DISS. ETH NO. 23359

VARIANCE CURVE MODELS: FINITE DIMENSIONAL REALIZATIONS AND BEYOND

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH ZURICH)

presented by

OZAN BARIŞ AKDOĞAN

Dipl. Ing., TU-Berlin
born on 08.01.1979
citizen of Germany

accepted on the recommendation of

Prof. Dr. Josef Teichmann examiner
Prof. Dr. Walter Farkas co-examiner

2016

To my sister and parents, for all their love and support.

Abstract

This thesis is devoted to the study of evolution equations for the term structure of forward variances, so-called *variance curves*. As done before in [15], it will be shown that under the assumption of no arbitrage and some mild regularity conditions, these processes can be understood as the (mild) solutions of certain stochastic partial differential equations satisfying the *HJM drift condition*, which for these models means the absence of a drift. In this generality, such solutions are infinite dimensional Markov processes and not very tractable for applications such as simulation or pricing. Starting from those equations, the aim of the thesis is twofold:

First, utilizing the Frobenius theory introduced in [40], conditions on the diffusion vector fields will be formulated such that the solutions can be parameterized by a finite dimensional diffusion process. In this case the solution is said to admit a (generic) *finite dimensional realization*. The analysis will be restricted to Markovian systems comprised of forward variance models with stochastic volatility of forward variance. Within this class of models, diffusion vector fields will be considered that correspond to *affine* and *exponentially affine* realizations. This part is similar to [40] and [12], where a similar analysis is done for forward interest rate processes and to [15], where for variance curve models (without stochastic volatility of forward variance) the same problem was investigated but from the perspective of *consistent factor models*, in which *given* a parameterization and a finite dimensional diffusion process conditions were derived such that these correspond to a finite dimensional realization. Thus, this part can be seen as an extension and completion of the corresponding part in [15]. At the end of this part, the relation to the recently introduced *fractional Bergomi* model (cf. [5]) will be highlighted and a finite dimensional approximation scheme suggested.

Second, motivated by the recent trend of looking at term structure models that do not admit (generic) finite dimensional realizations but are inherently infinte dimensional, a weak approximation scheme based on a Malliavin-Taylor expansion introduced in [68] will be applied to variance curve models. This expansion rests on iterative applications of Malliavin's integration by parts formula and as such requires the smoothness of the underlying density. For finite dimensional diffusion processes this can be checked with the well-known Hörmander condition (cf. [63]). For infinite dimensional evolution equations this however is essentially an open problem (cf. [47] for conditions for evolution equations with additive noise and [4] for conditions on the existence (but not necessarily smoothness) of a density that is absolutely continuous with respect to the Lebesgue measure of the projected solutions of more general equations). A class of equations will be identified that correspond to generalized versions of those that admit generic finite dimensional realizations, for which precise conditions can be given for the smoothness of the density of the process projected to a finite dimensional subspace. Furthermore we conjecture that this property holds as well, for a class of processes that

are inherently infinite dimensional. This class of processes are very similar to the *Consistent Re-Calibration* (CRC) processes introduced in [66] for forward interest rates and thus the name will be adopted. Given the existence of this expansion, the corresponding *variance-optimal Malliavin weights* will be derived in some cases by perturbation of the Kolmogorov PDE as was done in finite dimensions in [70] and in infinite dimensions in [9] albeit without the proof of existence of the expansion. With the representation of the *variance-optimal Malliavin weights* two applications will be considered, namely the pricing of put options on the log-price (which has now stochastic forward variance and stochastic volatility of forward variance) and options on the volatility swap rate (which correspond to options on the VIX).

Kurzfassung

Das Thema dieser Dissertation sind Evolutionsgleichungen für die Terminstruktur von Termin Varianzen, sogenannten *Varianzkurven*. Wie zuvor in [15], wird gezeigt das unter der Annahme von Arbitragefreiheit und einigen milden Regularitätsbedingungen, diese Prozesse als (milde) Lösungen von gewissen stochastischen partiellen Differentialgleichungen verstanden werden können, welche die *HJM Drift Bedingung* erfüllen, also vorliegend die Abwesenheit eines Driftes. In dieser Allgemeinheit sind Lösungen solcher Gleichungen unendlichdimensionale Markov Prozesse und damit nicht sehr nützlich für Anwendungen wie Simulationen oder Bewertungen. Ausgehend von diesen Gleichungen, werden in dieser Dissertation zwei Ziele verfolgt.

