N causes y_T^N to push up the expected value. This positive change is complemented by the pure discount bond D_{TS_M} owing to the direct relationship between D_{TS_M} and y_T^N, thereby explaining why for large M, in particular $M = 30$, $\Gamma_{M,N}$ is positive.

That $\Gamma_{M,M}$ is always negative for $M \in \{1, \ldots, \tilde{M}\}$, can be explained by the fact that the skew effect in this case is from the pure discount bond term and D_{TS_M} is inversely related to the variable y_T^M.

B.2 Finding the 'weight' terms $\xi_k^j(\cdot)$

We recall that under the assumption of separability, we set up a prior of the LIBORs using a PCA approach, and linked the model to the swap rates via a first-order Taylor approximation. We have the following approximation of the log of the swap rates expressed in terms of the log of the LIBORs:

$$\begin{aligned}
\ln(y_T^j) &\approx \ln(y_0^j) + \sum_{k=1}^{j} \frac{\tilde{\omega}_k^j(0)}{y_0^j} L_0^k(\ln(L_T^k) - \ln(L_0^k)) \\
&= \ln(y_0^j) + \sum_{k=1}^{j} \xi_k^j(0)(\ln(L_T^k) - \ln(L_0^k)),
\end{aligned}$$

where

$$\xi_k^j(0) := \frac{\tilde{\omega}_k^j(0)}{y_0^j} L_0^k, \tag{B.4}$$

and

$$\tilde{\omega}_k^j(0) = \omega_k^j(0) + \sum_{p=1}^{j} L_0^p \left(\frac{\partial \omega_p^j(T)}{\partial L_T^k} \right)_{T=0}. \tag{B.5}$$

We are interested in finding the exact form of $\xi_k^j(\cdot)$.

We begin by finding $\frac{\partial \omega_p^j(t)}{\partial L_t^k}$ explicitly.

We have that:

$$\omega_p^j(t) = \frac{\tau_p \prod_{i=1}^{p} \frac{1}{1+\tau_i L_t^i}}{\sum_{l=1}^{j} \tau_l \prod_{i=1}^{l} \frac{1}{1+\tau_i L_t^i}}.$$

Define $u^p(t) := \tau_p \prod_{i=1}^{p} \frac{1}{1+\tau_i L_t^i}$

$$\begin{aligned}
\frac{\partial u^p(t)}{\partial L_t^k} &= \frac{\partial}{\partial L_t^k} \left(\frac{\tau_p}{(1+\tau_1 L_t^1)(1+\tau_2 L_t^2)\dots(1+\tau_p L_t^p)} \right) \\
&= \begin{cases} \frac{-\tau_p \tau_k}{1+\tau_k L_t^k} \prod_{i=1}^{p} \frac{1}{1+\tau_i L_t^i} & \text{if } p \geq k \\ 0 & \text{if } p < k \end{cases}
\end{aligned}$$

Define $v(t) := \sum_{l=1}^{j} \tau_l \prod_{i=1}^{l} \frac{1}{1+\tau_i L_t^i}$

$$\begin{aligned}
\frac{\partial v(t)}{\partial L_t^k} &= \frac{\partial}{\partial L_t^k} \Bigg(\frac{\tau_1}{1+\tau_1 L_t^1} + \frac{\tau_2}{(1+\tau_1 L_t^1)(1+\tau_2 L_t^2)} + \dots \\
&\quad + \frac{\tau_k}{(1+\tau_1 L_t^1)(1+\tau_2 L_t^2)..(1+\tau_k L_t^k)} + \dots \frac{\tau_j}{(1+\tau_1 L_t^1)\dots(1+\tau_j L_t^j)} \Bigg) \\
&= \frac{-\tau_k.\tau_k}{1+\tau_k L_t^k} \prod_{i=1}^{k} \frac{1}{1+\tau_i L_t^i} + \dots + \frac{-\tau_k.\tau_j}{1+\tau_k L_t^k} \prod_{i=1}^{j} \frac{1}{1+\tau_i L_t^i} \\
&= \frac{-\tau_k}{1+\tau_k L_t^k} \left[\sum_{l=k}^{j} \tau_l \prod_{i=1}^{l} \frac{1}{1+\tau_i L_t^i} \right].
\end{aligned}$$

By the quotient rule, we have that:

$$\frac{\partial \omega_p^j(t)}{\partial L_t^k} = \frac{v(t) \frac{\partial u^p(t)}{\partial L_t^k} - u^p(t) \frac{\partial v(t)}{\partial L_t^k}}{(v(t))^2}.$$

If $p \geq k$,

$$\begin{aligned}\frac{\partial \omega_p^j(t)}{\partial L_t^k} &= \frac{\frac{\tau_p \tau_k}{1+\tau_k L_t^k}\prod_{i=1}^p \frac{1}{1+\tau_i L_t^i}\left[\sum_{l=k}^j \tau_l \prod_{i=1}^l \frac{1}{1+\tau_i L_t^i} - \sum_{l=1}^j \tau_l \prod_{i=1}^l \frac{1}{1+\tau_i L_t^i}\right]}{v(t).v(t)} \\ &= \frac{\tau_k}{1+\tau_k L_t^k}.\frac{\tau_p \prod_{i=1}^p \frac{1}{1+\tau_i L_t^i}}{\sum_{l=1}^j \tau_l \prod_{i=1}^l \frac{1}{1+\tau_i L_t^i}}\left[\frac{\sum_{l=k}^j \tau_l \prod_{i=1}^l \frac{1}{1+\tau_i L_t^i}}{\sum_{l=1}^j \tau_l \prod_{i=1}^l \frac{1}{1+\tau_i L_t^i}} - 1\right] \\ &= \frac{\tau_k \omega_p^j(t)}{1+\tau_k L_t^k}\left[\frac{\sum_{l=k}^j \tau_l \prod_{i=1}^l \frac{1}{1+\tau_i L_t^i}}{\sum_{l=1}^j \tau_l \prod_{i=1}^l \frac{1}{1+\tau_i L_t^i}} - 1\right].\end{aligned}$$

