Finite dimensional Markovian realizations for stochastic volatility forward rate models

Tomas Björk,
Department of Finance,
Stockholm School of Economics,
Box 6501,
SE-113 83 Stockholm, SWEDEN
tomas.bjork@hhs.se

Camilla Landén,
Department of Mathematics
Royal Insitute of Technology
SE-100 44 Stockholm, SWEDEN
camilla@math.kth.se

Lars Svensson
Department of Mathematics
Royal Insitute of Technology
SE-100 44 Stockholm, SWEDEN
larss@math.kth.se

March 2001

This version: May 6, 2002

Abstract

We consider forward rate rate models of Heath-Jarrow-Morton type, as well as more general infinite dimensional SDEs, where the volatility/diffusion term is stochastic in the sense of being driven by a separate hidden Markov process. Within this framework we use the previously developed Hilbert space realization theory in order provide general necessary and sufficent conditions for the existence of a finite dimensional Markovian realizations for the stochastic volatility models. We illustrate the theory by analyzing a number of concrete examples.

Keywords HJM models, stochastic volatility, factor models, forward rates, state space models, Markovian realizations, infinite dimensional SDEs.

JEL Classification: E43, G13

Contents

1 Introduction

The main object under study in this paper is a general forward rate model of Heath-Jarrow-Morton type (see [19]) with "stochastic volatility". The stochastic volatility is modeled by allowing the volatility term in the forward rate equation to depend on a hidden Markov process, as well as on the present forward rate curve. The goal of the paper is to investigate when and how the given, inherently infinite dimensional, stochastic volatility forward rate model, admits a finite dimensional Markovian realization in terms of a finite dimensional diffusion process. Since many of the results do not depend upon the particular structure of a forward rate model (where the drift is uniquely determined by the volatility through the HJM drift condition) the investigation below is to a large extent carried out for an arbitrary infinite dimensional stochastic differential equation where the diffusion term which is allowed to depend upon a hidden Markov process. The general results obtained are then specialized to the particular case of a forward rate model.

There exists a substantial literature on finite dimensional realizations (FDRs) for forward rate models with non-stochastic volatility. For special cases of the HJM volatility structure, the existence of an FDR is investigated in [1], [5], [9], [10], [14], [11], [21], [22], and [24]. In all these special cases, the existence of an FDR is proved by actually constructing a concrete realization.

A more systematic study of the general FDR problem from a geometric point of view was first undertaken in a series of papers [4], [7], and [6] (see [3] for an overview). In [7], Björk and Svensson provide the first general necessary and sufficient conditions for the existence of an FDR for SDEs in Hilbert space. The main technical tool is the Frobenius theorem. and the main result is that there exists an FDR if and only if the Lie algebra generated by the (Stratonovich) drift and diffusion terms is finite dimensional. This general result was then used in order to analyze a number of special cases, thereby including and extending the earlier results (see above) in the field.

The results in [7] were, however, pure existence results and no concrete realizations were constructed. The problem of actually constructing an FDR for a given model was then studied in [6], where the authors presented a systematic method for the construction of a concrete realization from a knowledge of the structure of the underlying Lie algebra.

The FDR problem is intimately related to the so called "consistency problem" for infinite dimensional SDEs. This problem was first formulated and discussed in [4], extended in [15], and then investigated in great detail in [16].

While in one sense the general FDR problem was more or less completely solved in [7], a major technical problem was still remaining. This had to do with the fact that in [7] the framework was that of strong solutions of infinite dimensional SDEs in Hilbert space and this forced Björk and Svensson to construct a particular Hilbert space of real analytic functions as their space of forward rate curves. While serving reasonably well, it was clear that this particular space was very small, and in particular it was pointed out by Filipović and Teichmann that the space does not include the forward rate curves generated by the Cox-Ingersoll-Ross model (see [12]). It was therefore necessary to extend the theory to a larger space but such an extension is far from to trivial to carry out, the problem being that on a larger Hilbert space you will loose the smoothness of the differential operator $\partial/\partial x$ appearing in the drift term of the forward rate equation. This problem was overcome with great elegance by Filipović and Teichmann who, partly building on the geometric and analytic results from [16], in [17] managed to extend the Lie algebraic FDR theory to a much larger space of forward rate curves than the one considered in [7]. In doing so, Filipović and Teichmann first extend the space of [7] to a larger Hilbert space. On the new space, however, the operator $\partial/\partial x$ becomes unbounded so they then change the topology on the space, thus making it into a Frechet space where the operator in fact is bounded. This approach, however, leads to new problems, since on a Frechet space there is no easy way of introducing differential calculus–in fact there is even no obvious way of defining the concept of smoothness which is necessary in order to have a Frobenius theorem. In order to overcome this problem, Filipović and Teichmann used the framework of so called "convenient spaces" developed some ten years ago (see [17] for references) in order to carry out analysis on the enlarged space. The main result of all this is that the Lie algebra conditions obtained by Björk and Svensson are shown to still hold in this more general setting. At this point it is worth mentioning that the technical price one has to pay for going into the deep parts of the theory of convenient analysis is quite high. It is therefore fortunate that the Lie algebraic machinery of [17] can be used without going into these (sometimes very hard) technical details. In fact, one of the main result of [17] can be formulated in the following pedestrian terms for the working mathematician: "When you are searching for FDRs for equations of HJM type, you can compute the relevant Lie algebra without worrying about the space". In [17] and in the follow up paper [18], the extended Lie algebra theory in [17] is used in to analyze a number of concrete problems concerning the forward rate equation, and in particular it was shown that any forward rate model admitting an FDR must necessarily have an affine term structure.

2 Basics

In this section we give the basic definitions, present the stochastic volatility model, and provide a precise formulation of the main problem to be treated.

2.1 Setup and model specification

As in Heath, Jarrow and Morton [20], we consider a default free bond market living on a filtered probability space $\{\Omega, \mathcal{F}, Q, \{\mathcal{F}_t\}_{t\geq 0}\}$ carrying an m-dimensional Wiener process W. Let $p(t,x)$ denote the price at time $t \geq 0$ of a zero-coupon bond with maturity $t+x$. Note that we use the so called Musiela parameterization (see [8] and [23]) with x denoting time **to** maturity, rather than the standard HJM parameterization with T denoting time **of** maturity. The instantaneous forward rate $r_t(x)$ is defined as usual by

$$r_t(x) = -\frac{\partial}{\partial x} \ln p(t,x),$$

and the short rate R is defined by

$$R_t = r_t(0).$$

We assume that the bond market is is arbitrage-free in the sense that the probability measure Q is a martingale measure for the model. In other words;

for each $T \geq 0$ we assume that $p(t, T-t)/B_t$ is a Q-martingale for $t \leq T$, where B denotes the money account defined by

$$B_t = \exp\left\{\int_0^t R_s ds\right\}.$$

In the sequel we will mainly concentrate on the entire forward rate curve $x \mapsto r_t(x)$, as opposed to the individual forward rate $r_t(x)$, and in order to emphasize this point of view the forward rate curve at time t will henceforth be denoted by r_t. The forward rate process $\{r_t;\ t \geq 0\}$ is thus a stochastic process taking values in a function space $\mathcal{H}$ of forward rate curves. In the present paper the precise choice of space $\mathcal{H}$ is in fact left to the reader. He/she can read the entire paper either within the restricted but technically simple framework of [7], or within the more general but technically more complicated framework of [17]. Regardless of the choice of space, all computations and all results (with one exception) will be the same (see the comment at the end of the introduction above). The only exception is that the discussion concerning the CIR model has to be read within the [17] framework. For easy reference we include an appendix with the main results of [7]. For all details concerning regularity requirements and the precise functional analytic setup we refer the reader to [7] and [17].

The main object under study in the present paper is a HJM model of the forward rates, with a stochastic volatility driven by a k dimensional hidden Markov process y. The model is defined as follows.

Definition 2.1 *The Itô formulation of the* **stochastic volatility model** *(henceforth SVM) is defined as the process pair (r, y), where the Q-dynamics of r and y are defined by the following system of SDEs.*

$$dr_t(x) = \left\{\frac{\partial}{\partial x} r_t(x) + \mathbf{H}\sigma(r_t, y_t, x)\right\} dt + \sigma(r_t, y_t, x) dW_t \tag{1}$$

$$dy_t = a^0(y_t)dt + b(y_t)dW_t, \tag{2}$$

where $\mathbf{H}$ is defined by

$$\mathbf{H}\sigma(r, y, x) = \sigma(r, y, x)\int_0^x \sigma^\star(r, y, s)ds, \tag{3}$$

and * denotes transpose.

In this specification we consider the following objects as given a priori:

- The volatility structure σ for the forward rates, i.e. a deterministic mapping
$$\sigma : \mathcal{H} \times R^k \times R_+ \to R^m.$$
- The drift vector field a^0 for y, i.e. a deterministic mapping
$$a^0 : R^k \to R^k.$$
(The superindex on a^0 will be explained below)

- The volatility vector field b for y, i.e. a deterministic mapping

$$b : R^k \to M(k, m).$$

where $M(k, m)$ denotes the set of $k \times m$ matrices.

We view σ as a row vector

$$\sigma(r, y, x) = [\sigma_1(r, y, x), \ldots, \sigma_m(r, y, x)],$$

the drift a^0 is viewed as a column vector

$$a^0(y) = \begin{bmatrix} a_1^0(y) \\ \vdots \\ a_k^0(y) \end{bmatrix},$$

and the volatility b is a matrix

$$b(y) = \begin{bmatrix} b_{11}(y) & b_{12}(y) & \cdots & b_{1m}(y) \\ b_{21}(y) & b_{22}(y) & \cdots & b_{2m}(y) \\ \vdots & \vdots & & \vdots \\ b_{k1}(y) & b_{k2}(y) & \cdots & b_{km}(y) \end{bmatrix},$$

We note in particular that the forward rate volatility σ is allowed to be an arbitrary functional of the entire forward rate curve r, as well as a function of the k-dimensional variable y. We may also view each component of σ as a mapping from $\mathcal{H} \times R^k$ to a space of functions (parameterized by x), and we will in fact assume that each σ_i, viewed in this way, is a smooth mapping with values in $\mathcal{H}$, i.e.

$$\sigma_i : \mathcal{H} \times R^k \to \mathcal{H}.$$

We make the following regularity assumptions.

Assumption 2.1 *From now on we assume that:*

- *The mappings* $\sigma_i : \mathcal{H} \times R^k \to \mathcal{H}$ *are smooth for* $i = 1, \ldots, m$.
- *The mapping* $\mathbf{H}\sigma : \mathcal{H} \times R^k \to \mathcal{H}$, *defined by (3) is smooth.*
- *The mappings* a^0 *and* b *are smooth on* R^k.

In the forward rate dynamics (1) we recognize the drift term in the r-dynamics above as the HJM drift condition, transferred into the Musiela parameterization. Note the particular structure of the equations (1)-(2): The y-process is feeding the drift and diffusion terms of the r-dynamics, but the r-process does not appear in the y-dynamics. Thus the y process is a Markov process in its own right, but this is not the case for the r-process. The extended process $\hat{r} = (r, y)$ is, however, Markovian. In many applications we want r and y to be driven by independent Wiener processes, and this case is of course included in our setup by choosing σ and b such that r and y are driven by different components of the (multidimensional) W process, i.e. by choosing b and σ such that $\sigma b^\star = 0$.

In many applications it is natural to study, not only the full SVM above but also a restricted model, where we forget about the dynamics of y and consider y as a constant parameter. In this way we obtain a *parameterized* model, and the formal definition is as follows.

Definition 2.2 *Consider the SVM defined by (1)-(2) above. For any* **fixed** *value of $y \in R^k$, the induced* **parameterized forward rate model** *is defined by the dynamics*

$$dr_t^y(x) = \left\{ \frac{\partial}{\partial x} r_t^y(x) + \mathbf{H}\sigma(r_t^y, y, x) \right\} dt + \sigma(r_t^y, y, x) dW_t. \tag{4}$$

Note that in the parameterized model, the forward rate process r^y itself is Markovian, whereas this is not the case in the full stochastic volatility model. For ease of reading we will sometimes drop the superscript ^{y}.

2.2 Problem formulation

The basic problem to be discussed in this paper is under what conditions the, inherently infinite dimensional, SVM defined above by (1)-(2), with given initial conditions $r_0 = r^0, y_0 = y^0$, admits a **generic finite dimensional Markovian realization** in the sense of [7]. More precisely we thus want to investigate under what conditions the extended process $\hat{r}_t = (r_t, y_t)$ possesses a local representation of the form

$$\hat{r}_t = \hat{G}(Z_t), \quad Q - a.s. \tag{5}$$

where, for some d, Z satisfies a d-dimensional SDE of the form

$$\left\{ \begin{array}{rcl} dZ_t & = & A_0(Z_t)dt + B(Z_t)dW_t, \\ Z_0 & = & z_0. \end{array} \right. \tag{6}$$

and where $\hat{G}$ is a smooth map $G : R^d \to \mathcal{H} \times R^k$. The drift and diffusion terms A_0 and B are assumed to be smooth and of suitable dimensions.

In a realization of this kind, the objects $\hat{G}$, A_0, B and z_0 will typically depend upon the choice of starting point (r^0, y^0). The term "generic" above means that we demand that there exists a realization, not only for the given initial point (r^0, y^0), but in fact for all initial points (r_0, y_0) in a neighborhood of (r^0, y^0). When we speak of realizations in the sequel we always intend this to mean generic realizations.

Note that the state process Z above is driven by the same Wiener process as the $\hat{r}$ system, and that the realization above is assumed to hold almost surely and trajectory wise.

We may now formulate some natural problems:

Main problems:

- Find necessary and sufficient conditions for the existence of an FDR for a given stochastic volatility model.
- Assuming the existence of an FDR has been guaranteed, how do you construct it?
- How is the existence of an FDR for the full stochastic volatility model related to the existence of an FDR for the induced parameterized model? More precisely: is the existence of an FDR for the parameterized model necessary and/or sufficient for the existence of an FDR for the full model?

2.3 Test examples: I.

To give more concretion to the discussion, and to illustrate technique, we now present four simple and natural test examples, which will be recurrent throughout the paper. In all cases we consider the case with a scalar driving Wiener process W^r for the forward rates, a scalar y process and a scalar driving Wiener process W^y for the y process. Furthermore we assume that W^r and W^y are independent. To motivate our choice of examples we recall (see [2]) the following well known (non stochastic) HJM volatilities for the forward rates.