Als erstes werden unter Benutzung der in [40] eingeführten Frobenius Theorie Bedingungen on die Diffusionsvektorfelder formuliert, so dass die zugehörigen Lösungen duch einen endlichdimensionalen Diffusionsprozess parametrisiert werden können. In solchen Fällen wird gesagt, dass die Lösung eine (generische) endlichdimensionale Realisierung zulässt. Die Analyse wird beschränkt auf Markovsche Systeme bestehend aus einem Termin Varianz modell mit stochastischer Volatilität der Termin Varianz. Innerhalb dieser Modell- klasse werden solche Diffusionsvektorfelder betrachtet, die zu *affinen* und *exponentiell-affinen* Realisierungen gehören. Dieser Teil ist vergleichbar zu [40] und [12], in welchen ähnliche Analysen für Termin Zinssätze gemacht werden und zu [15], wo für Varianzkurven (ohne stochastischer Volatilität) dasselbe Problem untersucht wird, allerdings aus dem Blickwinkel der *konsistenten Faktormodelle*. Hierbei werden für eine *gegebene* Parametrisierung und endlichdimensionalen Diffusionsprozess Bedingungen formuliert so dass diese einer endlichdimensionalen Realisierung entsprechen. Dieser Teil kann also als eine Erweiterung und Vervollständigung des entsprechenden Teils in [15] angesehen werden. Dieser Teil endet mit der Darstellung der Verbindung dieser Modelle zu dem kürzlich eingeführten *fractional Bergomi* Modells (cf. [5]) und der Einführung eines endlichdimensionalen Approximationsschemas.

Als zweites, motiviert von dem gegenwärtigen Trend der Betrachtung von Terminstruktur Modellen die keine endlichdimensionalen Realisierungen zulassen und inhärent unendlichdimensional sind, wird ein schwaches Approximationsschema, dass auf einer Malliavin-Taylor Expansion basiert und in [68] vorgestellt wurde, untersucht und auf Varianzkurven angewendet. Diese Expansion basiert auf Malliavins Partiellen Integrationsformel und benötigt daher die Glattheit der zugehörigen Dichte. Bei endlichdimensionalen Diffusionsprozessen lässt sich dies mit der bekannten Hörmander Bedingung überprüfen (cf. [63]). Für unendlichdimensionale Evolutionsgleichungen ist dies allerdings im wesentlichen ein offenes Problem (cf. [47] für Bedingungen für Evolutionsgleichungen mit additivem Rauschen und für allgemeinere Gleichungen [4] für Bedingungen für die Existenz (aber nicht notwendigerweise Glattheit) einer Dichte der endlichdimensional projizierten Lösung, die absolutstetig bezüglich des Lebesgue

Masses ist). Es wird eine Klasse von Prozessen identifiziert die verallgemeinerten Versionen von Prozessen entprechen, welche endlichdimensionale Realisierungen zulassen. Für diese Prozesse werden präzise Bedingungen für die Existenz einer glatten Dichte (des endlichdimensional projiezierten Prozesses) formuliert.Darüberhinaus wird eine Vermutung formuliert für eine Klasse von Prozessen die inherent unendlich dimensional sind. Diese Prozesse sind sehr ähnlich zu den in [66] für Termin Zinssätze eingeführten *Consistent Re-Calibration* (CRC) Prozessen und daher wird dieser Name übernommen. Unter der Annahme der Existenz dieser Expansion, werden die zugehörigen *varianzoptimalen Malliavin Gewichte* in einigen Fällen mittels Perturbation der Kolmogorov PDG gewonnen. Dies ist vergleichbar zu [70], in welchem endlichdimensionale Diffusionen betrachtet wurden, und zu [9], in welchem auch undendlichdimensionale Diffusionen betrachtet wurden, aller- dings ohne einen Beweis der Existenz der Expansion. Mit der gewonnen Darstellung der varianzoptimalen Malliavin Gewichte werden zwei Anwendungen betrachtet, nämlich die Bewertung von Put Optionen auf den log-Preis (der in diesem Fall stochastische Termin Varianz und stochastische Volatilität der Termin Varianz hat) und Optionen auf den Volatilitäts-Swapsatz (dies entspricht einer Option auf den VIX).