If $p < k$,

$$\begin{aligned}\frac{\partial \omega_p^j(t)}{\partial L_t^k} &= \frac{\tau_k \tau_p \prod_{i=1}^p \frac{1}{1+\tau_i L_t^i}}{(1+\tau_k L_t^k)\sum_{l=1}^j \tau_l \prod_{i=1}^l \frac{1}{1+\tau_i L_t^i}}.\frac{\sum_{l=k}^j \tau_l \prod_{i=1}^l \frac{1}{1+\tau_i L_t^i}}{\sum_{l=1}^j \tau_l \prod_{i=1}^l \frac{1}{1+\tau_i L_t^i}} \\ &= \frac{\tau_k \omega_p^j(t)}{1+\tau_k L_t^k}\left[\frac{\sum_{l=k}^j \tau_l \prod_{i=1}^l \frac{1}{1+\tau_i L_t^i}}{\sum_{l=1}^j \tau_l \prod_{i=1}^l \frac{1}{1+\tau_i L_t^i}}\right].\end{aligned}$$

We can therefore observe that

$$\frac{\partial \omega_p^j(t)}{\partial L_t^k} = \frac{\tau_k \omega_p^j(t)}{1+\tau_k L_t^k}\left[\frac{\sum_{l=k}^j \tau_l \prod_{i=1}^l \frac{1}{1+\tau_i L_t^i}}{\sum_{l=1}^j \tau_l \prod_{i=1}^l \frac{1}{1+\tau_i L_t^i}} - \mathbb{1}_{\{p \geq k\}}\right]. \tag{B.6}$$

Plugging equation (B.6) back into equation (B.5), we have that:

$$\begin{aligned}\tilde{\omega}_k^j(t) &= \omega_k^j(t) + \frac{\tau_k}{1+\tau_k L_t^k}\sum_{p=1}^j L_t^p \omega_p^j(t)\left[\frac{\sum_{l=k}^j \tau_l \prod_{i=1}^l \frac{1}{1+\tau_i L_t^i}}{\sum_{l=1}^j \tau_l \prod_{i=1}^l \frac{1}{1+\tau_i L_t^i}} - \mathbb{1}_{\{p \geq k\}}\right] \\ &= \omega_k^j(t) + \frac{\tau_k}{1+\tau_k L_t^k}\left[\sum_{p=1}^{k-1} L_t^p \omega_p^j(t)\left(\frac{\sum_{l=k}^j \tau_l \prod_{i=1}^l \frac{1}{1+\tau_i L_t^i}}{\sum_{l=1}^j \tau_l \prod_{i=1}^l \frac{1}{1+\tau_i L_t^i}}\right)\right. \\ &\left. - \sum_{p=k}^j L_t^p \omega_p^j(t)\left(\frac{\sum_{l=1}^{k-1} \tau_l \prod_{i=1}^l \frac{1}{1+\tau_i L_t^i}}{\sum_{l=1}^j \tau_l \prod_{i=1}^l \frac{1}{1+\tau_i L_t^i}}\right)\right],\end{aligned} \tag{B.7}$$

with the convention that $\sum_{i=1}^0 \ldots$ is an empty sum and is equal to 0. This gives us the exact formula for $\xi_k^j(t)$ by replacing equation (B.7) in equation (B.4).

Given initial conditions on L_0^j, we can easily derive the value at time 0 for $\xi_k^j(\cdot)$, for $k = 1, \ldots, j$, by setting $t = 0$ in equation (B.7), and using the fact that

$$\omega_k^j(0) = \frac{\tau_k \prod_{i=1}^k \frac{1}{1+\tau_i L_0^i}}{\sum_{l=1}^j \tau_l \prod_{i=1}^l \frac{1}{1+\tau_i L_0^i}} \quad \text{and} \quad y_0^j = \sum_{k=1}^j \omega_k^j(0) L_0^k.$$

C A crude approximation for the initial swap rates under shifted log-normal assumption

In section 4.3, we assume swap rates are (shifted) log-normally distributed under their respective swaption measures. To set up the 1F smcMFL model, we first fix a maturity S_n, and we postulate functional forms for the swap rate y_T^n and the PVBP P_T^n, in terms of a one-dimensional model driver. we recall, we model the PVBP as follows:

$$P_T^n = a_n \hat{P}_T^n(\eta_n),$$

where

$$\hat{P}_T^n(\eta_n) := \sum_{k=1}^{n} \alpha_k \bigg(\prod_{j=k}^{n} \frac{1}{1+\alpha_j \hat{y}_T^j} \bigg). \tag{C.1}$$

For $j \in \{1, \dots, n\}$, we express $\hat{y}_T^j$ as follows:

$$\hat{y}_T^j(\eta_n) := \theta_j + (\hat{y}_0^j - \theta_j) \exp(\hat{\sigma}_j \eta_n - \frac{1}{2}\hat{\sigma}_j^2),$$

where $\hat{y}_0^j$ is the convexity-adjusted forward swap rate. We now derive an approximation for $\hat{y}_0^j$ below. By a change of measure, we have that:

$$P_0^n \mathbb{E}_{\mathbb{S}^n}[y_T^j] = P_0^j \mathbb{E}_{\mathbb{S}^j}\left[y_T^j \frac{P_T^n}{P_T^j} \right].$$