I. Hull-White extended Vasiček:

$$\sigma(r, x) = \sigma e^{-ax}. \tag{7}$$

Here a and the right hand side occurrence of σ are real constants. This HJM model has a short rate realization of the form.

$$dR_t = \{\Phi(t) - aR_t\}\, dt + \sigma dW_t, \tag{8}$$

where the deterministic function Φ depends on the initial term structure. The parameters σ and a are the same as in (7).

II. Hull-White extended Cox-Ingersoll-Ross:

$$\sigma(r, x) = \sigma\sqrt{r(0)} \cdot \lambda(x, \sigma, a), \tag{9}$$

Here a and the right hand side occurrence of σ are real constants, whereas λ is given by

$$\lambda(x, \sigma, a) = -\frac{\partial}{\partial x}\left(\frac{2(e^{\gamma x} - 1)}{(\gamma + a)(e^{\gamma x} - 1) + 2\gamma}\right), \tag{10}$$

where

$$\gamma = \sqrt{a^2 + 2\sigma^2}.$$

Also this HJM model admits a short rate realization, namely

$$dR_t = \{\Phi(t) - aR_t\}\, dt + \sigma\sqrt{R_t} dW_t \tag{11}$$

The role of Φ is as in the extended Vasiček model above.

It is now natural to ask if we can extend these models by allowing one or several parameters to be stochastic, and still retain the existence of a finite dimensional realization.

We consider the following extensions of the above volatility structures. In all cases we assume that the scalar y process has dynamics of the form

$$dy_t = a_0(y_t)dt + b(y_t)dW_t^y,$$

with $b(y) \neq 0$ for all y.

1. HW with stochastic a:

$$\sigma(r, y, x) = \sigma e^{-yx} \tag{12}$$

2. HW with stochastic σ:

$$\sigma(r, y, x) = ye^{-ax} \tag{13}$$

3. CIR with stochastic σ:

$$\sigma(r, y, x) = y\sqrt{r(0)} \cdot \lambda(x, y, a) \tag{14}$$

4. CIR with stochastic a:

$$\sigma(r, y, x) = \sigma\sqrt{r(0)} \cdot \lambda(x, \sigma, y) \tag{15}$$

For all these models, the induced parameterized model admits, by construction, an FDR. It is now reasonable to ask if this also holds for the corresponding stochastic volatility models.

3 Finite realizations for general stochastic volatility models

In order to solve the FDR problem for stochastic volatility models we will need the Lie algebra theory for the existence of FDRs in Hilbert space, developed in [7] and extended in [17]. The main result that we will use is Theorem 3.2 of [7] (or the corresponding result in [17]). This result basically says that, for an SDE of HJM type on a Hilbert space, there exists a generic FDR if and only if the Lie algebra generated by the Stratonovich drift and diffusion terms is locally of finite dimension. In Appendix A we provide a brief recapitulation of that theory and the reader is referred to [7] and [17] for proofs and details.

3.1 Lie algebra conditions for the existence of an FDR

Our problem is to study the existence of an FDR for a stochastic volatility model of the form

$$\begin{aligned} dr_t &= \mu_0(r_t, y_t)dt + \sigma(r_t, y_t)dW_t && (16) \\ dy_t &= a^0(y_t)dt + b(y_t)dW_t. && (17) \end{aligned}$$

In the particular case of a forward rate model, the drift term is given by

$$\mu_0(r, y, x) = \frac{\partial}{\partial x} r(x) + \mathbf{H}\sigma(r, y, x) \tag{18}$$

but none of the results in this section does in fact depend upon this particular structure of μ_0. Therefore we will, for the rest of the section, consider a general abstract stochastic volatility model of the form (16)-(17). Within the framework of [7] the drift μ_0 has to be assumed to be smooth. If instead, we use the framework of [17], then μ_0 is allowed to be of the form

$$\mu_0(r, y) = \mathbf{F}r + \alpha(r, y).$$

Here $\mathbf{F}$ is assumed to be a linear densely defined operator, generating a strongly continuous semigroup on $\mathcal{H}$, whereas α is assumed to be smooth.

To apply the Lie algebra results of [7] and [17] to the present situation we proceed in the following way.

- Define the Hilbert space $\hat{\mathcal{H}}$ by $\hat{\mathcal{H}} = \mathcal{H} \times R^k$.
- Define the $\hat{\mathcal{H}}$-valued process $\hat{r}$ by

$$\hat{r}_t = \begin{bmatrix} r_t \\ y_t \end{bmatrix} \tag{19}$$

- Write the dynamics of $\hat{r}$ on Stratonovich form instead of the original Itô form.
- Use the abstract Lie algebra results from [7] (see Appendix A) and [17] on the process $\hat{r}$.

We will thus view $\hat{r}$ as an infinite dimensional "column vector" process, and we will henceforth always write it on block vector form as above.

The Stratonovich dynamics of $\hat{r}$ are routinely derived as

$$\begin{aligned} dr_t &= \mu(r_t, y_t)dt + \sigma(r_t, y_t) \circ dW_t && (20) \\ dy_t &= a(y_t)dt + b(y_t) \circ dW_t, && (21) \end{aligned}$$

where

$$\begin{aligned} \mu(r, y) &= \mu_0(r, y) - \frac{1}{2}\sigma_r(r, y)\sigma(r, y) - \frac{1}{2}\sigma_y(r, y)b(y) && (22) \\ a(y) &= a^0(y) - \frac{1}{2}b_y(y)b(y). && (23) \end{aligned}$$

Here σ_r denotes the partial Frechet derivative of σ w.r.t. the vector variable r and similarly for the other terms.

Written as a single equation on $\hat{\mathcal{H}}$ we thus have

$$d\hat{r}_t = \hat{\mu}(\hat{r})dt + \hat{\sigma}(\hat{r}) \circ dW_t, \tag{24}$$

where $\hat{\mu}$ and $\hat{\sigma}$ are given by

$$\hat{\mu}(r,y) = \begin{bmatrix} \mu(r,y) \\ a(y) \end{bmatrix} \tag{25}$$

$$\hat{\sigma}(r,y) = \begin{bmatrix} \hat{\sigma}_1(r,y), & \dots, & \hat{\sigma}_m(r,y) \end{bmatrix} \tag{26}$$

Here the vector fields $\hat{\sigma}_1, \dots, \hat{\sigma}_m$ are defined by

$$\hat{\sigma}_i(r,y) = \begin{bmatrix} \sigma_i(r,y) \\ b_i(y) \end{bmatrix} \tag{27}$$

where b_i is the $i:th$ column of the b matrix, i.e.

$$b_i(y) = \begin{bmatrix} b_{1i}(y) \\ \vdots \\ b_{ki}(y) \end{bmatrix} \tag{28}$$

We make the following standing regularity assumption which is assumed to hold throughout the entire paper.

Assumption 3.1 *We assume that the dimension of the Lie algebra*

$$\{\hat{\mu}, \hat{\sigma}_1, \dots, \hat{\sigma}_m\}_{LA} < \infty, \tag{29}$$

is constant in a neighborhood of $\hat{r}_0 \in \hat{\mathcal{H}}$

Our first general result now follows immediately from the Lie algebra results of [7] and [17].

Theorem 3.1 *Under Assumption 3.1, the stochastic volatility model (16)-(17) will have a generic FDR at the point* $\hat{r}_0$ *if and only if*

$$dim\, \{\hat{\mu}, \hat{\sigma}_1, \dots, \hat{\sigma}_m\}_{LA} < \infty, \tag{30}$$

in a neighbourhood of $\hat{r}_0 \in \hat{\mathcal{H}}$.

For simplicity of notation we will often use the shorthand notation $\{\hat{\mu}, \hat{\sigma}\}_{LA}$ for the Lie algebra $\{\hat{\mu}, \hat{\sigma}_1, \dots, \hat{\sigma}_m\}_{LA}$.

3.2 Geometric intuition

At this level of generality it is hard to obtain more concrete results. As an example: there seems to be no simple result connecting the existence of an FDR for the full model with existence of an FDR for the parameterized model. The geometric intuition behind this is roughly as follows.

- From [7] we know that existence of an FDR for $\hat{r}$ is equivalent to the existence of a finite dimensional invariant manifold in $\hat{\mathcal{H}}$ passing through $\hat{r}_0$. See Appendix A.
- If the parameterized model admits a generic FDR then, for every fixed y near y^0, there exists an invariant manifold $\mathcal{G}$ in $\mathcal{H}$ through r_0. Thus one would perhaps guess that the manifold $\mathcal{G} \times R^k$ would be invariant for $\hat{r}$, thus implying the existence of an FDR for $\hat{r}$.
- However, the manifold $\mathcal{G}$ above will generically depend on y. Writing it as $\mathcal{G}^y$, what may (and generically will) happen is that, as $\hat{r}_t$ moves around in $\hat{\mathcal{H}}$, y_t will move in R^k and the family $\{\mathcal{G}^{y_t};\ t \geq 0\}$ may sweep out an infinite dimensional manifold in $\hat{\mathcal{H}}$. Thus the existence of an FDR for the parameterized model is not sufficient for the existence of an FDR for the full model.
- Conversely, the existence of an FDR for the parameterized model does not even seem to be necessary for the existence of an FDR for the full model. Suppose for example that, for each y, there does not exist an invariant manifold for the parameterized model. This means that the parameterized model does not possess an FDR. Despite this it could well happen that the process $\hat{r}$ **does** live on a finite dimensional invariant manifold (and thus possesses an FDR). The reason for this is that there could be a subtle interplay between the dynamics of r and y, and in particular one might intuitively expect this interplay to be possible if there is strong correlation between the Wiener process components driving r and y.
- From the argument above we are led to guess that the simplest structural situation occurs when r and y are driven by independent Wiener processes. Since in this case, the evolution of y is independent of the present state of r, we may even guess (bravely) that any FDR properties of the full model will be "uniform" w.r.t. y in the sense that the results will not depend much on the particular dynamics of y.

As we shall see below, the intuition outlined above is basically substantiated.

4 General orthogonal noise models

Based on the informal arguments in the previous section we now go on to study the case when r and y are driven by independent Wiener processes. We will refer to this type of model as an "orthogonal noise model". As in the previous section, we will not use the particular structure of the drift term that is induced by the HJM drift condition for a forward rate model. Thus, in the present section we consider a general drift term.

4.1 Model specification and preliminary results

Assumption 4.1 *For the rest of the section we assume that we can write the Wiener process W on block vector form as*

$$W_t = \begin{bmatrix} W_t^r \\ \\ W_t^y \end{bmatrix}$$

where W^r and W^y are vector Wiener processes of dimensions m_r and m_y respectively. Furthermore we assume that the (r, y) dynamics are of the particular form

$$\begin{aligned} dr_t &= \mu_0(r_t, y_t)dt + \sigma(r_t, y_t)dW_t^r \quad (31) \\ dy_t &= a^0(y_t)dt + b(y_t)dW_t^y, \quad (32) \end{aligned}$$

where the coefficients satisfy suitable smoothness conditions (see Section 3).

Under this assumption r and y are driven by orthogonal noise terms, and this leads to an important simplification of the geometric structure of the model.

Lemma 4.1 *The Stratonovich formulation of (31)-(32) is given by*

$$\begin{aligned} dr_t &= \mu(r_t, y_t)dt + \sigma(r_t, y_t) \circ dW_t^r \quad (33) \\ dy_t &= a(y_t)dt + b(y_t) \circ dW_t^y, \quad (34) \end{aligned}$$

where

$$\mu(r, y) = \mu_0(r, y) - \frac{1}{2}\sigma_r(r, y)\sigma(r, y) \quad (35)$$

$$a(y) = a^0(y) - \frac{1}{2}b_y(y)b(y). \quad (36)$$

Proof. In order to find the Stratonovich form of the r dynamics we need to compute

$$d\langle \sigma, W^r \rangle_t = d\sigma(r_t, y_t).$$

The infinite dimensional Itô formula gives us

$$d\sigma(r_t, y_t) = (dt\text{-terms}) + \sigma_r(r_t, y_t)\sigma(r_t, y_t)dW_t^r + \sigma_y(r_t, y_t)b(y_t)dW_t^y$$

We thus obtain

$$d\langle \sigma, W^r \rangle_t = \sigma_r(r_t, y_t)\sigma(r_t, y_t)d\langle W^r, W^r \rangle_t + \sigma_y(r_t, y_t)b(y_t)d\langle W^y, W^r \rangle_t$$

Since W^r and W^y are independent this simplifies to

$$d\langle \sigma, W^r \rangle_t = \sigma_r(r,y)\sigma(r,y)dt. \quad \blacksquare$$

In order to see more clearly the geometric structure of the orthogonal noise model we write it on block operator form as

$$d\begin{bmatrix} r_t \\ y_t \end{bmatrix} = \begin{bmatrix} \mu(r_t, y_t) \\ a(y_t) \end{bmatrix} dt + \begin{bmatrix} \sigma(r_t, y_t) \\ 0 \end{bmatrix} \circ dW_t^r + \begin{bmatrix} 0 \\ b(y_t) \end{bmatrix} \circ dW_t^y \quad (37)$$

We thus have the following immediate and preliminary result.

Proposition 4.1 *The orthogonal noise model (31)-(32) admits an FDR if and only if the Lie algebra generated by the vector fields*

$$\begin{bmatrix} \mu(r,y) \\ a(y) \end{bmatrix}, \begin{bmatrix} \sigma_1(r,y) \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} \sigma_{m_r}(r,y) \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ b_1(y) \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ b_{m_y}(y) \end{bmatrix}$$

is finite dimensional at $\hat{r}_0$.

More compactly we will often write the generators of the Lie algebra above as

$$\hat{\mu}(r,y), \quad \hat{\sigma}(r,y), \quad \hat{b}(y)$$

where, in obvious shorthand notation,

$$\hat{\mu}(r,y) = \begin{bmatrix} \mu(r,y) \\ a(y) \end{bmatrix}, \quad \hat{\sigma}(r,y) = \begin{bmatrix} \sigma(r,y) \\ 0 \end{bmatrix}, \quad \hat{b}(y) = \begin{bmatrix} 0 \\ b(y) \end{bmatrix} \quad (38)$$

A very useful property of the orthogonal noise model is the simple structure of the Stratonovich formulation of the parameterized model. The proof is trivial.

Lemma 4.2 *For the orthogonal noise model (31)-(32), the Itô formulation of the parameterized model is defined by*

$$dr_t = \mu_0(r_t, y)dt + \sigma(r_t, y)dW_t^r, \quad (39)$$

and the Stratonovich formulation of the parameterized model is given by

$$dr_t = \mu(r_t, y)dt + \sigma(r_t, y) \circ dW_t^r, \quad (40)$$

with μ *defined by (35).*

The point of this Lemma is that it shows that, for orthogonal noise models, the operations "restrict to the parameterized model" and "compute the Stratonovich dynamics" commute, i.e. the Stratonovich formulation of the parameterized model is identical to the parameterized version of the Stratonovich formulation of the original model.