Acknowledgements

First of all, I would like to express my deep gratitude to my supervisor Prof. Josef Teichmann. With his expertise on the one hand and encouragement, patience and kindness on the other hand, he made the completion of my thesis an enjoyable task. I appreciate not only his vast knowledge in so many different areas of Mathematics but also the intuitive way of teaching which made many complicated areas of my thesis easily accessible. Finally, the many social activities within his working group, which he joined - if not organized - many times, created a very friendly and social working atmosphere which without doubt added to the quality of my research.

I am also thankful to Prof. Walter Farkas, for kindly accepting to act as the co-examiner of my thesis and his understanding and flexibility during the last weeks of the completion.

Thanks also go to the previous and current members of "Group 3", for interesting discussions and enjoyable times, with a special thank to my several room mates: Stefan Tappe, Anja Richter, Blanka Horvarth and Philipp Harms for great times, discussions and fun.

I am also thankful to the Credit Suisse for accepting me in the Academic Research program that provided the opportunity of starting this PhD project in a part-time role and included a scholarship.

Finally, I would like to thank to Prof. Alex Schied for raising my interest in Mathematical Finance at TU-Berlin and his encouragement and support during my Diploma Thesis.

Contents

I Introduction **1**

II HJM Models for forward variances **7**
- II.1 Variance Swaps . . . 7
- II.2 Idealized Variance Swaps . . . 8
- II.3 Stochastic Volatility Models . . . 10
- II.4 HJM Theory . . . 13
- II.5 HJM Musiela Theory . . . 17

III Finite dimensional Realizations **27**
- III.1 Markovian HJMM Models . . . 27
- III.2 The Frobenius Theorem . . . 28
- III.3 Affine realizations . . . 30
- III.4 Exponentially-Affine Realizations . . . 42
- III.5 Mixed Models . . . 52
- III.6 Relation to the Fractional Bergomi Model . . . 59

IV Weak Taylor Expansions for SPDEs **63**
- IV.1 Introduction . . . 63
- IV.2 The Malliavin Derivative . . . 63
- IV.3 Asymptotic Expansions . . . 77

V Consistent Recalibration Models **79**
- V.1 Introduction . . . 79
- V.2 CRC Forward Variance Models . . . 82
- V.3 Hypoellipticity . . . 87

VI Weak Taylor PDE Expansions **93**
- VI.1 Introduction . . . 93
- VI.2 Implied Volatility Expansion . . . 94
- VI.3 Term Structure Expansion . . . 105

A Stochastic Processes in Hilbert Spaces **111**
- I.1 The Itô Integral . . . 111
- I.2 Itô's Formula . . . 112
- I.3 Stochastic Equations . . . 113

Bibliography **117**

Chapter I

Introduction

The assumption of constant (instantaneous) volatility in the celebrated Black & Scholes model (see [14]) was soon rejected as it is not consistent with the observed *volatility smile* in the markets of plain vanilla options. This was the advent of so-called *stochastic volatility models* (as given for example by the Heston model [52]) on the one hand and volatility derivatives (see for example [20] for an overview) on the other hand. Among the simplest volatility derivatives written on a stock-index is the *variance swap.* It pays at maturity the *variance notional* times the difference between the *realized variance* of that stock-index and the fixed *variance swap rate.* Its popularity for investors stems from the possibility to have a direct exposure to realized variance (see also [3, Part 3]) and for dealers from the fact that it can be replicated in an essentially model free way using traded plain vanilla options (see for example [22] and [21]). In May 2014 the newly issued variance swaps on the S&P 500 had a *vega notional* (see Section II.1) of 85.2 million USD ([61, Table 1]) and the total gross outstanding vega notional was 1.512 billion USD ([61, Table 2]). Although these contracts are traded over the counter (OTC) its specifications are highly standardized. Major dealers publish on a daily basis indicative variance swap rates for a fixed tenor of *time-to-maturities* ranging (usually) from one month to five years. This gives rise to a daily observed term structure of variance swap rates and due to the relationship to the plain vanilla options through the replication formula this term structure can be seen also on a much finer time grid. Typically for long maturities this term structure is pretty stable while the shorter end is strongly related to the asset's instantaneous variance and accordingly the term structure is upward sloping in low volatility regimes and downward sloping in high volatility regimes. The prime theme of this thesis will be to discuss models for the evolution of this term structure that are both, analytically tractable and rich in the sense that empirically observed static and dynamic features of this term structure are well reflected.