Assuming all swap rates are equal, we can approximate

$$D_{tS_j} = \prod_{k=1}^{j} (1 + \alpha_k y_T^j)^{-1}.$$

We obtain:

$$\begin{aligned}
\log(D_{tS_j}) &= \sum_{k=1}^{j} \log(1 + \alpha_k y_t^j)^{-1} \\
\frac{\mathrm{d}D_{tS_j}}{D_{tS_j}} &= -\bigg(\sum_{k=1}^{j} \frac{\alpha_k}{1+\alpha_k y_t^j} \bigg) \mathrm{d}y_t^j + \text{f.v.}
\end{aligned}$$

We have that:

$$\begin{aligned}
y_T^j P_T^j &= 1 - D_{TS_j} \qquad \text{therefore,} \\
y_t^j \mathrm{d}P_t^j + P_t^j \mathrm{d}y_t^j &= -\mathrm{d}D_{tS_j} + \text{f.v} \\
y_T^j \mathrm{d}P_t^j &= -(P_t^j \mathrm{d}y_t^j + \mathrm{d}D_{tS_j}) + \text{f.v} \\
&= -\left(P_t^j - D_{tS_j} \sum_{k=1}^{j} \frac{\alpha_k}{1+\alpha_k y_t^j}\right) \mathrm{d}y_t^j + \text{f.v} \\
\frac{\mathrm{d}P_t^j}{P_t^j} &= \left(\frac{D_{tS_j}}{P_t^j} \sum_{k=1}^{j} \frac{\alpha_k}{1+\alpha_k y_t^j} - 1\right) \frac{\mathrm{d}y_t^j}{y_t^j} + \text{f.v.}
\end{aligned}$$

Assuming $\mathrm{d}y_t^j = \hat{\sigma}_j (y_t^j - \theta_j) \mathrm{d}W_t$ it follows that:

$$\begin{aligned}
\frac{\mathrm{d}P_t^j}{P_t^j} &= \left(\frac{D_{tS_j}}{P_t^j} \sum_{k=1}^{j} \frac{\alpha_k}{1+\alpha_k y_t^j} - 1\right) \frac{\hat{\sigma}_j (y_t^j - \theta_j)}{y_t^j} \mathrm{d}W_t + \text{f.v} \\
&\approx \left[\left(\frac{D_{0S_j}}{P_0^j} \sum_{k=1}^{j} \frac{\alpha_k}{1+\alpha_k y_0^j} - 1\right) \frac{\hat{\sigma}_j (y_0^j - \theta_j)}{y_0^j}\right] \mathrm{d}W_t + \text{f.v.}
\end{aligned}$$

Similarly,

$$\frac{\mathrm{d}P_t^n}{P_t^n} \approx \left[\left(\frac{D_{0S_n}}{P_0^n} \sum_{k=1}^{n} \frac{\alpha_k}{1+\alpha_k y_0^n} - 1\right) \frac{\hat{\sigma}_n (y_0^n - \theta_n)}{y_0^n}\right] \mathrm{d}W_t + \text{f.v.}$$

By Ito's lemma, we have that:

$$\begin{aligned}
\mathrm{d}\left(\frac{P_t^n}{P_t^j}\right) &= \frac{\mathrm{d}P_t^n}{P_t^j} - \frac{P_t^n}{(P_t^j)^2} \mathrm{d}P_t^j + \text{f.v} \\
&= \left(\frac{P_t^n}{P_t^j}\right) \frac{\mathrm{d}P_t^n}{P_t^n} - \left(\frac{P_t^n}{P_t^j}\right) \frac{\mathrm{d}P_t^j}{P_t^j} + \text{f.v} \\
&= \left(\frac{P_t^n}{P_t^j}\right) \left[\left(\frac{D_{0S_n}}{P_0^n} \sum_{k=1}^{n} \frac{\alpha_k}{1+\alpha_k y_0^n} - 1\right) \frac{\hat{\sigma}_n (y_0^n - \theta_n)}{y_0^n} - \right. \\
&\left. \left(\frac{D_{0S_j}}{P_0^j} \sum_{k=1}^{j} \frac{\alpha_k}{1+\alpha_k y_0^j} - 1\right) \frac{\hat{\sigma}_j (y_0^j - \theta_j)}{y_0^j}\right] + \text{f.v.}
\end{aligned}$$

Finally, we have that:

$$\begin{aligned}
\hat{y}_0^j = \mathbb{S}^n[y_T^j] &= \frac{P_0^j}{P_0^n}\mathbb{E}_{\mathbb{S}^j}\left[y_T^j \frac{P_T^n}{P_T^j}\right] \\
= \theta_j + (y_0^j - \theta_j)\exp\Bigg(\hat{\sigma}_j\Bigg[&\left(\frac{D_{0S_n}}{P_0^n}\sum_{k=1}^{n}\frac{\alpha_k}{1+\alpha_k y_0^n} - 1\right)\frac{\hat{\sigma}_n(y_0^n - \theta_n)}{y_0^n} - \\
&\left(\frac{D_{0S_j}}{P_0^j}\sum_{k=1}^{j}\frac{\alpha_k}{1+\alpha_k y_0^j} - 1\right)\frac{\hat{\sigma}_j(y_0^j - \theta_j)}{y_0^j}\Bigg]T\Bigg).
\end{aligned}$$

We observe that for $j = n$, $\hat{y}_0^n = \theta_n + (y_0^n - \theta_n) = y_0^n$ and the choice of functional form for $\hat{y}_T^n$ matches that of y_T^n, i.e., $\hat{y}_T^n = f_n(\cdot)$. For $\theta_j = 0$, i.e., the log-normal case, we have that:

$$\hat{y}_0^j = y_0^j \exp\left(\sigma_j\left[\sigma_n\left(\frac{D_{0S_n}}{P_0^n}\sum_{k=1}^{n}\frac{\alpha_k}{1+\alpha_k y_0^n} - 1\right) - \sigma_j\left(\frac{D_{0S_j}}{P_0^j}\sum_{k=1}^{j}\frac{\alpha_k}{1+\alpha_k y_0^j} - 1\right)\right]T\right).$$

We tabulate below the values used for $\hat{\sigma}_j$ when assuming that the swap rates are shifted log-normal in their own swaption measures.