In order to obtain easily verifiable necessary and sufficient conditions for the existence of an FDR we will in the next sections introduce some further structural assumptions. In doing this we will have to deal with Lie brackets in several spaces, so we have to clarify some notation.

Definition 4.1 *From now on, the following notation is in force:*

- *For any vector smooth fields $\hat{f}(r,y)$ and $\hat{g}(r,y)$ on $\hat{\mathcal{H}}$, the expression $\left[\hat{f},\hat{g}\right]$ denotes the Lie bracket in $\hat{\mathcal{H}}$.*
- *For any smooth mapping $f(r,y)$ where $f:\hat{\mathcal{H}}\rightarrow\mathcal{H}$ and for any fixed $y\in R$, the* **parameterized vector field** *$f^y:\mathcal{H}\rightarrow\mathcal{H}$ is defined by $f^y(r)=f(r,y)$*
- *For any smooth mappings $f,g:\hat{\mathcal{H}}\rightarrow\mathcal{H}$, the expression $[f^y,g^y]$ denotes the Lie bracket on $\mathcal{H}$ between f^y and g^y. This Lie bracket will sometimes also be denoted by $[f(\cdot,y),g(\cdot,y)]_{\mathcal{H}}$.*
- *For vector fields $c(y)$ and $d(y)$ on R^k, the notation $[c,d]$ denotes the Lie bracket on R^k.*

4.2 Necessary conditions

It turns out that, in order to obtain easy necessary condition, a crucial role is played by the geometric relation between the drift vector field $a(y)$ and the Lie algebra on R^k generated by the diffusion vector fields $b_1(y),\ldots,b_{m_y}$.

Our first result relates the stochastic volatility model to the corresponding parameterized model.

Proposition 4.2 *Consider the model (31)-(32). Assume that*

$$a\in\left\{b_1,\ldots,b_{m_y}\right\}_{LA} \tag{41}$$

in a neighbourhood of y^0. Under this assumption, a necessary condition for the existence of an FDR for the stochastic volatility model is that the corresponding parameterized model

$$dr_t=\mu(r_t,y)dt+\sigma(r_t,y)dW_t^r \tag{42}$$

admits a generic FDR at y_0.

Proof. We assume that the full stochastic volatility model admits an FDR, and we also assume that (41) is satisfied. We now have to show that, under these assumptions, the parameterized model admits and FDR, i.e. that the Lie algebra (on $\mathcal{H}$) of the parameterized model is finite dimensional near r_0, for every fixed y near y_0. From Lemma 4.2 we know that the Stratonovich formulation of the parameterized model is given by

$$dr_t = \mu(r_t, y)dt + \sigma(r_t, y) \circ dW_t^r, \tag{43}$$

which we write as

$$dr_t = \mu^y(r_t)dt + \sigma^y(r_t) \circ dW_t^r. \tag{44}$$

Our task is now to show that

$$\{\mu^y, \sigma^y\}_{LA}$$

is finite dimensional near r_0 for all y near y_0.

Since we assumed that the full model possessed an FDR we know that the Lie algebra

$$\left\{\hat{\mu}, \hat{\sigma}, \hat{b}\right\}_{LA} = \left\{ \begin{bmatrix} \mu(r,y) \\ a(y) \end{bmatrix}, \begin{bmatrix} \sigma(r,y) \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ b(y) \end{bmatrix} \right\}_{LA}$$

is finite dimensional near $\hat{r}_0$. We now have the trivial inclusion

$$\left\{\hat{b}\right\}_{LA} \subseteq \left\{\hat{\mu}, \hat{\sigma}, \hat{b}\right\}_{LA},$$

and we go on to compute $\left\{\hat{b}\right\}_{LA} = \left\{\hat{b}_1, \ldots, \hat{b}_{m_y}\right\}_{LA}$. For any i and j, let us thus compute the Lie bracket $\left[\hat{b}_i, \hat{b}_j\right]$. Since

$$\hat{b}_i(y) = \begin{bmatrix} 0 \\ b_i(y) \end{bmatrix}$$

we easily obtain the block matrix form for the Frechet derivative of $\hat{b}_i$ on $\hat{\mathcal{H}}$ as

$$\hat{b}_i'(y) = \begin{bmatrix} 0 & 0 \\ 0 & b_i'(y) \end{bmatrix}$$

where b_i' denotes the Frechet derivative on R^k of the vector field b_i. Performing the same calculation for $\hat{b}_j$ we obtain

$$\left[\hat{b}_i, \hat{b}_j\right]_{\hat{\mathcal{H}}} = \begin{bmatrix} 0 & 0 \\ 0 & b_i' \end{bmatrix} \begin{bmatrix} 0 \\ b_j \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & b_j' \end{bmatrix} \begin{bmatrix} 0 \\ b_i \end{bmatrix} = \begin{bmatrix} 0 \\ [b_i, b_j]_{R^k} \end{bmatrix}.$$

Continuing in this way by taking repeated brackets, wee see that if $\hat{\beta}$ denotes a generic element of $\left\{\hat{b}\right\}_{LA}$ then it has the form

$$\hat{\beta} = \begin{bmatrix} 0 \\ \beta \end{bmatrix}$$

where β denotes a generic element of $\{b\}_{LA}$. We can formally write this as

$$\left\{\hat{b}\right\}_{LA} = \left\{\hat{b}_1, \ldots, \hat{b}_{m_y}\right\}_{LA} = \begin{bmatrix} 0 \\ \{b_1, \ldots, b_{m_y}\}_{LA} \end{bmatrix} = \begin{bmatrix} 0 \\ \{b\}_{LA} \end{bmatrix}$$

We assumed that $a \in \{b\}_{LA}$, so there exists vector fields $c_1(y), \ldots, c_n(y)$ in $\{b\}_{LA}$ and scalar fields $\alpha_1(y), \ldots, \alpha_n(y)$ on R^k such that

$$a(y) = \sum_1^n \alpha_i(y) c_i(y)$$

for all y near y_0. Since $\left\{\hat{b}\right\}_{LA} \subseteq \left\{\hat{\mu}, \hat{\sigma}, \hat{b}\right\}_{LA}$ we see from the above that the vector fields $\hat{c}_1, \ldots, \hat{c}_n$ where

$$\hat{c}_i = \begin{bmatrix} 0 \\ c_i(y) \end{bmatrix}$$

all lie in $\left\{\hat{\mu}, \hat{\sigma}, \hat{b}\right\}_{LA}$. We may now invoke Lemma A.1 to perform Gaussian elimination. More precisely, we may replace $\hat{\mu}$ by $\hat{\mu} - \sum_1^n \alpha_i \hat{c}_i$, and we obtain

$$\hat{\mu} - \sum_1^n \alpha_i \hat{c}_i = \begin{bmatrix} \mu \\ a \end{bmatrix} - \sum_1^n \alpha_i \begin{bmatrix} 0 \\ c_i \end{bmatrix} = \begin{bmatrix} \mu \\ 0 \end{bmatrix}.$$

From this we see that the Lie algebra $\left\{\hat{\mu}, \hat{\sigma}, \hat{b}\right\}_{LA}$ for the full model is in fact generated by the much simpler system $\hat{m}$, $\hat{\sigma}$ and $\hat{b}$ where $\hat{m}$ is defined by

$$\hat{m} = \begin{bmatrix} \mu \\ 0 \end{bmatrix}.$$

Since we assumed that $\left\{\hat{\mu}, \hat{\sigma}, \hat{b}\right\}_{LA}$ was finite dimensional, then also the smaller Lie algebra

$$\{\hat{m}, \hat{\sigma}\}_{LA} = \left\{ \begin{bmatrix} \mu \\ 0 \end{bmatrix}, \quad \begin{bmatrix} \sigma \\ 0 \end{bmatrix} \right\}_{LA}$$

is necessarily also finite dimensional. In computing this latter Lie algebra we may now argue as for $\{b\}_{LA}$ above. Let us, for example, compute the Lie bracket $[\hat{m}, \hat{\sigma}_i]$. The Frechet derivatives (in $\hat{\mathcal{H}}$) are given by

$$\hat{m}' = \begin{bmatrix} \mu_r & \mu_y \\ 0 & 0 \end{bmatrix}, \quad , \hat{\sigma}_i' = \begin{bmatrix} \sigma_{ir} & \sigma_{iy} \\ 0 & 0 \end{bmatrix}$$

where subindex r and y denotes the partial Frechet derivative w.r.t r and y respectively. We thus obtain

$$[\hat{m}, \hat{\sigma}_i] = \begin{bmatrix} \mu_r & \mu_y \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \sigma_i \\ 0 \end{bmatrix} - \begin{bmatrix} \sigma_{ir} & \sigma_{iy} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mu \\ 0 \end{bmatrix} = \begin{bmatrix} \mu_r \sigma_i - \sigma_{ir} \mu \\ 0 \end{bmatrix}$$

Now we observe that $\mu_r(r,y)\sigma_i(r,y) - \sigma_{ir}(r,y)\mu(r,y) = [\mu^y, \sigma_i^y](r)$ so we have

$$[\hat{m}, \hat{\sigma}_i](r,y) = \begin{bmatrix} [\mu^y, \sigma_i^y] \\ 0 \end{bmatrix}(r),$$

and continuing in this way we obtain

$$\{\hat{m}, \hat{\sigma}\}_{LA}(r,y) = \left\{ \begin{bmatrix} \mu \\ 0 \end{bmatrix}, \quad \begin{bmatrix} \sigma \\ 0 \end{bmatrix} \right\}_{LA}(r,y) = \begin{bmatrix} \{\mu^y, \sigma^y\}_{LA} \\ 0 \end{bmatrix}(r)$$

Since $\{\hat{m}, \hat{\sigma}\}_{LA}$ is finite dimensional for all (r,y) near (r_0, y_0) we thus see that $\{\mu^y, \sigma^y\}_{LA}$ has to be finite dimensional near r_0 for all y near y_0. This however is equivalent to the existence of an FDR for the parameterized model. ∎

We have the following obvious corollary, which seems to be enough for many concrete applications.

Corollary 4.1 *Assume that the Lie algebra generated by b in R^k is full, i.e. that*

$$\{b_1, \ldots, b_{m_y}\}_{LA} = R^k. \tag{45}$$

Then, regardless of the form of a, the existence of an FDR for the parameterized model is necessary for the existence of an FDR for the full model. In particular, the assumption above is valid, and thus the conclusion holds, for the following special cases.

- *$m_y = k$ and the $k \times k$ diffusion matrix $b(y)$ is invertible near y_0.*
- *y is scalar and driven by a scalar Wiener process (i.e. $k = m_y = 1$), and the scalar field $b(y)$ is nonzero near y_0.*

We now go on to obtain more precise (but still easily verifiable) necessary conditions, and the simplest case is when the diffusion matrix b is square and invertible. Since the multidimensional case is a bit messy we start with the scalar case, and we will in fact use the scalar result in the proof of the multidimensional case.

Proposition 4.3 *Assume that y and W^y are scalar and that the (scalar) diffusion term $b(y)$ is nonzero near y_0. Then the following conditions are necessary for the existence of an FDR for the full model.*

- *For every fixed r and y near (r_0, y_0) the partial derivatives of μ and $\sigma_i(r,y)$ $i = 1, \ldots, m_r$ w.r.t y span a finite dimensional space in $\mathcal{H}$. Formally, for every (r,y)*

$$\dim\ span \left\{ \frac{\partial^n \mu}{\partial y^n}(r,y); \quad n = 0, 1, 2, \ldots \right\} < \infty \tag{46}$$

  and

$$\dim\ span \left\{ \frac{\partial^n \sigma_i}{\partial y^n}(r,y); \quad n = 0, 1, 2, \ldots \right\} < \infty \tag{47}$$

  for every $i = 1, \ldots m_r$.

- *The drift term μ, and each volatility component σ_i have the form*

$$\mu(r, y, x) = \sum_{j=1}^{n_0} c_{0j}(r, y)\lambda_{0j}(r, x). \tag{48}$$

and

$$\sigma_i(r, y, x) = \sum_{j=1}^{n_i} c_{ij}(r, y)\lambda_{ij}(r, x). \tag{49}$$

Proof. In order to obtain necessary conditions we assume that the full model admits an FDR, and for simplicity of notation we assume that $m_r = 1$ (this will not affect the proof). The Lie algebra for the full model is then finite dimensional and it is generated by

$$\hat{\mu} = \begin{bmatrix} \mu(r, y) \\ a(y) \end{bmatrix}, \quad \hat{\sigma} = \begin{bmatrix} \sigma \\ 0 \end{bmatrix}, \quad \hat{b} = \begin{bmatrix} 0 \\ b(y) \end{bmatrix}.$$

Since b is scalar and nonzero we can use Gaussian elimination (Lemma A.1) and locally replace $\hat{b}$ by

$$\frac{1}{b(y)}\hat{b}(y) = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

and, with further elimination, we see that the full Lie algebra is in fact generated by

$$\hat{\mu} = \begin{bmatrix} \mu(r, y) \\ 0 \end{bmatrix}, \quad \hat{\sigma} = \begin{bmatrix} \sigma \\ 0 \end{bmatrix}, \quad \hat{1} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

We start by proving (47), the proof for (46) being identical. Since the full algebra is finite dimensional, also the smaller Lie algebra generated by $\hat{\sigma}$ and $\hat{1}$ has to be finite dimensional. In particular the space spanned in $\hat{\mathcal{H}}$ by the vector fields

$$\hat{\sigma}, \quad \left[\hat{\sigma}, \hat{1}\right], \quad \left[\left[\hat{\sigma}, \hat{1}\right], \hat{1}\right], \quad \left[\left[\left[\hat{\sigma}, \hat{1}\right], \hat{1}\right], \hat{1}\right], \ldots$$

obtained by starting with $\hat{\sigma}$ and then taking repeated brackets with $\hat{1}$, has to be finite dimensional at every point (r, y) near $\hat{r}_0$. We can write these vectors more compactly as

$$Ad^0_{\hat{1}}(\sigma), \quad Ad^1_{\hat{1}}(\sigma), \quad Ad^2_{\hat{1}}(\sigma), \ldots$$

where for any vector field $\hat{f}$ the operators $Ad^n_{\hat{f}} : \hat{\mathcal{H}} \to \hat{\mathcal{H}}$ are defined recursively by

$$\begin{aligned} Ad^0_{\hat{f}}(\hat{g}) &= \hat{g}, \\ Ad^1_{\hat{f}}(\hat{g}) &= \left[\hat{g}, \hat{f}\right], \\ Ad^{n+1}_{\hat{f}}(\hat{g}) &= \left[Ad^n_{\hat{f}}(\hat{g}), \hat{f}\right]. \end{aligned}$$

We easily obtain the Frechet derivatives of $\hat{\sigma}$ and $\hat{1}$ as

$$\hat{\sigma}' = \begin{bmatrix} \partial_r \sigma & \partial_y \sigma \\ 0 & 0 \end{bmatrix}, \quad \hat{1}' = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

where ∂_r and ∂_r denotes the corresponding partial Frechet derivatives. Thus we have

$$Ad^1_{\hat{1}}(\hat{\sigma}) = [\hat{\sigma}, \hat{1}] = \begin{bmatrix} \partial_r \sigma & \partial_y \sigma \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \sigma \\ 0 \end{bmatrix} = \begin{bmatrix} \partial_y \sigma \\ 0 \end{bmatrix}$$

Similarly we have

$$\left\{ Ad^1_{\hat{1}}(\hat{\sigma}) \right\}' = \begin{bmatrix} \partial_r \partial_y \sigma & \partial_y^2 \sigma \\ \\ 0 & 0 \end{bmatrix}$$

and thus

$$Ad^2_{\hat{1}}(\hat{\sigma}) = \left[Ad_{\hat{1}^1}(\hat{\sigma}), \hat{1} \right] = \begin{bmatrix} \partial_r \partial_y \sigma & \partial_y^2 \sigma \\ \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \partial_y \sigma \\ 0 \end{bmatrix} = \begin{bmatrix} \partial_y^2 \sigma \\ 0 \end{bmatrix}$$

Continuing this way we see by induction that

$$Ad^n_{\hat{1}}(\hat{\sigma}) = \begin{bmatrix} \partial_y^n \sigma \\ 0 \end{bmatrix}.$$

Since, by the argument above, $\left\{ Ad^n_{\hat{1}}(\hat{\sigma})(r, y); \quad n \geq 0 \right\}$ span a finite dimensional subspace of $\hat{\mathcal{H}}$ for all (r, y) near $\hat{r}_0$, we thus see that

$$\left\{ \partial_y^n \sigma(r, y), \quad n \geq 0 \right\}$$

must span a finite dimensional subspace in $\mathcal{H}$ for all (r, y) near $\hat{r}_0$. We have thus proved (47) for the case when W^r is scalar. The general case is proved by applying the above argument for each component of σ.

We now go on to prove the necessary condition (49) and we will in fact show that (49) follows from (47). Again we carry out a separate argument for each component σ_i, so without loss of generality we may assume that σ only has a single component (i.e that $m_r = 1$). Now, if (47) holds and we denote the dimension of the spanned subspace by $n+1$, there exists scalar fields $a_j(r, y); \quad j = 0, \dots n$, such that we have the following $\mathcal{H}$-valued vector identity holding locally at $\hat{r}_0$

$$\partial_y^{n+1} \sigma(r, y) = \sum_{j=0}^{n} a_j(r, y) \partial_y^j \sigma(r, y) \tag{50}$$

We now fix an arbitrary r, and for this fixed r we define the $\mathcal{H}$-vector functions $Z_0(y), Z_1(y), \dots Z_n(y)$ by

$$Z_0(y) \quad = \quad \sigma(r, y),$$

$$\begin{array}{rcl} Z_1(y) & = & \partial_y \sigma(r,y), \\ \vdots & & \vdots \\ Z_n(y) & = & \partial_y^n \sigma(r,y), \end{array}$$

and the $\mathcal{H}^{n+1}$-valued block vector function $Z(y)$ by

$$Z(y) = \begin{bmatrix} Z_0(y) \\ Z_1(y) \\ \vdots \\ Z_n(y) \end{bmatrix}$$

The point of this is that we can now write equation (50) as the linear ODE

$$\frac{d}{dy}\begin{bmatrix} Z_0(y) \\ Z_1(y) \\ \vdots \\ Z_n(y) \end{bmatrix} = \begin{bmatrix} 0 & I & 0 & \dots & 0 \\ 0 & 0 & I & \dots & 0 \\ \vdots & \vdots & & & I \\ a_0(y)I & a_1(y)I & a_2(y)I & \dots & a_n(y)I \end{bmatrix}\begin{bmatrix} Z_0(y) \\ Z_1(y) \\ \vdots \\ Z_n(y) \end{bmatrix}$$

where I denotes the identity on $\mathcal{H}$. More compactly we can thus write it as

$$\frac{dZ(y)}{dy} = (A(y) \otimes I)\, Z(y) \tag{51}$$

where $\otimes$ denotes the Kronecker product, and the $(n+1)\times(n+1)$ matrix function A is defined as the companion matrix

$$A(y) = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & & & 1 \\ a_0(y) & a_1(y) & a_2(y) & \dots & a_n(y) \end{bmatrix}.$$

As one would perhaps guess, the solution of (51) can be shown (see Lemma 4.3 below) to have the representation

$$Z(y) = [\Phi(y,y_0) \otimes I]\, Z(y_0), \tag{52}$$

where Φ is the transition matrix induced by A. In particular we thus obtain

$$Z_0(y) = \sum_{j=0}^{n} c_j(y) Z_j(0)$$

where $c_j(y) = \Phi(y,y_0)_{1,j}$. Recalling that there is a suppressed r and that $Z_j(y) = \partial_y^j \sigma(r,0)$ we obtain

$$\sigma(r,y) = \sum_{j=0}^{n} c_j(r,y) \partial_y^j \sigma(r,0), \tag{53}$$

which proves (49). The proof for (48) is identical. ∎

Lemma 4.3 *The solution of the linear ODE (51) has the representation*

$$Z(y) = [\Phi(y, y_0) \otimes I]\, Z(y_0), \tag{54}$$

Here the $(n+1) \times (n+1)$ *matrix function* Φ *is the transition matrix for the ODE*

$$\frac{dz(y)}{dy} = A(y)z(y),$$

i.e. $\Phi(t, s)$ *is the unique solution of the linear matrix ODE*

$$\begin{aligned} \frac{\partial \Phi(t,s)}{\partial t} &= A(t)\Phi(t,s), \\ \Phi(s,s) &= I_{n+1}, \quad \forall s, \end{aligned}$$

where I_{n+1} *is the identity matrix on* R^{n+1}.

Proof. Let us define Z^0 by $Z^0(y) = [\Phi(y, y_0) \otimes I]\, Z(y_0)$. Using the formula $CD \otimes EF = (C \otimes E)(D \otimes F)$ we obtain

$$\begin{aligned} \frac{dZ^0(y)}{dy} &= \left(\frac{d}{dy}\Phi(y, y_0) \otimes I \right) Z(y_0) \\ &= \{A(y)\Phi(y, y_0) \otimes I\}\, Z(y_0) \\ &= \{A(y) \otimes I\}\, \{\Phi(y, y_0) \otimes I\}\, Z(y_0) \\ &= \{A(y) \otimes I\}\, Z^0(y). \end{aligned}$$

Thus Z^0, defined by (54), satisfies (51) and, since the initial value is the correct one, we have by uniqueness $Z = Z^0$ thus finishing the proof of the Lemma. ∎

In order to state the corresponding multidimensional result we need to introduce some notation.

Definition 4.2 *A* **multi index** $\alpha \in Z_+^k$ *is any* k*-vector with nonnegative integer elements. For a multi index* $\alpha = (\alpha_1, \ldots, \alpha_k)$ *the differential operator* ∂_y^α *is defined by*

$$\partial_y^\alpha = \frac{\partial^{\alpha_1}}{\partial y_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial y_2^{\alpha_2}} \cdots \frac{\partial^{\alpha_k}}{\partial y_k^{\alpha_k}}$$

We can now state and prove a multidimensional version of the theorem above. The crucial assumption needed is that the Lie algebra generated by the diffusion matrix $b(y)$ spans the entire space R^k.

Proposition 4.4 *Assume that the condition*

$$\left\{b_1, \ldots b_{m_y}\right\}_{LA} = R^k, \tag{55}$$

is satisfied near y_0.

Then the following conditions are necessary for the existence of an FDR for the stochastic volatility model.

- *For every fixed r and y near (r_0, y_0) the partial derivatives of $\mu(r, y)$ and $\sigma_i(r, y)$ w.r.t y span a finite dimensional space in $\mathcal{H}$. Formally, for every (r, y)*

$$dim\ span \left\{ \partial_y^\alpha \mu(r, y); \quad \alpha \in Z_+^k \right\} < \infty \tag{56}$$

and

$$dim\ span \left\{ \partial_y^\alpha \sigma_i(r, y); \quad \alpha \in Z_+^k \right\} < \infty \tag{57}$$

for every $i = 1, \ldots m_r$.

- *The drift μ and every volatility component σ_i have the form*

$$\mu(r, y, x) = \sum_{j=1}^{n_i} c_{ij}(r, y) \lambda_{ij}(r, x). \tag{58}$$

$$\sigma_i(r, y, x) = \sum_{j=1}^{n_i} c_{ij}(r, y) \lambda_{ij}(r, x). \tag{59}$$

Proof. We confine ourselves to proving (57) and (59), the proof of (56) and (58) being identical. We assume that the full model admits and FDR i.e. that the full Lie algebra $\left\{ \hat{\mu}, \hat{\sigma}, \hat{b} \right\}_{LA}$ is finite dimensional. From the spanning assumption (55) it follows easily that, after Gaussian elimination, this Lie algebra is in fact generated by the vector fields

$$\hat{f}_0 = \begin{bmatrix} \mu(r, y) \\ 0 \end{bmatrix}, \quad \hat{\sigma} = \begin{bmatrix} \sigma \\ 0 \end{bmatrix}, \quad \hat{I} = \begin{bmatrix} 0 \\ I_k \end{bmatrix},$$

where I_k denotes the identity matrix in R^k. The smaller Lie algebra obtained by selecting one fixed component of $\hat{\sigma}$ (say the $i : th$ component) and then taking successive Lie brackets with different columns of $\hat{I}$ is included in $\left\{ \hat{\mu}, \hat{\sigma}, \hat{b} \right\}_{LA}$ and thus also finite dimensional. As in the proof of Proposition 4.3 it is however easy to see that these repeated brackets will be of the form

$$\begin{bmatrix} \partial_y^\alpha \sigma_i(r, y) \\ 0 \end{bmatrix},$$

which proves (57).

We now go on to show that (59) follows from (57). We thus assume that (57) holds and we will in fact prove that for each component σ_i and each natural number $n \le k$ we have a representation of the form

$$\sigma_i(r, y) = \sum_{\alpha \in Z_+^n} c_\alpha(r, y) \partial_y^\alpha \sigma_i(r, 0_n, y^{n+1}) \tag{60}$$

where the sum only contains a finite number of terms. In the expression above, the differential operator ∂_y^α for $\alpha \in R^n$ will only contain partial derivatives w.r.t the n first variables $y_1, \dots y_n$ and c_α is some scalar field. The expression 0_n denotes the zero vector $(0, \dots, 0)$ in R^n, and for any $y \in R^k$ y^n denotes the vector $(y_n, \dots, y_k)$.

We prove (60) by induction and for notational simplicity we suppress the subindex i in σ_i. The case $n = 1$ is easily proved in exactly the same way as when we proved (53). For the induction step, let us assume that (60) holds for a fixed n. From the assumption (57) it follows in particular that the space spanned of the vector fields

$$\left\{\partial^j_{y_{n+1}} \sigma(r, y); \quad j = 0, 1, \dots\right\}$$

is finite dimensional near (r_0, y_0). Again adapting the proof of (53) (keep r and all y components except y_{n+1} fixed) to the present situation, we obtain a representation of σ as a finite sum of the form

$$\sigma(r, y) = \sum_{j=1}^{N} \gamma_j(r, y) \partial^j_{y_{n+1}} \sigma(r, y_1, \dots, y_n, 0, y^{n+2}), \tag{61}$$

where $\gamma_j(r, y)$ is a scalar field. We can now apply $\partial^j_{y_{n+1}}$ to (60) and set $y_{n+1} = 0$ to obtain

$$\partial^j_{y_{n+1}} \sigma(r, y_1, \dots, y_n, 0, y^{n+2}) = \sum_{\alpha \in Z^{n+1}_+} \beta_\alpha(r, y) \partial^\alpha_y \sigma_i(r, 0_{n+1}, y^{n+2}).$$

If we plug this into (61) we obtain an expression of the form

$$\sigma(r, y) = \sum_{\alpha \in Z^{n+1}_+} c_\alpha(r, y) \partial^\alpha_y \sigma(r, 0_{n+1}, y^{n+2})$$

and we have thus proved the induction step. ∎

We now go on to study the more complicated, but also more interesting, situation when the Lie algebra $\left\{b_1, \dots b_{m_y}\right\}_{LA}$ does not span the whole of R^k. We will need the following geometric result.

Lemma 4.4 *Define the integer* l *by*

$$l = dim \left\{b_1, \dots, b_{m_y}\right\}_{LA},$$

and choose any vector fields $f_1, \dots, f_l$ *such that*

$$span\left\{f_1, \dots, f_l\right\} = \left\{b_1, \dots, b_{m_y}\right\}_{LA}$$

near a given point $y^0 \in R^k$*. Choose furthermore any vectors* $g_{l+1}, \dots, g_k$ *such that* $R^k = Rf_1(y^0) \oplus \dots \oplus Rf_1(y^0) \oplus Rg_{l+1} \oplus \dots \oplus Rg_k$*. Then the mapping* $\Psi : R^k \to R^k$ *defined by*

$$\Psi(s_1, \dots, s_k) = e^{f_1 s_1} \dots e^{f_l s_l} \left(y^0 + \sum_{n=l+1}^{k} s_n g_n\right).$$

is a local diffeomorphism near y^0. Furthermore, defining φ as $\varphi = \Psi^{-1}$ near y^0, we have

$$\varphi_\star \left\{b_1, \ldots, b_{m_y}\right\}_{LA} = \bigoplus_{n=1}^{l} Re_n$$

where $e_1, \ldots, e_l$ are the l first unit vectors in R^k.

Proof. The proof is a fairly straightforward extension of the proofs of Theorem 2.1 and Proposition 2.1 in [7]. ∎

Using this Lemma we may now formulate our final necessary condition.