$\theta = -0.05$			
j	$\hat{\sigma}_j$	j	$\hat{\sigma}_j$
1	0.0864	16	0.0804
2	0.0858	17	0.0801
3	0.0853	18	0.0798
4	0.0849	19	0.0795
5	0.0845	20	0.0792
6	0.0840	21	0.0789
7	0.0837	22	0.0786
8	0.0833	23	0.0784
9	0.0829	24	0.0781
10	0.0825	25	0.0779
11	0.0821	26	0.0776
12	0.0818	27	0.0774
13	0.0814	28	0.0771
14	0.0811	29	0.0769
15	0.0808	30	0.0767

$\theta = 0.02$			
j	$\hat{\sigma}_j$	j	$\hat{\sigma}_j$
1	0.2140	16	0.1983
2	0.2124	17	0.1975
3	0.2111	18	0.1967
4	0.2100	19	0.1959
5	0.2088	20	0.1951
6	0.2078	21	0.1944
7	0.2067	22	0.1936
8	0.2057	23	0.1929
9	0.2047	24	0.1923
10	0.2037	25	0.1916
11	0.2028	26	0.1910
12	0.2018	27	0.1904
13	0.2009	28	0.1898
14	0.2000	29	0.1892
15	0.1992	30	0.1886

Table 2: Implied Volatility parameters when assuming shifted log-normal distribution

D Correlation between the log of the swap rates

In Chapters 3 and 5, we have shown how to set up the correlation between the log of the swap rates. In the table below, we report the correlation used in the numerical analyses.

N = 2	$\lambda = 0.1$		$\lambda = 0.3$	
M	$\lambda = 0.02$	$\lambda_2 = 0.38$	$\lambda_2 = 0.02$	$\lambda_2 = 0.5$
1	1.000	0.999	0.998	0.998
2	1.000	1.000	1.000	1.000
3	1.000	0.999	0.999	0.998
4	0.999	0.998	0.996	0.993
5	0.998	0.995	0.991	0.985
6	0.997	0.991	0.987	0.973
7	0.996	0.986	0.982	0.959
8	0.995	0.981	0.978	0.943
9	0.994	0.974	0.974	0.926
10	0.992	0.967	0.970	0.908
11	0.991	0.959	0.967	0.890
12	0.990	0.951	0.964	0.873
13	0.988	0.942	0.961	0.855
14	0.987	0.934	0.959	0.839
15	0.986	0.925	0.957	0.823
16	0.984	0.916	0.955	0.809
17	0.983	0.907	0.953	0.795
18	0.982	0.898	0.951	0.782
19	0.981	0.889	0.950	0.770
20	0.980	0.881	0.949	0.759
21	0.979	0.872	0.947	0.749
22	0.978	0.864	0.946	0.740
23	0.977	0.857	0.945	0.731
24	0.976	0.849	0.944	0.723
25	0.976	0.842	0.944	0.716
26	0.975	0.836	0.943	0.709
27	0.974	0.829	0.942	0.703
28	0.973	0.823	0.942	0.697
29	0.973	0.817	0.941	0.691
30	0.972	0.812	0.940	0.686

Table 3: $\mathrm{Corr}(\log(y_T^2), \log(y_T^M))$
The parameter λ_2 is the variance of the second component of the driver, and is used to control the correlation structure; The parameter λ controls the shape of the second eigenvector

N = 10	$\lambda = 0.1$		$\lambda = 0.3$	
M	$\lambda = 0.02$	$\lambda_2 = 0.38$	$\lambda_2 = 0.02$	$\lambda_2 = 0.5$
1	0.989	0.958	0.954	0.884
2	0.992	0.967	0.970	0.908
3	0.995	0.975	0.982	0.931
4	0.996	0.982	0.989	0.951
5	0.998	0.988	0.994	0.967
6	0.999	0.992	0.997	0.980
7	0.999	0.996	0.998	0.989
8	1.000	0.998	0.999	0.996
9	1.000	1.000	1.000	0.999
10	1.000	1.000	1.000	1.000
11	1.000	1.000	1.000	0.999
12	1.000	0.998	1.000	0.997
13	1.000	0.997	0.999	0.994
14	0.999	0.994	0.999	0.990
15	0.999	0.991	0.999	0.985
16	0.999	0.988	0.998	0.981
17	0.998	0.984	0.998	0.976
18	0.998	0.980	0.998	0.971
19	0.997	0.976	0.997	0.966
20	0.997	0.972	0.997	0.962
21	0.997	0.968	0.997	0.958
22	0.996	0.964	0.996	0.953
23	0.996	0.960	0.996	0.949
24	0.996	0.956	0.996	0.946
25	0.995	0.952	0.996	0.942
26	0.995	0.948	0.995	0.939
27	0.995	0.944	0.995	0.936
28	0.994	0.941	0.995	0.933
29	0.994	0.937	0.995	0.930
30	0.994	0.934	0.995	0.928

Table 4: $\mathrm{Corr}(\log(y_T^{10}), \log(y_T^M))$
The parameter λ_2 is the variance of the second component of the driver, and is used to control the correlation structure; The parameter λ controls the shape of the second eigenvector

Note that the results reported here are obtained numerically from the single-time MFM (instead of using the approximate formula).