Proposition 4.5 *Assume that $\dim \{b\}_{LA} = l$, and let $f_1, \ldots, f_l$ be as in Lemma 4.4. For a fixed but arbitrary point $y \in R^k$ and for $s \in R^l$ define the functions $\bar{\mu}(r, s, x; y)$ and $\bar{\sigma}_i(r, s, x; y)$, for $i = 1, \ldots, m_r$ by*

$$\begin{aligned} \bar{\sigma}_i(r, s, x; y) &= \sigma_i(r, e^{f_1 s_1} \ldots e^{f_l s_l} y, x), \\ \bar{\mu}(r, s, x; y) &= \mu(r, e^{f_1 s_1} \ldots e^{f_l s_l} y, x) \end{aligned}$$

Then the following conditions are necessary for the existence of an FDR for the stochastic volatility model.

- *For every (r, y) near (r_0, y_0), it holds that*

$$dim\ span \left\{\partial_s^\alpha \bar{\sigma}_i(r, s; y); \quad \alpha \in Z_+^l\right\} < \infty, \quad i = 1, \ldots m_r \tag{62}$$

  and

$$dim\ span \left\{\partial_s^\alpha \bar{\mu}(r, s; y); \quad \alpha \in Z_+^l\right\} < \infty \tag{63}$$

- *For every (r, y) near (r_0, y_0) the drift and volatility terms have the form*

$$\bar{\sigma}_i(r, s, x; y) = \sum_{j=1}^{n_i} c_{ij}(r, s; y) \lambda_{ij}(r, x; y). \tag{64}$$

  and

$$\bar{\mu}(r, s, x; y) = \sum_{j=1}^{n_0} c_{0j}(r, s; y) \lambda_{0j}(r, x; y). \tag{65}$$

Proof. Choose a fixed y and define Ψ and φ as in Lemma 4.4. We now change coordinates from y to u on R^k by setting $u = \varphi(y)$. We do not however change coordinates on $\mathcal{H}$, so in the new coordinate system we have the variables (r, u) instead of the former variables (r, y). In terms of transformations on $\hat{\mathcal{H}} = \mathcal{H} \times R^k$ we have thus defined a diffeomorphism $\hat{\varphi} : \hat{\mathcal{H}} \to \hat{\mathcal{H}}$ by

$$\hat{\varphi} \begin{bmatrix} r \\ y \end{bmatrix} = \begin{bmatrix} r \\ \varphi(y) \end{bmatrix}$$

and in block operator form we can write

$$\hat{\varphi} = \begin{bmatrix} I \\ \varphi \end{bmatrix}$$

where I denotes the identity mapping on $\mathcal{H}$.

Under this transformation, the original SDE

$$d \begin{bmatrix} r_t \\ y_t \end{bmatrix} = \begin{bmatrix} \mu(r_t, y_t) \\ a(y_t) \end{bmatrix} dt + \begin{bmatrix} \sigma(r_t, y_t) \\ 0 \end{bmatrix} \circ dW_t^r + \begin{bmatrix} 0 \\ b(y_t) \end{bmatrix} \circ dW_t^y$$

will be transformed into

$$d \begin{bmatrix} r_t \\ u_t \end{bmatrix} = \hat{\varphi}_\star \begin{bmatrix} \mu(r_t, u_t) \\ a(u_t) \end{bmatrix} dt + \hat{\varphi}_\star \begin{bmatrix} \sigma(r_t, u_t) \\ 0 \end{bmatrix} \circ dW_t^r + \hat{\varphi}_\star \begin{bmatrix} 0 \\ b(u_t) \end{bmatrix} \circ dW_t^y$$

and it is easily seen that

$$\hat{\varphi}_\star \begin{bmatrix} \sigma(r, u) \\ 0 \end{bmatrix} = \begin{bmatrix} \sigma(r, \varphi^{-1}(u)) \\ 0 \end{bmatrix},$$

and

$$\hat{\varphi}_\star \begin{bmatrix} 0 \\ b(u) \end{bmatrix} = \begin{bmatrix} 0 \\ \varphi_\star b(u) \end{bmatrix}.$$

Since we are looking for necessary conditions we assume that the original model possesses an FDR and thus that the Lie algebra generated by $\hat{\mu}$, $\hat{\sigma}$, and $\hat{b}$ is finite dimensional. Since $\hat{\varphi}_\star$ is a Lie algebra homomorphism, this implies that the Lie algebra generated by $\hat{\varphi}_\star\hat{\mu}$, $\hat{\varphi}_\star\hat{\sigma}$, and $\hat{\varphi}_\star\hat{b}$ is finite dimensional. This Lie algebra obviously includes the subalgebra generated by $\hat{\varphi}_\star\hat{b}$, and again using the fact that $\hat{\varphi}_\star$ preserves the Lie bracket we have

$$\left\{\hat{\varphi}_\star\hat{b}\right\}_{LA} = \hat{\varphi}_\star \left\{\hat{b}\right\}_{LA} = \begin{bmatrix} 0 \\ \varphi_\star \{b\}_{LA} \end{bmatrix}.$$

From this and Lemma 4.4 it now follows that $\left\{\hat{\varphi}_\star\hat{b}\right\}_{LA}$ is in fact spanned by

$$\begin{bmatrix} 0 \\ e_1 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ e_l \end{bmatrix}.$$

where e_i is the i:th unit column vector in R^k. Defining $\tilde{\sigma}$ by $\tilde{\sigma}(r, u) = \sigma(r, \varphi^{-1}(u))$ we thus see that the Lie algebra generated by the vector fields

$$\begin{bmatrix} \tilde{\sigma}_1(r, u) \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} \tilde{\sigma}_{m_r}(r, u) \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ e_1 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ e_l \end{bmatrix}$$

will be finite dimensional. An argument almost identical to the one in the proof of Proposition 4.4 now shows that for each i, and at each point (r, u), we have

$$dim\ span \left\{\partial_u^\alpha \tilde{\sigma}_i(r, u); \quad \alpha \in Z_+^l\right\} < \infty, \tag{66}$$

where the point is that the partial derivatives are only taken over the variables $u_1, \ldots, u_l$. Exactly as in the proof of Proposition 4.4 this implies that there is a representation of the form

$$\tilde{\sigma}_i(r,u) = \sum_{j=1}^{n_i} c_{ij}(r,u)\lambda_{ij}(r, u_{l+1}, \ldots, u_k), \tag{67}$$

where the point is that the vector field λ_{ij} does not depend upon the coordinates $u_1, \ldots, u_l$. Setting $u_{l+1} = \ldots = u_k = 0$ and denoting $u_1, \ldots, u_l$ by $s_1, \ldots, s_l$ in (66) and (67) gives us (62) and (64). The corresponding results for μ are proved similarly. ∎

4.3 Test examples: II.

We illustrate the necessary conditions obtained so far by studying the test examples of Section 2.3. We recall the volatility structures as

1. HW with stochastic a:

$$\sigma(r,y,x) = \sigma e^{-yx} \tag{68}$$

2. HW with stochastic σ:

$$\sigma(r,y,x) = y e^{-ax} \tag{69}$$

3. CIR with stochastic σ:

$$\sigma(r,y,x) = y\sqrt{r(0)} \cdot \lambda(x,y,a) \tag{70}$$

4. CIR with stochastic a:

$$\sigma(r,y,x) = \sigma\sqrt{r(0)} \cdot \lambda(x,\sigma,y) \tag{71}$$

By the assumptions of Section 2.3, all three examples are within the class of orthogonal noise models. We may thus directly apply Proposition 4.2, or (since we have a scalar model) Corollary 4.1 and check whether the corresponding parameterized models possess finite dimensional realizations. In all these cases, however, this test is trivially satisfied since the volatility structures were constructed directly from HJM models possessing short rate realizations. Thus all the models pass this necessary conditions.

We now go on to the necessary conditions of Proposition (4.3). From (49) and ocular inspection of the examples above we immediately have the following result.

Proposition 4.6 *Assuming a scalar y-process with non zero diffusion term, the stochastic volatilities in (68), (70) and (71) do not admit an FDR.*

Thus (68), (70) and (71) are out of the race. In particular it is noteworthy (and perhaps surprising) that there is no stochastic volatility extension of the CIR model (in the sense above) for which there exists a finite dimensional realization. In fact, it is easy to see that we in fact have the following stronger result where we allow both the parameters a and σ to depend upon the process y.

Proposition 4.7 *Consider any stochastic volatility extension of the CIR model of the form*

$$\sigma(r, y, x) = \sigma(y)\sqrt{r(0)} \cdot \lambda(x, \sigma(y), a(y)) \tag{72}$$

where the functions $\sigma(y)$ and $a(y)$ are assumed to be non-constant and where the y process is assumed to have non zero diffusion term. Then the stochastic volatility model does not possess an FDR.

It remains to study the volatility structure (68) in more detail, and this will be done below.

4.4 Necessary and sufficient conditions

In this section we provide necessary and sufficient conditions for the existence of an FDR in the case of an orthogonal noise model, thus improving upon the general results of Theorem 3.1.

We need the following definition.

Definition 4.3 *Define, for each y, the parameterized Lie algebra $\mathcal{L}^y$ on $\mathcal{H}$ by*

$$\mathcal{L}^y = \left\{\partial_y^\alpha \mu^y, \partial_y^\alpha \sigma_1^y, \ldots, \partial_y^\alpha \sigma_{m_r}^y; \quad \alpha \in Z_+^k\right\}_{LA}$$

In this expression $\partial_y^\alpha \mu^y$ is, for each fixed y, considered as a (parameterized) vector field on $\mathcal{H}$, and correspondingly for the σ components.

In order to obtain reasonably concrete results we need to assume that the Lie algebra generated by the b matrix is full dimensional, leaving the general case as an open problem.

Proposition 4.8 *Assume that*

$$dim \left\{b_1, \ldots, b_{m_y}\right\}_{LA} = k. \tag{73}$$

Under this assumption, a necessary and sufficient condition for the existence of an FDR for the stochastic volatility model is that, for each y, we have

$$dim\ \mathcal{L}^y < \infty \tag{74}$$

near r^0.

Proof. From proposition 4.1 we know that there exists an FDR if and only if the Lie algebra $\mathcal{L}$ on $\hat{\mathcal{H}}$ generated by

$$\begin{bmatrix} \mu(r,y) \\ a(y) \end{bmatrix}, \begin{bmatrix} \sigma_1(r,y) \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} \sigma_{m_r}(r,y) \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ b_1(y) \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ b_{m_y}(y) \end{bmatrix}$$

is finite dimensional. Under the assumption (73), and using Gaussian elimination, we see that $\mathcal{L}$ is generated by

$$\begin{bmatrix} \mu(r,y) \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1(r,y) \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} \sigma_{m_r}(r,y) \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ I_k \end{bmatrix},$$

where I_k is the identity matrix on R^k. Using the fact that repeated bracketing of a vector field of the form

$$\begin{bmatrix} f(r,y) \\ 0 \end{bmatrix}$$

with different columns in

$$\begin{bmatrix} 0 \\ I_k \end{bmatrix}$$

will produce a vector field of the form

$$\begin{bmatrix} \partial_y^\alpha f(r,y) \\ 0 \end{bmatrix}$$

it now follows that $\mathcal{L}$ is in fact generated by

$$\begin{bmatrix} \partial_y^\alpha \mu(r,y) \\ 0 \end{bmatrix}, \begin{bmatrix} \partial_y^\alpha \sigma(r,y) \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} \partial_y^\alpha \sigma_{m_r}(r,y) \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ I_k \end{bmatrix}; \quad \alpha \in Z_+^k$$

From this it is clear that $\mathcal{L}$ is generated by

$$\begin{bmatrix} \mathcal{L}^y \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ I_k \end{bmatrix}; \quad y \in R^k,$$

and the proof is finished if we can show that for each multi index α we have

$$\partial_y^\alpha \ \mathcal{L}^y \subseteq \mathcal{L}^y. \tag{75}$$

It follows by induction that in order to prove (75) we may WLOG assume that $k = 1$ (i.e.y is scalar) and that it is in fact enough to prove that

$$\partial_y \ \mathcal{L}^y \subseteq \mathcal{L}^y. \tag{76}$$

Now, it is easily seen that

$$\mathcal{L}^y = \bigcup_{k=0}^{\infty} L_k^y,$$

where

$$\begin{aligned} L_0^y &= span\left\{\partial_y^n \mu^y, \partial_y^n \sigma_1^y, \ldots, \partial_y^n \sigma_{m_r}^y; \quad n \geq 0\right\} \\ L_{k+1}^y &= span\left\{L_k^y, [L_k^y, L_k^y]\right\}, \quad k = 0, 1, \ldots \end{aligned}$$

so it is enough to prove that each L_k^y is invariant under ∂_y and we prove this by induction. The case $k = 0$ is clear, so assume that

$$\partial_y L_n^y \subseteq L_n^y$$

for all $n \leq k$. Now fix an arbitrary $f \in L_{k+1}^y$. We start by considering two cases: the case when $f \in L_k^y$ and the case when $f = [g, h]$ with $g, h \in L_k^y$. If $f \in L_k^y$ then $\partial_y \in L_k^y$ by the induction assumption, so $\partial_y \in L_{k+1}^y$. If $f = [g, h]$ with $g, h \in L_k^y$ then an easy calculation shows that

$$\partial_y f = \left[\partial_y g, h\right] + \left[g, \partial_y h\right]$$

which is in $[L_k^y, L_k^y]$ by the induction assumption. Thus also in this case we have $\partial_y f \in L_{k+1}^y$. A generic $f \in L_{k+1}^y$ is, by definition, a linear combination of terms of the above type so we are finished. ∎

4.5 A simple sufficient condition

The object of this section is to show that, under some rather restrictive but nontrivial assumptions, it is possible to derive an extremely simple sufficient condition for the existence of an FDR for the full stochastic volatility model in terms of the FDR for the parameterized model. Furthermore; under these assumptions the realization for the full model can be constructed directly, and in a trivial manner, from the realization for the parameterized model.

Assumption 4.2

1. *The Ito formulation of the r-dynamics of the stochastic volatility model is of the form*
$$dr_t = \mu_0(r_t, y_t)dt + \sigma_t(r_t, y_t)dW_t. \tag{77}$$

2. *We assume that y is independent of W. Apart from this assumption, the process y is allowed to be an arbitrary semimartingale with values in R^k.*

3. *For any fixed y, the parameterized r-model is assumed to possess an FDR of the form*
$$r_t^y = G(Z_t^y), \tag{78}$$
$$dZ_t^y = A(Z_t^y, y)dt + B(Z_t^y, y) \circ dW_t, \tag{79}$$
where Z^y is R^d valued and G is a smooth mapping $G : R^d \to \mathcal{H}$.

The important part of this assumption is that, for the parameterized model, the parameter y only appears in the Z^y dynamics, but not the output mapping G. We will discuss the geometric significance of this below, but first we state the result.

Proposition 4.9 *Under Assumption 4.2, the stochastic volatility model possesses an FDR, and a concrete realization is in fact given by*

$$
\begin{aligned}
r_t &= G(Z_t), && (80)\\
dZ_t &= A(Z_t, y_t)dt + B(Z_t, y_t) \circ dW_t, && (81)
\end{aligned}
$$

With G, A and B as in (78)-(79).

Proof. From the independence between y and W it follows that the Stratonovich formulation of the r-dynamics is given by

$$
dr_t = \mu(r_t, y_t)dt + \sigma(r_t, y_t) \circ dW_t, \tag{82}
$$

where

$$
\mu(r, y) = \mu_0(r, y) - \frac{1}{2}\sigma_r(r, y)\sigma(r, y).
$$

Now let us consider (80)-(81) as an Ansatz. The r-dynamics induced by (80)-(81) are given by

$$
dr_t = G'(Z_t)A(Z_t, y_t)dt + G'(Z_t)B(Z_t, y_t) \circ dW_t, \tag{83}
$$

so it follows that (80)-(81) is a realization of (82) if and only if

$$
\begin{aligned}
\mu(r, y) &= G_\star A(r, y), && (84)\\
\sigma(r, y) &= G_\star B(r, y). && (85)
\end{aligned}
$$

We thus have to prove that (84)-(85) hold, and to this end we use the fact that, by assumption, (78)-(79) is a realization for the parameterized model. The Stratonovich formulation for the parameterized model is easily seen to be given by

$$
dr^y_t = \mu(r^y_t, y)dt + \sigma(r^y_t, y) \circ dW_t, \tag{86}
$$

and the important point here is that this is precisely the parameterized version of the Stratonovich formulation of the original r-dynamics. The r^y-dynamics induced by (78)-(79) are given by

$$
dr^y_t = G'(Z^y_t)A(Z^y_t, y)dt + G'(Z^y_t)B(Z^y_t, y) \circ dW_t, \tag{87}
$$

and since this was assumed to be a realization of (86) we thus have

$$
\begin{aligned}
\mu(r, y) &= G_\star A(r, y),\\
\sigma(r, y) &= G_\star B(r, y),
\end{aligned}
$$

which was to be proved. ∎

Remark 4.1 If the Stratonovich differential in (79) is replaced by an Itô differential i.e. by

$$dZ_t^y = A(Z_t^y, y)dt + B(Z_t^y, y)dW_t,$$

then the conclusion of Proposition 4.9 still holds if the Stratonovich differential in (81) is replaced by an Itô differential, i,.e. by

$$dZ_t = A(Z_t, y_t)dt + B(Z_t, y_t)dW_t.$$

This is useful if the realization of the parameterized model is originally given in Itô form.

This, very strong but also very restrictive, result has a clear and simple geometric interpretation. First, we know from general (orthogonal noise) theory that a necessary condition for an FDR is that the parameterized model possesses an FDR. In general, the realization for the parameterized model will of course be of the form

$$r_t^y = G(Z_t^y, y), \tag{88}$$

$$dZ_t^y = A(Z_t^y, y)dt + B(Z_t^y, y) \circ dW_t, \tag{89}$$

where the output function G as well as the drift term A and diffusion term B depend upon y, but in Proposition 4.9 we have assumed that G does not in fact depend on y. To understand the geometric meaning of this assumption we recall from [7] that the parameterized model, for a fixed y, admits an FDR if and only if there exists an invariant manifold $\mathcal{G}^y$ passing through r^0, and in the generic case this invariant manifold will of course depend upon y. The relation between $\mathcal{G}^y$ and the realization (88)-(89) is that

$$\mathcal{G}^y = Im\ G^y,$$

where the mapping $G^y : R^d \rightarrow \mathcal{H}$ is defined by $G^y(z) = G(z, y)$. Thus; assuming that G does not depend upon the parameter y is equivalent to assuming that the invariant manifold for the parameterized model passing through r^0 does not depend upon y. In that case, denoting the invariant manifold by $\mathcal{G}$ it is of course geometrically obvious that $\mathcal{G} \times R^k$ will be a finite dimensional invariant manifold for the process (r_t, y_t) thus guaranteeing the existence of an FDR for the full model.

Furthermore, it follows from Theorem A.2 that the invariant manifold $\mathcal{G}^y$ is determined uniquely by the parameterized Lie algebra

$$\mathcal{L}^y = \left\{\mu^y, \sigma_1^y, \ldots, \sigma_{m_r}^y\right\}_{LA}, \tag{90}$$

so if $\mathcal{L}^y$ does not depend upon y then neither will $G(z, y)$. We thus have the following result.

Proposition 4.10 *Assume that*

- *The process y is an R^k-valued semimartingale which is independent of W.*

- *The parameterized model admits an FDR for every fixed y.*
- *Lie algebra $\mathcal{L}^y$ defined in (90) does not depend upon the parameter y.*

Then the full model will possess an FDR.

We finish this discussion by noticing that for the general Lie algebraic machinery of [7] and [17] to work it is essential that all processes are Wiener driven. The geometric reason for this is that the Wiener process acts locally in space (the infinitesimal generator is a partial differential operator) and this allows us to analyze the realization problems using differential geometry (i.e. local analysis). It is therefore noteworthy that in the simple situation discussed above in this section, we did not have to assume that y is driven by a Wiener process – it can also have jumps.

4.6 An example

As an application of the results in Section 4.5, we consider the following volatility structure for a standard forward rate model driven by a scalar Wiener process W^r,

$$\sigma(r, x) = \varphi(r)e^{-\alpha x}. \tag{91}$$

Here φ is assumed to be an arbitrarily chosen smooth scalar field, and α is a positive constant. This is an extension of the model investigated in [24], where an FDR was constructed for the case when φ was assumed to be of the particular form $\varphi(r) = g\,(r(0))$, for some smooth function $g : R \to R$. As was shown in [7], also the extended model admits an FDR, and from [6] a realization is easily obtained in the following way.

Define the mapping $G : R_+ \times R^2 \to \mathcal{H}$ by

$$G(t, z_1, z_2)(x) = r^0(x + t) + z_1 e^{-\alpha x} + z_2 e^{-2\alpha x} \tag{92}$$

The realization is then given by

$$r_t(x) = G\,(t, Z_1(t), Z_2(t))\,(x), \tag{93}$$

$$dZ_1(t) = \left\{ \frac{1}{\alpha}\varphi^2\,[G_t] - \alpha Z_1(t) \right\} dt + \varphi\,[G_t]\,dW_t^r, \tag{94}$$

$$dZ_2(t) = -\left\{ 2\alpha Z_1(t) + \frac{1}{\alpha}\varphi^2\,[G_t] \right\} dt. \tag{95}$$

where we have used the shorthand notation

$$G_t = G\,(t, Z_1(t), Z_2(t))\,.$$

The important point to notice is that the mapping G in (92) does not involve φ. We may now extend the model above to a stochastic volatility model with an arbitrary scalar y-process (assumed to be independent of W^r), by defining the volatility structure as

$$\sigma(r, y, x) = \varphi(r, y)e^{-\alpha x}. \quad (96)$$

where φ is an arbitrarily chosen scalar field.

By construction, the parameterized model admits an FDR of the form (93)-(95) where G is exactly as above, and where $\varphi[G_t]$ is replaced by $\varphi[G_t, y]$. The point is again that G does not involve y , so it now follows immediately from Proposition 4.9 that a realization for the stochastic volatility model is given by

$$\begin{aligned} r_t(x) &= G(t, Z_1(t), Z_2(t))(x), \\ dZ_1(t) &= \left\{ \frac{1}{\alpha}\varphi^2[G_t, y_t] - \alpha Z_1(t) \right\} dt \\ &+ \varphi[G_t, y_t] dW_t^r, \\ dZ_2(t) &= -\left\{ 2\alpha Z_1(t) + \frac{1}{\alpha}\varphi^2[G_t, y_t] \right\} dt. \end{aligned}$$

Remark 4.2 In this example we have used the Itô dynamics instead of the Stratonovich dynamics. The reason is that the Itô dynamics of the realization are simpler than the Stratonovich dynamics.

5 Forward rate models

We now go on to apply the general results above to the more concrete case of forward rate models. we recall that the Ito formulation of the stochastic volatility forward rate model is given by

$$dr_t(x) = \left\{ \frac{\partial}{\partial x} r_t(x) + \mathbf{H}\sigma(r_t, y_t, x) \right\} dt + \sigma(r_t, y_t, x)dW_t \quad (97)$$

$$dy_t = a^0(y_t)dt + b(y_t)dW_t, \quad (98)$$

where $\mathbf{H}$ is defined in (3). On Stratonovich form the model has the form

$$\begin{aligned} dr_t &= \mu(r_t, y_t)dt + \sigma(r_t, y_t) \circ dW_t & (99) \\ dy_t &= a(y_t)dt + b(y_t) \circ dW_t, & (100) \end{aligned}$$

where

$$\mu(r, y) = \mathbf{F}r + \mathbf{H}\sigma(r, y) - \frac{1}{2}\sigma_r(r, y)\sigma(r, y) - \frac{1}{2}\sigma_y(r, y)b(y) \quad (101)$$

$$a(y) = a^0(y) - \frac{1}{2}b_y(y)b(y). \quad (102)$$

As usual $\mathbf{F}$ denotes the operator $\frac{\partial}{\partial x}$, σ_r denotes the partial Frechet derivative of σ w.r.t. the vector variable r and similarly for σ_y.

5.1 Necessary conditions for orthogonal noise models

In the orthogonal noise case the model has the following Stratonovich form

$$
\begin{aligned}
dr_t &= \mu(r_t, y_t)dt + \sigma(r_t, y_t) \circ dW_t^r && (103)\\
dy_t &= a(y_t)dt + b(y_t) \circ dW_t^y, && (104)
\end{aligned}
$$

where

$$
\begin{aligned}
\mu(r, y) &= \mathbf{F}r + \mathbf{H}\sigma(r, y) - \frac{1}{2}\sigma_r(r, y)\sigma(r, y) && (105)\\
a(y) &= a^0(y) - \frac{1}{2}b_y(y)b(y). && (106)
\end{aligned}
$$

We now have the following surprisingly restrictive result.

Proposition 5.1 *Assume the following:*

- *The model is an orthogonal noise model.*
- *The condition*
$$\left\{b_1, \ldots b_{m_y}\right\}_{LA} = R^k, \tag{107}$$
is satisfied near y^0.

Then, a necessary condition for the existence of an FDR is that the volatility structure has the form

$$\sigma_i(r, y, x) = \sum_{j=1}^{N} \varphi_{ij}(r, y)\lambda_j(x), \quad i = 1, \ldots, m_r, \tag{108}$$

where $\lambda_1, \ldots, \lambda_N$ are constant vector fields, and φ_{ij} are smooth scalar fields.

Proof. Since we have assumed orthogonal noise, Proposition 4.2 implies that a necessary condition for the existence of an FDR is that the parameterized model admits an FDR. Furthermore; applying Theorem 4.13 of [17] to the parameterized model it follows that the volatility must be of the form

$$\sigma_i(r, y, x) = \sum_{j=1}^{N} \varphi_{ij}(r, y)\lambda_j(y, x). \tag{109}$$

Given this expression, an application of Proposition 4.4 finishes the proof. ∎

Given a volatility structure of the form (108) we now go on to find sufficient conditions for the existence of an FDR.

5.2 Sufficient conditions for the general noise models

We now consider a multidimensional forward rate model of the form

$$\begin{aligned} dr_t &= \mu(r_t, y_t)dt + \sigma(r_t, y_t) \circ dW_t && (110) \\ dy_t &= a(y_t)dt + b(y_t) \circ dW_t. && (111) \end{aligned}$$

where W is assumed to be m-dimensional, and y is as usual k-dimensional. We will assume that the volatility structure is of the form (108), but we stress the fact that we do not restrict ourselves to the orthogonal noise model.

We recall from [7] that a real valued function $f : R \to R$ is said to be *quasi exponential* if it can be written as

$$f(x) = ce^{Ax}b,$$

where c is a row vector, b is a column vector and A is a matrix. It is easy to see that a function is quasi exponential if and only if it satisfies a linear ordinary differential equation with constant coefficients. The general form of a quasi exponential function is given by

$$f(x) = \sum_i e^{\gamma_i x} + \sum_j e^{\alpha_j x} \left[p_j(x) \cos(\omega_j x) + q_j(x) \sin(\omega_j x) \right], \tag{112}$$

where $\gamma_i, \alpha_j, \omega_j$ are real numbers, whereas p_j and q_j are real polynomials.

The main result is as follows.

Proposition 5.2 *Consider the model (110)-(111) and assume that the components of σ are of the form*

$$\sigma_i(r, y, x) = \sum_{j=1}^{N} \varphi_{ij}(r, y)\lambda_j(x), \quad i = 1, \ldots, m. \tag{113}$$

Under this assumption a sufficient condition for the existence of an FDR is that $\lambda_1(x), \ldots, \lambda_m(x)$ are quasi exponential. The scalar fields $\varphi_{ij}(x)$ are allowed to be arbitrary.