D.1 Correlation between $\log(y_T^N)$ and $\log(D_{TS_M})$

We tabulate below the correlation between the log of the reference swap rate y_T^N and the log of D_{TS_M} for $N = 2, 10$ and $M \in \{1, \ldots, 30\}$. We use this correlation structure as input to get the convexity corrections results reported in Figures 5.11- 5.14; the values are obtained numerically from the 2F single-time MFM.

N = 2	$\lambda = 0.1$		$\lambda = 0.3$	
M	$\lambda = 0.02$	$\lambda_2 = 0.38$	$\lambda_2 = 0.02$	$\lambda_2 = 0.5$
1	-0.918	-0.918	-0.917	-0.917
2	-0.928	-0.928	-0.928	-0.928
3	-0.938	-0.937	-0.936	-0.936
4	-0.946	-0.944	-0.941	-0.938
5	-0.953	-0.948	-0.944	-0.935
6	-0.959	-0.950	-0.945	-0.926
7	-0.963	-0.949	-0.945	-0.912
8	-0.967	-0.946	-0.944	-0.893
9	-0.969	-0.940	-0.943	-0.871
10	-0.971	-0.931	-0.941	-0.846
11	-0.971	-0.921	-0.939	-0.819
12	-0.971	-0.908	-0.937	-0.791
13	-0.970	-0.893	-0.934	-0.763
14	-0.968	-0.876	-0.931	-0.736
15	-0.966	-0.858	-0.927	-0.709
16	-0.963	-0.839	-0.923	-0.683
17	-0.959	-0.820	-0.919	-0.658
18	-0.955	-0.799	-0.915	-0.634
19	-0.950	-0.778	-0.911	-0.611
20	-0.945	-0.757	-0.906	-0.590
21	-0.940	-0.736	-0.901	-0.569
22	-0.934	-0.714	-0.895	-0.549
23	-0.928	-0.693	-0.890	-0.531
24	-0.921	-0.672	-0.884	-0.513
25	-0.914	-0.652	-0.878	-0.497
26	-0.907	-0.632	-0.871	-0.481
27	-0.900	-0.612	-0.865	-0.466
28	-0.892	-0.593	-0.858	-0.451
29	-0.885	-0.574	-0.851	-0.438
30	-0.877	-0.556	-0.844	-0.425

Table 5: $\mathrm{Corr}(\log(y_T^2), \log(D_{TS_M}))$
The parameter λ_2 is the variance of the second component of the driver, and is used to control the correlation structure; The parameter λ controls the shape of the second eigenvector

N = 10	$\lambda = 0.1$		$\lambda = 0.3$	
M	$\lambda = 0.02$	$\lambda_2 = 0.38$	$\lambda_2 = 0.02$	$\lambda_2 = 0.5$
1	-0.910	-0.881	-0.878	-0.813
2	-0.923	-0.899	-0.903	-0.845
3	-0.934	-0.916	-0.923	-0.876
4	-0.945	-0.932	-0.939	-0.905
5	-0.954	-0.946	-0.951	-0.929
6	-0.962	-0.957	-0.961	-0.949
7	-0.969	-0.967	-0.968	-0.962
8	-0.975	-0.973	-0.974	-0.971
9	-0.979	-0.978	-0.978	-0.975
10	-0.982	-0.979	-0.981	-0.974
11	-0.985	-0.979	-0.983	-0.969
12	-0.986	-0.975	-0.984	-0.962
13	-0.986	-0.970	-0.984	-0.952
14	-0.986	-0.963	-0.984	-0.941
15	-0.985	-0.954	-0.983	-0.929
16	-0.983	-0.943	-0.981	-0.916
17	-0.980	-0.931	-0.979	-0.902
18	-0.977	-0.918	-0.976	-0.888
19	-0.974	-0.904	-0.972	-0.874
20	-0.969	-0.889	-0.969	-0.860
21	-0.965	-0.873	-0.964	-0.846
22	-0.960	-0.857	-0.960	-0.833
23	-0.954	-0.841	-0.955	-0.819
24	-0.948	-0.825	-0.949	-0.806
25	-0.942	-0.808	-0.943	-0.793
26	-0.935	-0.792	-0.937	-0.780
27	-0.928	-0.775	-0.931	-0.767
28	-0.921	-0.759	-0.924	-0.755
29	-0.914	-0.743	-0.917	-0.743
30	-0.906	-0.727	-0.910	-0.731

Table 6: $\mathrm{Corr}(\log(y_T^{10}), \log(D_{TS_M}))$
The parameter λ_2 is the variance of the second component of the driver, and is used to control the correlation structure; The parameter λ controls the shape of the second eigenvector

E A guide to the numerical implementation of the models

E.1 A one-dimensional piecewise polynomial fit

The implementation of the models described in this thesis - and in particular, the single-time MFM, depends primarily on the ability to approximate a smooth-enough function on a bounded interval. We do so using a linear combination of piecewise polynomial functions. Let $[a, b]$ be a given interval and $a = z_1 < z_2 < \ldots < z_n = b$ be a partition of $[a, b]$. A *piecewise polynomial function* is a set of $n-1$ polynomial functions of a specified order q used to approximate a smooth-enough function. We can define a piecewise polynomial interpolation functional with respect to the above partition of a smooth function f on a bounded interval $[a, b]$ as follows:

$$I(f)(z) := \sum_{i=1}^{n} P_i(z),$$

where,

$$P_i(z) := \sum_{k=0}^{q} b_{i:k} z^k \mathbb{1}\{z_i \leq z < z_{i+1}\},$$

for $b_{i:k} \in \mathbb{R}$ are the coefficients of the polynomial. The polynomial function P_i is constructed such that $P_i(z_i) = f(z_i)$, for $i = 1, \ldots, n$. So we ensure that the resulting polynomial interpolation functional passes through every given coordinate. We also observe that the i^{th} q-order polynomial exists only on the i^{th} sub-interval and is zero everywhere else.

To provide a concrete example of where this algorithm can be applied, we recall from section 2.1.2 of Chapter 2 that we construct the functional form of the forward swap rate by calibrating the model to the known market-implied distribution of the swap rate in its own swaption measure. We know the values of the functional form of the swap rate, which we denoted by $f_.(z)$ for a given $z \in \mathbb{R}$ in equation (2.11). The question therefore is how can we interpolate and approximate the function $f_.$ given a set of coordinates $(z_i, f_.(z_i))$, $i \in \{1, \ldots, n\}$. The algorithm presented here allows us to approximate the functional form.

The construction of the piecewise polynomial interpolation functional is as follows: Suppose we are given a set of coordinates $(z_i, f(z_i)), i = 1, \ldots, n$, where $z_1 < z_2 < \ldots < z_n$ is a partition of a bounded interval and $f : \mathbb{R} \mapsto \mathbb{R}$ is the function we wish to approximate. We consider each sub-interval $[z_j, z_{j+1}]$, for $j = 1, \ldots, n-1$ to which we will fit the j^{th} q-order polynomial function. For j satisfying

$$\left\lfloor \frac{q-1}{2} \right\rfloor < j < \left\lceil m - \frac{q+1}{2} \right\rceil$$

we choose a set of coordinates $\left(z_{j+k}, f(z_{j+k})\right)_{k=-\lfloor\frac{q-1}{2}\rfloor}^{\lceil\frac{q+1}{2}\rceil}$, and we fit a q-order polynomial by solving the *Vandermonde system of equations* as described in Press, Teukolsky, Vetterling, and Flannery (1992).

When j does not satisfy the inequality - we observe that this relates to either ends of the function, i.e the first set of intervals or the last set of intervals; we choose the first $(q+1)$ coordinates that we will use to fit a q-order polynomial to the first sub-interval $[z_1, z_2]$ and fit the same polynomial to the j^{th} interval for $j < \lfloor\frac{q-1}{2}\rfloor$. On the other hand, for $j > \lceil m - \frac{q+1}{2}\rceil$, we fit the same polynomial to the j^{th} interval as the one that we fit over the interval $[z_{n-1}, z_n]$ using the last $q+1$ coordinates. By construction, $I(f)$ is the piecewise polynomial approximation of degree q of the function f defined on a bounded interval $[a, b]$ passing through $(z_i, f(z_i)), i = 1, \ldots, n$.

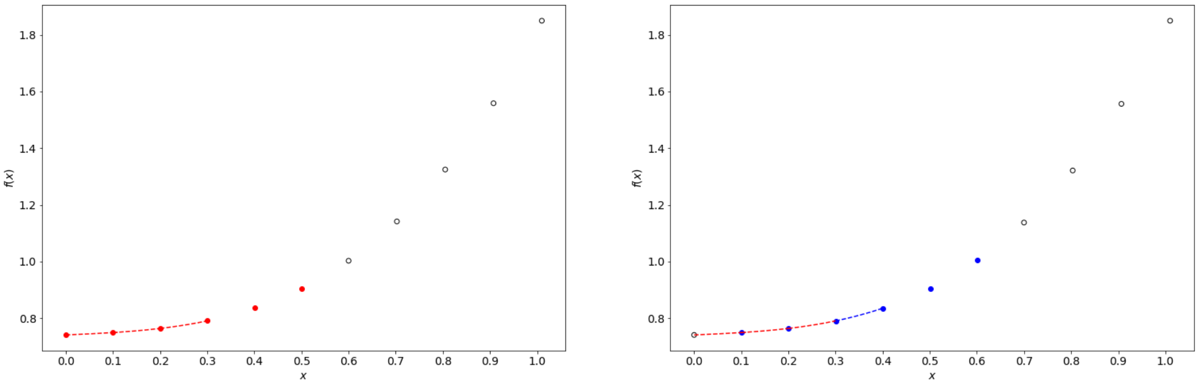

Figure 20: Piecewise polynomial fit

Figure 20 illustrates how the piecewise polynomial fit (of order 5) algorithm works. For the first three intervals, the same six coordinates shown in red are used to construct the polynomial that is fitted to the intervals $[0, 0, 1]$, $[0.1, 0.2]$ and $[0.2, 0.3]$. For the interval $[0.3, 0.4]$, the choice of coordinates used to fit the piecewise polynomial is illustrated in blue.

E.2 A two-dimensional piecewise polynomial fit

We have shown above how we do the piecewise polynomial fit in the one-dimensional case. In the thesis, we have developed two-factor models. If we use the model to price the payoff of a CMS or any related product, we will have to evaluate an expectation of the form $\mathbb{E}[f(X, Y)]$, which will involve a two-dimensional integration. We explain below how we extend the piecewise polynomial functional fit defined in the previous section to the two-dimensional case and carry out the required integration.