Proof. In order to avoid to much and messy notation, we give the proof only for the simplified case when

$$\sigma_i(r, y, x) = \varphi_i(r, y)\lambda_i(x).$$

The arguments in the general case are almost identical. Under the given assumption the Stratonovich drift term of r is given by

$$\mu = \mathbf{F}r + \sum_{i=1}^{m} \Phi_i D_i - \frac{1}{2} \sum_{i=1}^{m} \varphi_{ir}[\lambda_i]\varphi_i\lambda_i - \frac{1}{2} \sum_{i=1}^{m} \varphi_{iy}[b_i]\lambda_i \tag{114}$$

where b_i denotes the i.th column of the matrix b. The Lie algebra $\mathcal{L}$ under study is the one generated by the vector fields

$$\begin{bmatrix} \mu \\ a \end{bmatrix}, \begin{bmatrix} \varphi_1\lambda_1 \\ b_1 \end{bmatrix}, \ldots, \begin{bmatrix} \varphi_m\lambda_m \\ b_m \end{bmatrix}.$$

Obviously, $\mathcal{L}$ is included in the larger algebra $\mathcal{L}_1$, generated by

$$\begin{bmatrix} \mu \\ 0 \end{bmatrix}, \begin{bmatrix} \varphi_1\lambda_1 \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} \varphi_m\lambda_m \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ a \end{bmatrix}, \begin{bmatrix} 0 \\ b_1 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ b_m \end{bmatrix}.$$

Using the structure of μ we can reduce this generator system to

$$\begin{bmatrix} \mathbf{F}r + \sum_{i=1}^m \Phi_i D_i \\ 0 \end{bmatrix}, \begin{bmatrix} \lambda_1 \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} \lambda_m \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ a \end{bmatrix}, \begin{bmatrix} 0 \\ b \end{bmatrix}.$$

From this we see that $\mathcal{L}_1$ is included in the algebra $\mathcal{L}_2$, generated by

$$\begin{bmatrix} \mathbf{F}r \\ 0 \end{bmatrix}, \begin{bmatrix} D_1 \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} D_m \\ 0 \end{bmatrix}, \begin{bmatrix} \lambda_1 \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} \lambda_m \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ a \end{bmatrix}, \begin{bmatrix} 0 \\ b \end{bmatrix}.$$

As in the previous section, it is now easily seen (see Section 5 in [7]) that $\mathcal{L}_2$ is finite dimensional if and only if $\lambda_1, \ldots, \lambda_m$ are quasi exponential. ∎

5.3 The scalar case

We finish by a reasonably complete investigation of the most important special case, which occurs when y is scalar, r and y are driven by scalar Wiener processes, and the volatility has the form

$$\sigma(r, y, x) = \varphi(r, y)\lambda(x). \tag{115}$$

Such a model will have the form

$$\begin{aligned} dr_t(x) &= \{\mathbf{F}r_t(x) + \Phi(r, y)D(x)\}\, dt + \varphi(r, y)\lambda(x)dW_t^r \\ dy_t &= a^0(y_t)dt + b(y_t)dW_t^y. \end{aligned}$$

where

$$\begin{aligned} \Phi(r, y) &= \varphi^2(r, y), \\ D(x) &= \lambda(x) \int_0^x \lambda(s)ds. \end{aligned}$$

In order to allow for a correlation, ρ, between W^r and W^y we write them as

$$\begin{aligned} W_t^r &= \rho W_t^1 + \sqrt{1 - \rho^2} W_t^2, \\ W_t^y &= W_t^1 \end{aligned}$$

where W^1 and W^2 are independent Wiener processes. We then have the dynamics

$$\begin{aligned} dr_t &= \{\mathbf{F}r_t + \Phi D\} \, dt + \varphi\lambda\rho W_t^1 + \varphi\lambda\sqrt{1-\rho^2} W_t^2 \\ dy_t &= a^0 dt + b dW_t^1. \end{aligned}$$

We can now prove the following main result for the scalar case.

Proposition 5.3 *Assume that $\varphi_y(r, y) \neq 0$, and that $b(y) \neq 0$ i.e. that the model is non trivial. Then the following hold.*

- *In the non-perfectly correlated case $|\rho| < 1$, a necessary and sufficient condition for the existence of an FDR is that the vector field λ is quasi exponential. The scalar field $\varphi(r, y)$ is allowed to be arbitrary.*
- *In the perfectly correlated case $|\rho| = 1$, the condition above is sufficient.*

Proof. The Stratonovich dynamics of the model are given by

$$\begin{aligned} dr_t &= \left\{ \mathbf{F}r_t + \Phi D - \frac{1}{2}\varphi_r[\lambda]\varphi\lambda - \frac{1}{2}\varphi_y b\lambda \right\} dt + \varphi\lambda \circ W_t^1 + \sqrt{1-\rho^2}\varphi\lambda \circ W_t^2 \\ dy_t &= a dt + b \circ dW_t^1. \end{aligned}$$

Thus the relevant Lie algebra $\mathcal{L}$ on $\hat{\mathcal{H}}$ is generated by the vector fields

$$\begin{bmatrix} \mathbf{F}r + \Phi D - \frac{1}{2}\varphi_r[\lambda]\varphi\lambda - \frac{1}{2}\varphi_y b\lambda \\ a \end{bmatrix}, \begin{bmatrix} \rho\varphi\lambda \\ b \end{bmatrix}, \begin{bmatrix} \sqrt{1-\rho^2}\varphi\lambda \\ 0 \end{bmatrix},$$

We start with the non-perfectly correlated case, so we assume that $|\rho| < 1$. Then, by Gaussian elimination, the system of generators can immediately be reduced to

$$\begin{bmatrix} \mathbf{F}r + \Phi D \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} \lambda \\ 0 \end{bmatrix}$$

The Lie bracket between the first two vector fields gives us

$$\begin{bmatrix} \Phi_y D \\ 0 \end{bmatrix},$$

so after reducing this field we have the generators

$$\begin{bmatrix} \mathbf{F}r + \Phi D \\ 0 \end{bmatrix}, \begin{bmatrix} D \\ 0 \end{bmatrix}, \begin{bmatrix} \lambda \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

which finally reduce to

$$\begin{bmatrix} \mathbf{F}r \\ 0 \end{bmatrix}, \begin{bmatrix} D \\ 0 \end{bmatrix}, \begin{bmatrix} \lambda \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

From this it follows immediately (see [7] section 5) that the Lie algebra is finite dimensional if an only if the linear span of

$$\{\mathbf{F}^n\lambda, \quad \mathbf{F}^n D; \quad n \geq 0\}$$

is a finite dimensional subspace in $\mathcal{H}$. It is however easily seen that this happens if and only if λ is quasi exponential.

In the perfectly correlated case $|\rho| = 1$ we can WLOG assume that $\rho = 1$ and we are left with the following generators for the Lie algebra $\mathcal{L}$.

$$\left[\begin{array}{c} \mathbf{F}r + \Phi D - \frac{1}{2}\varphi_r[\lambda]\varphi\lambda - \frac{1}{2}\varphi_y b\lambda \\ a \end{array}\right], \left[\begin{array}{c} \varphi\lambda \\ b \end{array}\right],$$

There seems tho be no easy way of reducing this set of generators, but it is obvious that $\mathcal{L}$ is included in the Lie algebra $\mathcal{L}_{ext}$ generated by the fields

$$\left[\begin{array}{c} \mathbf{F}r + \Phi D - \frac{1}{2}\varphi_r[\lambda]\varphi\lambda - \frac{1}{2}\varphi_y b\lambda \\ a \end{array}\right], \left[\begin{array}{c} \varphi\lambda \\ 0 \end{array}\right], \left[\begin{array}{c} 0 \\ b \end{array}\right]$$

Thus a sufficient condition for an FDR is that the larger Lie algebra $\mathcal{L}_{ext}$ is finite dimensional. It is however easily seen that $\mathcal{L}_{ext}$ is identical with the algebra discussed in the non-perfectly correlated case above, so we are finished. ∎

5.4 Test examples: III.

We can now continue our study of the test examples of Section 2.3. In fact, only one example is left in the race, namely

2. HW with stochastic σ:

$$\sigma(r, y, x) = ye^{-ax}. \tag{116}$$

We now have the following result, which is immediately obtained from Proposition 5.3.

Proposition 5.4 *The stochastic volatility version of the Hull-White extended Vasiček volatility structure with stochastic σ, as in (116) admits an FDR.*

5.5 Construction of realizations

In the previous sections we have provided existence results for FDRs, but so far we have not constructed any concrete realizations. The object of this section is to present a general method for the construction of an FDR for any multidimensional forward rate model of the form (110)-(111), for which there exists an FDR. We will then apply this methodology to concrete cases. The methodology is basically the one presented in [6], so we will be rather brief.

5.5.1 General theory, results and examples

A general method for constructing an FDR of a forward rate system, for which such a realization exists, was presented in [6]. This method involved the following steps:

- Choose a finite number of vector fields $f_1, \ldots, f_d$ such that

$$\{\hat{\mu}, \hat{\sigma}\}_{LA} \subseteq span\{f_1, \ldots, f_d\}. \tag{117}$$

- Compute the mapping $\widehat{G} : R^d \to \mathcal{H} \times R^k$ using the formula

$$\widehat{G}(z_1, \ldots, z_d) = e^{f_d z_d} \ldots e^{f_1 z_1} \hat{r}_0. \tag{118}$$

  Here $\hat{r}_0$ denotes an initial point admitting an FDR and for the definition of $e^{ft}\hat{r}$ we refer to Definition A.3 in the Appendix.

- We now have that $\hat{r} = \widehat{G}(Z)$. Make the following *Ansatz* for the dynamics of the state space variables Z

$$dZ = A(Z)dt + B(Z) \circ dW_t. \tag{119}$$

  It must then hold that

$$\widehat{G}_* A = \hat{\mu}, \quad \widehat{G}_* B = \hat{\sigma}. \tag{120}$$

  Now use the equations in (120) to obtain the vector fields A and B.

To insure that the model we are considering possesses an FDR we recall Proposition 5.2 and assume that the components of σ are of the form

$$\sigma_i(r, y, x) = \varphi_i(r, y)\lambda_i(x), \quad i = 1, \ldots, m, \tag{121}$$

where φ_i is an arbitrary smooth functional $i = 1, \ldots, m$ and λ_i is given by

$$\lambda_i(x) = p_i(x)e^{\alpha_i x}, \quad i = 1, \ldots, m. \tag{122}$$

Here p_i is a polynomial of degree n_i and α_i is a scalar constant. Furthermore we will assume that

$$\{b_1, \ldots, b_m\}_{LA} = R^k. \tag{123}$$

Working through the steps outlined above will prove the proposition stated below, provided you choose the vector fields mentioned in the first step as follows

$$\begin{bmatrix} \mathbf{F}r \\ 0 \end{bmatrix}, \begin{bmatrix} \mathbf{F}^j\lambda_i \\ 0 \end{bmatrix}, \begin{bmatrix} \mathbf{F}^l\widetilde{D}_i \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ e_i \end{bmatrix}. \tag{124}$$

Here $\widetilde{D}_i$ is defined by

$$\widetilde{D}_i(x) = D_i(x) - \gamma_i\lambda_i(x), \quad i = 1, \ldots, m, \tag{125}$$

where the constant γ_i is given by

$$\gamma_i = \sum_{j=0}^{n_i} \left(\frac{-1}{\alpha_i} \right)^{j+1} \mathbf{F}^j p_i(0), \quad i = 1, \ldots, m, \tag{126}$$

furthermore, e_i denotes the i:th unit vector in R^k, $i = 1, \ldots, m$; $j = 0, 1, \ldots, n_i$ and $l = 0, 1, \ldots, q_i = 2n_i$.

Proposition 5.5 *Given the initial point $\hat{r}_0$ the forward rate system (110)-(111) with volatility defined by equations (121)-122 and satisfying the condition (123) has a finite dimensional realization given by*

$$\hat{r}_t = \widehat{G}(Z_t), \tag{127}$$

where $\widehat{G}$ is given by

$$\widehat{G}(z_0, z_{ij}^1, z_{il}^2, z_p^3) = \left[\begin{array}{c} G(z_0, z_{ij}^1, z_{il}^2, z_p^3) \\ y_0 + z^3 \end{array} \right]. \tag{128}$$

Here G is given by

$$G(z_0, z_{ij}^1, z_{il}^2, z_p^3)(x) = r_0(x+z_0) + \sum_{i=1}^{m} \sum_{j=0}^{n_i} \mathbf{F}^j \lambda_i(x) z_{ij}^1 + \sum_{i=1}^{m} \sum_{l=0}^{2n_i} \mathbf{F}^l \widetilde{D}_i(x) z_{il}^2, \tag{129}$$

z^3 denotes the column vector $z^3 = (z_1^3, \ldots, z_k^3)^*$, $i = 1, \ldots, m$; $j = 0, 1, \ldots, n_i$; $l = 0, 1, \ldots, q_i = 2n_i$ and $p = 1, \ldots, k$.

The dynamics of the state space variables are given by

$$\left\{ \begin{array}{rcl} dZ_0 &=& dt, \\ dZ_{i0}^1 &=& [c_{i0} Z_{in_i}^1 + \gamma_i \Phi_i(\widehat{G}(Z))]dt + \varphi_i(\widehat{G}(Z))dW_t, \\ dZ_{ij}^1 &=& (c_{ij} Z_{in_i}^1 + Z_{i,j-1}^1)dt \\ dZ_{i0}^2 &=& (d_{i0} Z_{iq_i}^2 + \Phi_i(\widehat{G}(Z)))dt \\ dZ_{il}^2 &=& (d_{il} Z_{il}^2 + Z_{i,l-1}^2)dt, \\ dZ^3 &=& a_0(Z_t^3)dt + b(Z_t^3)dW_t. \end{array} \right. \tag{130}$$

Here c_{ij} and d_{il} denote the constants

$$c_{ij} = -\left(\begin{array}{c} n_i + 1 \\ j \end{array} \right) (-\alpha_i)^{n_i+1-j}, \tag{131}$$

and

$$d_{il} = -\left(\begin{array}{c} 2n_i + 1 \\ l \end{array} \right) (-2\alpha_i)^{2n_i+1-l}, \tag{132}$$

respectively, the constants γ_i *were defined in (126),* $\Phi_i = \varphi_i^2$*,* $a_0(y)$ *is given by* $a_0(y) = a(y) + \frac{1}{2}b_y(y)b(y)$*,* $i = 1, \ldots, m$*;* $j = 0, 1, \ldots, n_i$ *and* $l = 0, 1, \ldots, q_i = 2n_i$*.*

Remark 5.1 *The reader familiar with the paper [6] will recognize that the realization of the model with stochastic volatility is almost identical to the realization of the model without stochastic volatility, except of course for the inclusion of the* y*-dynamics in the form of* Z^3*.*

5.5.2 A simple special case: Hull-White

To see the general results above in a very simple special case, we now apply Proposition 5.5 to the Hull-White extended Vasiček model with stochastic volatility, i.e. the model with volatilities given by

$$\sigma(r, y, x) = ye^{-\alpha x}. \tag{133}$$

Note that we have $k = 1$, $m = 1$ and $n = 0$. If we assume that the dynamics of y satisfy the condition (123) Proposition 5.5 gives us that, given $\hat{r}_0 = (r_0, y_0)^*$, the forward rate model of the form (110)-(111), with volatilities given by (133), has a finite dimensional realization given by

$$\hat{r}_t = \widehat{G}(Z_t). \tag{134}$$

Here $\widehat{G}$ is defined by

$$\widehat{G}(z_0, z_1, z_2, z_3) = \begin{bmatrix} G(z_0, z_1, z_2, z_3) \\ y_0 + z_3 \end{bmatrix}, \tag{135}$$

where G is given by

$$G(z_0, z_1, z_2, z_3)(x) = r_0(x + z_0) + e^{-\alpha x} z_1 - \frac{e^{-2\alpha x}}{\alpha} z_2. \tag{136}$$

The dynamics of the state space variables are given by

$$\left\{ \begin{array}{lcl} dZ_0 & = & dt, \\ dZ_1 & = & \left[-\alpha Z_1 + \frac{1}{\alpha}(y_0 + Z_3)^2) \right] dt + (y_0 + Z_3) dW_t, \\ dZ_2 & = & \left[-2\alpha Z_2 + (y_0 + Z_3)^2 \right] dt \\ dZ_3 & = & a_0(Z_3)dt + b(Z_3)dW_t. \end{array} \right. \tag{137}$$

Here $a_0(y) = a(y) + \frac{1}{2}b_y(y)b(y)$.