Assume we know the function values for a given set of tuples $(x_i, y_j), i = 1, \ldots, n$ and $j = 1, \ldots, m$. We define a two-dimensional piecewise polynomial functional as follows:

For each fixed x_i, we have a set of coordinates $\big(x_i, f(x_i, y_j)\big)_{j=1}^m$. We can use the one-dimensional piecewise polynomial fit to define an interpolation functional $I_i^y(f)$.

Remark 31*: Note firstly, we could have done the same step but fixing y_j and defining a piecewise polynomial functional in terms of the variable x. Secondly, we only define the piecewise polynomial functional within the interval specified. For any value outside of the interval the piecewise polynomial functional is assigned a value of zero.*

Suppose now we want to evaluate the function at (x, y) for some given $x_1 < x < x_n$ and $y_1 < y < y_m$. For each x_i , $i = 1, \ldots, n$, we evaluate $I_i^y(f)(y)$. We therefore have a set of coordinates $(x_i, I_i^y(f)(y))$, to which we can fit a one-dimensional piecewise polynomial functional, $I^x(f)$, and we can then evaluate $I^x(f)(x)$.

Similarly, if we want to compute the expectation $\mathbb{E}[f(X, Y)]$, by Tower property, we have that:

$$\begin{aligned}\mathbb{E}[f(X,Y)] &= \mathbb{E}\big[\mathbb{E}[f(X,Y)|X]\big] \\ &\approx \mathbb{E}\big[\mathbb{E}[I_i^y(f)]\big].\end{aligned}$$

We specify a vector of values $(x_i)_{i=1}^n$ [1], and we observe that the inner expectation involves the one-dimensional integration of the piecewise polynomial interpolation functional $I_i^y(f)$. For each x_i, we can compute the inner one-dimensional integration, the value of which we denote by $\iota(x_i)$. We can then construct a piecewise polynomial functional using $\big(x_i, \iota(x_i)\big)_{i=1}^n$, and we perform a one-dimensional integration again.

F A note on model performance and efficiency

We include below a brief summary of the computational performance of the two-factor models; in particular, we compare the two-factor single-time MFM to the two-factor refined forward measure calibrated MF-Lite model (2F refined fmcMFL model). Note that we only consider the two-factor case since we are more concerned with the computation efficiency of higher-factor models than one-factor models. We provide a comparison for the memory usage (heap memory allocation: the models were implemented in C++) and the duration to compute the price of a single payment of a CMS with reference index $N = 2$ and payment index $M = 29$, using the same parameters and model setup as discussed in Section 5.2.2. Recall that we assume the swap rates are log-normally distributed in their own swaption measures and the model driver is jointly Gaussian under the forward measure $\mathbb{F}$.

[1] Say X is standard Gaussian, we would take a partition of the interval [-5,5] for instance, we know that more than 99% of the probability mass will lie within this interval

Model	Memory usage	Duration
2F single-time MFM	22.1MB	34 seconds
2F refined fmcMFL model : Exact Fit	6.6 MB	16 seconds
2F refined fmcMFL model : Alternate Fit	7.3MB	22 seconds

Table 7: Performance of the two-factor models to compute $D_{0T}\mathbb{E}_{\mathbb{F}}[y_T^2 D_{TS_{29}}]$

We note that the 2F refined fmcMFL model takes less that half the time of the 2F single-time MFM, which is expected, by construction. We equally point out that the exact fit approach using the 2F refined fmcMFL model uses the least memory space, since to compute any given expectation, we only set the model up for the maturities we consider. In other words, we only store the functional forms required for these maturities. On the other hand, for the single-time MFM, since we construct the model by unfolding forward in maturity and the functional forms of the PVBP for any given maturity is dependent on the functional forms of the preceding swap rates down to S_1, we need to store all the functional forms. The performance for the alternate fit model is only negligibly worse than the exact fit approach (since we do the single-maturity setup for three maturities instead of two). We should point out that for this numerical analysis, we are only considering a joint Gaussian distribution on the model driver. Going back to the numerical implementation discussed in the section above, when constructing the functional forms, we consider an equally-spaced partition of $[-4.8, 6.9]$ with intervals of size 0.3 for both components of the driver. For heavier-tailed distributions, we might need a wider interval and smaller partitions, hence the need to evaluate the functional forms for a bigger set of values, which would increase the computation time. Nonetheless, the difference in model performance will stay qualitatively the same. In other words, we expect the 2F refined fmcMFL model to be faster the 2F single-time MFM model.

Bibliography

Andersen, L., & Andreasen, J. (2000). Volatility skews and extensions of the LIBOR market model. *Applied Mathematical Finance*, *7*(1), 1–32.

Andersen, L., & Piterbarg, V. (2010a). *Interest rate modeling volume I: Foundations and vanilla models* (1st). Atlantic Financial Press.

Andersen, L., & Piterbarg, V. (2010b). *Interest rate modeling, volume III: Products and Risk Management* (1st Edition.). Atlantic Financial Press.

Bank for international settlements. (2022). https://www.bis.org/statistics/derstats.htm?m=2071

Benhamou, E. (2000). *A martingale result for convexity adjustment in the Black pricing model* (LSE Working Paper) [http://dx.doi.org/10.2139/ssrn.265276].

Bennett, M., & Kennedy, J. (2005). A comparison of Markov-functional and Market Models : The one-dimensional case. *The Journal of Derivatives*, *13*, 22–43.

Bermin, H., & Williams, G. (2017). On cash settled IRR-swaptions and Markov-functional modeling. *World Scientific*, *20*((02)), 1750009.

Brace, A., Gatarek, D., & Musiela, M. (1997). The Market Model of Interest Rate Dynamics. *Mathematical Finance*, *7*, 127–147.

Brigo, D., & Mercurio, F. (2006). *Interest rate models: Theory and practice with smile, inflation and credit* (Second Edition.). Springer Finance.

Cedervall, S., & Piterbarg, V. (2012). Full implications for CMS convexity. *Asia Risk*, 45.

Chen, B., Oosterlee, C., & Van Weeren, S. (2010). Analytical approximation to constant maturity swap convexity corrections in a multi-factor SABR model. *International Journal of Theoretical and Applied Finance*, *13*(07), 1019–1046.

Choy, B., Dun, T., & Schlögl, E. (2003). Correlating market models. *Available at SSRN 395640*.

Cox, J., & Ross, S. (1976). The valuation of options for alternative stochastic processes. *Journal of Financial Economics*, *3*(1), 145–166.

Delbaen, F., & Schachermayer, W. (1994). A general version of The Fundamental Theorem of Asset Pricing. *Mathematische annalen*, *300*(1), 463–520.

Delbaen, F., & Schachermayer, W. (1999a). The Fundamental Theorem of Asset pricing for unbounded stochastic processes.

Delbaen, F., & Schachermayer, W. (1999b). Non-arbitrage and the Fundamental Theorem of Asset Pricing: Summary of main results. *Proceedings of Symposia in Applied Mathematics*, *57*, 49–58.

Devroye, L. (1986). *Non-uniform random variate generation*. Springer-Verlag.

Di Francesco, M., & Kamm, K. (2022). On the deterministic-shift extended CIR model in a negative interest rate framework. *International Journal of Financial Studies*, *10*(2), 38.

Filipovic, D. (2009). *Term-structure models. a graduate course*. Springer.

Gatheral, J., & Jacquier, A. (2014). Arbitrage-free SVI volatility surfaces. *Quantitative Finance*, *14*(1), 59–71.

Guo, C. (2016). *The stochastic volatility Markov-functional model*, University of Warwick. http://wrap.warwick.ac.uk/91418/

Harrison, J., & Kreps, D. (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic theory*, *20*(3), 381–408.

Henrard, M. (2007). *CMS swaps in separable one-factor gaussian LLM and HJM model* (Munich Personal RePEc Archive) [https://mpra.ub.uni-muenchen.de/3228/].

Henrard, M. (2014). *Interest rate modelling in the multi-curve framework: Foundations, evolution and implementation*. Springer.

Hull, J. (1997). *Options, futures and other derivatives* (Third Edition.). Prentice Hall.

Hunt, P., & Kennedy, J. (2004). *Financial derivatives in theory and practice* (Second Edition). John Wiley; Sons.

Hunt, P., & Kennedy, J. (2000). *Financial derivatives in theory and practice* (First Edition.). Wiley Series in Probability & Statistics.

Hunt, P., Kennedy, J., & Pelsser, A. (2000). Markov-functional interest rate models. *Finance and Stochastics, Springer*, *4*((4)), 391–408.

Kaisajuntti, L., & Kennedy, J. (2014). Stochastic volatility for interest rate derivatives. *Quantitative Finance*, *14*(3), 457–480.

Karatzas, I., & Shreve, S. (1991). *Brownian motion and stochastic calculus* (Vol. 113). Springer Science & Business Media.

Kennedy, J. (2010). Markov functional models. *Encyclopedia of Quantitative Finance*.

Kennedy, J., Mitra, S., & Pham, D. (2012). On the approximation of the SABR model: A probabilistic approach. *Applied Mathematical Finance*, *19*(6), 553–586.

Lord, R., & Pelsser, A. (2007). Level–slope–curvature–fact or artefact? *Applied Mathematical Finance*, *14*, 105–130.

Miltersen, K., Sandmann, K., & Sondermann, D. (1997). Closed form solutions for term structure derivatives with log-normal interest rates. *Journal of Finance*, *52*, 409–430.

Murgoci, A., & Gaspar, R. M. (2016). Convexity adjustments for affine term structure models [https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1399323]. *SSRN: Social Science Research Network*.

Pedersen, M. B. (1998). Calibrating LIBOR market models. *Available at SSRN 111828*.

Pelsser, A. (2000). *Efficient methods for valuing interest rate derivatives* (First Edition.). Springer.

Pietersz, R., Pelsser, A., & Van Regenmortel, M. (2004). Fast drift approximated pricing in the BGM model. *Journal of Computational Finance*, *8*(1).

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1992). *Numerical recipes in C*. Cambridge university press Cambridge.

Rebonato, R. (1998). Interest-rate option models: Understanding, analysing and using models for exotic interest-rate options.

Rebonato, R. (2002). *Modern pricing of interest rate derivatives: The LIBOR market model and beyond*. Princeton University Press: Princeton, NJ.

Rebonato, R. (2005). *Volatility and correlation: The perfect hedger and the fox*. John Wiley & Sons.

Revuz, D., & Yor, M. (2013). *Continuous martingales and Brownian motion* (Vol. 293). Springer Science & Business Media.

Russo, V., & Torri, G. (2019). Calibration of one-factor and two-factor Hull-White models using swaptions. *Computational Management Science*, *16*(1-2), 275–295.

Schachermayer, W. (1994). Martingale measures for discrete-time processes with infinite horizon. *Mathematical Finance*, *4*(1), 25–55.

Wu, T., & Chen, S. N. (2010). Modifying the LMM to price constant maturity swaps. *The Journal of Derivatives*, *18*((2)), 20–32.