A Appendix: Realization theory in Hilbert space

In this appendix we will give a brief recapitulation of Lie algebra theory for the existence of FDRs in Hilbert space, developed in [7]. See [7] for proofs and details and [3] for an overview.

A.1 Problem statement

Take as given an m-dimensional standard Wiener process W with components $W^1, \ldots, W^m$, and a separable Hilbert space $\hat{\mathcal{H}}$, where a generic point will be denoted by $\hat{r}$. (In our applications to interest rates we will of course choose $\hat{r}$ as the extended process (r, y) like in Section 2.1.) Let furthermore $\hat{\mu}, \hat{\sigma}_1, \ldots, \hat{\sigma}_m$ be smooth vector fields on $\hat{\mathcal{H}}$. For a given initial point $\hat{r}^o \in \hat{\mathcal{H}}$ we can then consider the following SDE on $\mathcal{H}$.

$$\begin{cases} d\hat{r}_t &= \hat{\mu}(\hat{r}_t)dt + \hat{\sigma}(\hat{r}_t) \circ dW_t, \\ \hat{r}_0 &= \hat{r}^o, \end{cases} \tag{138}$$

Here $\hat{\sigma}(\hat{r}_t) \circ dW_t = \sum_1^m \hat{\sigma}_i(\hat{r}_t) \circ dW_t^i$ and $\circ$ denotes the Stratonovich integral (see Remark A.1 below). For information on SDEs in Hilbert space see [13].

Definition A.1 *We say that the SDE (138) has a* **finite dimensional realization** *(FDR) if there exists a point $z_0 \in R^d$, smooth vector fields $A, B_1, \ldots, B_m$ on some open subset $\mathcal{Z}$ of R^d and a smooth (submanifold) map $\hat{G} : \mathcal{Z} \to \hat{\mathcal{H}}$, such that $\hat{r}$ has the local representation*

$$\hat{r}_t = \hat{G}(Z_t), \quad P - a.s. \tag{139}$$

where Z is the strong solution of the d-dimensional Stratonovich SDE

$$\begin{cases} dZ_t &= A(Z_t)dt + B(Z_t) \circ dW_t, \\ Z_0 &= z_0, \end{cases} \tag{140}$$

and where the driving Wiener process W in (140) is the same as in (138). The prefix "local" above means that the representation is assumed to hold for all t with $0 \leq t < \tau(r^o)$, P-a.s. where, for each $\hat{r}^o \in \hat{\mathcal{H}}$, $\tau(\hat{r}^o)$ is a strictly positive stopping time.

Remark A.1 Note the use of the Stratonovich integral. The reason for this is that the main theorems are most naturally formulated (and proved) within the Stratonovich framework. If (as is the case for us) the original problem is stated in Itô terms, this simply means that you translate your Itô equations into the corresponding Stratonovich ones. For that purpose we recall that, if X is a semimartingale of suitable dimension, the Stratonovich integral/differential concept is related to the corresponding Itô concept by the formula

$$X_t \circ dW_t = X_t dW_t + \frac{1}{2} d\langle X, W \rangle_t \tag{141}$$

We also recall that for the Stratonovich integral, the Itô formula for continuous semimartingales takes the form

$$dF(t, X_t) = \frac{\partial F}{\partial t}(t, X_t)dt + \frac{\partial F}{\partial x}(t, X_t) \circ dX_t \tag{142}$$

for any $F \in C^{1,3}$.

The main problem is now to find necessary and sufficient conditions on $\hat{\mu}$ and $\hat{\sigma}$ for the existence of an FDR. For this we need some basic terminology from differential geometry.

A.2 Basic concepts in differential geometry

We recall the following concepts from infinite dimensional differential geometry. We only give the global definitions, but there are also local versions (see [7]) which will use without comment.

Consider a real Hilbert space $\hat{\mathcal{H}}$. By an n-dimensional **distribution** we mean a mapping F, which to each $\hat{r} \in \hat{\mathcal{H}}$ associates an n-dimensional subspace $F(\hat{r}) \subseteq \hat{\mathcal{H}}$. A mapping (vector field) $f : \hat{\mathcal{H}} \to \hat{\mathcal{H}}$, is said to **lie in** F if $f(\hat{r}) \in F(\hat{r})$ for every $\hat{r} \in \hat{\mathcal{H}}$. A collection $f_1, \ldots, f_n$ of vector fields lying in F **generates** (or spans) F if $span\,\{f_1(\hat{r}), \ldots, f_n(\hat{r})\} = F(\hat{r})$ for every $\hat{r} \in \hat{\mathcal{H}}$, where *span* denotes the linear hull over the real field. The distribution is **smooth** if, to every $\hat{r} \in \hat{\mathcal{H}}$, there exist smooth vector fields $f_1, \ldots, f_n$ spanning F. A vector field is smooth if it belongs to C^∞. If F and G are distributions and $G(\hat{r}) \subseteq F(\hat{r})$ for all $\hat{r}$ we say that F **contains** G, and we write $G \subseteq F$. The **dimension** of a distribution F is defined pointwise as $dim F(\hat{r})$.

Let f and g be smooth vector fields on U. Their **Lie bracket** is the vector field

$$[f, g](\hat{r}) = f'(\hat{r})g(\hat{r}) - g'(\hat{r})f(\hat{r}),$$

where $f'(\hat{r})$ denotes the Frechet derivative of f at $\hat{r}$, and similarly for g'. We will sometimes write $f'(\hat{r})[g(\hat{r})]$ instead of $f'(\hat{r})g(\hat{r})$ to emphasize that the Frechet derivative is operating on g. A distribution F is called **involutive** if for all smooth vector fields f and g lying in F on U, their lie bracket also lies in F, i.e.

$$[f, g](\hat{r}) \in F(\hat{r}) \qquad \forall \hat{r} \in \hat{\mathcal{H}}.$$

We are now ready to define the concept of a Lie algebra which will play a central role in what follows.

Definition A.2 *Let F be a smooth distribution on $\hat{\mathcal{H}}$. The* **Lie algebra** *generated by F, denoted by $\{F\}_{LA}$ or by $\mathcal{L}\{F\}$, is defined as the minimal (under inclusion) involutive distribution containing F.*

If, for example, the distribution F is spanned by the vector fields $f_1, \ldots, f_n$ then, to construct the Lie algebra $\{f_1, \ldots, f_n\}_{LA}$, you simply form all possible brackets, and brackets of brackets, etc. of the fields $f_1, \ldots, f_n$, and adjoin these to the original distribution until the dimension of the distribution is no longer increased.

When one tries to compute a concrete Lie algebra the following observations are often very useful. Taken together they basically say that, when computing a Lie algebra, you are allowed to perform Gaussian elimination.

Lemma A.1 *Take the vector fields $f_1, \ldots, f_k$ as given. It then holds that the Lie algebra $\{f_1, \ldots, f_k\}_{LA}$ remains unchanged under the following operations.*

- *The vector field f_i may be replaced by αf_i, where α is any smooth nonzero scalar field.*
- *The vector field f_i may be replaced by*

$$f_i + \sum_{j \neq i} \alpha_j f_j,$$

where $\alpha_1, \ldots, \alpha_k$ are any smooth scalar fields.

Let F be a distribution and let $\varphi : \hat{\mathcal{H}} \to \hat{\mathcal{H}}$ be a diffeomorphism on $\hat{\mathcal{H}}$. Then we can define a new distribution $\varphi_\star F$ on $\hat{\mathcal{H}}$ by

$$(\varphi_\star F)(\varphi(\hat{r})) = \varphi'(\hat{r}) F(\hat{r}).$$

For any smooth vector field f on $\hat{\mathcal{H}}$ the field $\varphi_\star f$ is defined analogously. It is straightforward to verify that

$$\varphi_\star [f, g] = [\varphi_\star f, \varphi_\star g]. \tag{143}$$

A.3 Existence of an FDR

We can now formulate the abstract Hilbert space results concerning the existence of an FDR. There are two main results and the first one gives us he general necessary and sufficient conditions for existence.

Theorem A.1 (Björk and Svensson) *Consider the SDE in (138) and assume that the dimension of the Lie algebra $\{\hat{\mu}, \hat{\sigma}_1, \ldots, \hat{\sigma}_m\}_{LA}$ is constant near the initial point $\hat{r}^0 \in \hat{\mathcal{H}}$. Then (138) possesses an FDR if and only if*

$$dim \{\hat{\mu}, \hat{\sigma}_1, \ldots, \hat{\sigma}_m\}_{LA} < \infty$$

in a neighbourhood of $\hat{r}^0$

The second theorem gives us a parameterization of the forward rate curves produced by the model. To state this theorem we need the following definition.

Definition A.3 *Let f be a smooth vector field on $\hat{\mathcal{H}}$, and let $\hat{r}$ be a fixed point in $\mathcal{H}$. Consider the ODE*

$$\begin{cases} \dfrac{d\hat{r}_t}{dt} &= f(\hat{r}_t), \\ \hat{r}_0 &= \hat{r}. \end{cases}$$

We denote the solution $\hat{r}_t$ as $\hat{r}_t = e^{ft}\hat{r}$.

The second theorem now reads as follows.

Theorem A.2 (Björk and Svensson) *Assume that the Lie algebra $\{\hat{\mu}, \hat{\sigma}\}_{LA}$ is spanned by the smooth vector fields $\hat{f}_1, \ldots, \hat{f}_d$. Then, for the initial point r^0, all forward rate curves produced by the model will belong to the manifold $\hat{\mathcal{G}} \in \hat{\mathcal{H}}$, which can be parameterized as $\hat{\mathcal{G}} = Im[\hat{G}]$, where*

$$\hat{G}(z_1, \ldots, z_d) = e^{f_d z_d} \ldots e^{f_1 z_1} \hat{r}^0,$$

and where the operator $e^{f_i z_i}$ is given in Definition A.3

The manifold $\hat{\mathcal{G}}$ in the above theorem is obviously invariant under the forward rate dynamics. It will be therefore be referred to as the **invariant manifold** in the sequel.

References

[1] Bhar, R., and Chiarella, C. Transformation of Heath-Jarrow-Morton models to Markovian systems. *The European Journal of Finance 3* (1997), 1–26.

[2] Björk, T. *Arbitrage Theory in Continuous Time.* Oxford University Press, 1998.

[3] Björk, T. A geometric view of interest rate theory. In *Option Pricing, Interest Rates and Risk Management.*, E. Jouini, J. Cvitanic, and M. Musiela, Eds. Cambridge University Press, 2001.

[4] Björk, T., and Cristensen, B. Interest rate dynamics and consistent forward rate curves. *Mathematical Finance 9*, 4 (1999), 323–348.

[5] Björk, T., and Gombani, A. Minimal realizations of interest rate models. *Finance and Stochastics 3*, 4 (1999), 413–432.

[6] Björk, T., and Landén, C. On the construction of finite dimensional realizations for nonlinear forward rate models. Forthcoming in *Finance and Stochstics.*, 2000.

[7] Björk, T., and Svensson, L. On the existence of finite dimensional realizations for nonlinear forward rate models. *Mathematical Finance 11*, 2 (2001), 205–243.

[8] Brace, A., and Musiela, M. A multifactor Gauss Markov implementation of Heath, Jarrow, and Morton. *Mathematical Finance 4* (1994), 259–283.

[9] Carverhill, A. When is the spot rate Markovian? *Mathematical Finance 4* (1994), 305–312.

[10] Cheyette, O. Markov representation of the heath-jarrow-morton model, 1996. Preprint. BARRA Inc.

[11] Chiarella, C., and Kwon, O. K. Forward rate dependent Markovian transformations of the Heath-Jarrow-Morton term structure model. *Finance and Stochastics 5* (2001), 237–257.

[12] Cox, J., Ingersoll, J., and Ross, S. A theory of the term structure of interest rates. *Econometrica 53* (1985), 385–408.

[13] Da Prato, G., and Zabzcyk, J. *Stochastic Equations in Infinite Dimensions*. Cambridge University Press, 1992.

[14] Eberlein, E., and Raible, S. Term structure models driven by general Levy processes. *Mathematical Finance 9*, 1 (1999), 31–53.

[15] Filipović, D. A note on the nelson-siegel family. *Mathematical Finance 9*, 4 (1999), 349–359.

[16] Filipović, D. *Consistency Problems for Heath-Jarrow-Morton Interest Rate Models*. Springer Lecture Notes in Mathematics, Vol. 1760. Springer Verlag., 2001.

[17] Filipović, D., and Teichmann, J. Existence of finite dimensional realizations for stochastic equations, 2001. Forthcoming in *J. Funct. Anal.*

[18] Filipović, D., and Teichmann, J. On finite dimensional term structure models, 2002. Working paper.

[19] Heath, D., Jarrow, R., and Morton, A. Bond pricing and the term structure of interest rates. *Econometrica 60*, 1 (1992), 77–106.

[20] Heath, D., Jarrow, R., and Morton, A. Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. *Econometrica 60* (1992), 77–105.

[21] Inui, K., and Kijima, M. A Markovian framework in multi-factor Heath-Jarrow-Morton models. *Journal of Financial and Quantitative Analysis 33* (1998), 423–440.

[22] Jeffrey, A. Single factor Heath-Jarrow-Morton term structure models based on Markov spot interest rate dynamics. *Journal of Financial and Quantitative Analysis 30* (1995), 619–642.

[23] Musiela, M. Stochastic PDE:s and term structure models. Preprint, 1993.

[24] Ritchken, P., and Sankarasubramanian, L. Volatility structures of forward rates and the dynamics of the term structure. *Mathematical Finance 5*, 1 (1995), 55–72.