Consistent Variance Curve Models

Instead of modeling the term structure of the variance swap rates directly, we will look at models for the evolution of the term structure of *forward variances*, which corresponds to the instantaneous variance swap rates (see Definition II.2.8). The term structure of the forward variances will be called *variance curve.* Looking at forward variances instead of variance swap rates is standard and adopted for example in [7], [15] in the context of HJM models (that will be discussed in the following) and in [36], [2], [31] in the context of models for the market price of variance risk. The short-end of this curve corresponds to the (instantaneous) variance (henceforth *spot-variance*) of the underlying asset's returns. On the other hand, under the martingale measure, the entire variance curve is determined from the spot-variance process. Using this relationship, we will see that the variance curves implied from popular stochastic volatility models can take only very limited shapes and in particular will (generally) not fit the variance curves observed in the markets. Therefore we will adopt the HJM Musiela framework for variance curves introduced in [15] in which the variance curve is modeled directly, as opposed to the case where it is derived from the asset's stochastic volatility process. At each instant of time, the variance curve is understood as a function mapping the *time-to-maturity* to the corresponding forward variance and the variance curve will be assumed to be given as the mild solution of a stochastic partial differential equation (SPDE), that is subject to the *HJM-drift condition* (see Proposition II.4.4), in a suitable (usually infinite dimensional) Hilbert space H of variance curves. That is, the variance curve at time t, denoted by u_t, will be given as the solution of

$$\begin{cases} du_t &= \frac{d}{dx} u_t \, dt + \sum_{i=1}^d \Sigma_i(u_t) \, d\beta_t^i \\ u_0 &\in H, \end{cases} \tag{0.1}$$

where the operator $\frac{d}{dx}$ generates the strongly continuous shift-semigroup $\{S_t \,|\, t \geq 0\}$, $\Sigma_1, ..., \Sigma_d$ are smooth vector fields and $\beta^1, ..., \beta^d$ are independent standard Brownian motions, all defined on a suitable filtered probability space. Such solutions will generally correspond to infinite dimensional Markov processes in the spirit of [26] and the spot-variance process will be derived from the short-end condition. This specification is obviously very general but lack the analytic tractability required for real-life applications. Therefore, as a compromise between one dimensional diffusion models for the spot-variance and infinite dimensional diffusion models for the variance curves, the main focus in [15] (see also [32] for the corresponding situation in interest rates) was in the construction of finite dimensional factor models. The investigated problem was the following: given a variance curve model u as the solution of (0.1), an m-dimensional diffusion process Z (with $m \in \mathbb{N}$) and a smooth map $G : \mathbb{R}^m \to H$, what are the conditions such that $u_t = G(Z_t), \quad t \geq 0$, holds true? This problem is (completely) solved (see also [37] for a more general version) and under its conditions the resulting variance curve process is now completely determined by a finite dimensional diffusion process. In this situation (subject to some mild regularity conditions), the solution u stays in the submanifold $\mathcal{M} := G(\mathbb{R}^m)$ of H.

Generic Finite Dimensional Realizations

In the consistent variance curve models problem, the parameterization G of $\mathcal{M}$ must be given a-priori. Also, the possible set of initial curves lies necessarily in the finite dimensional submanifold $\mathcal{M}$ as well. By employing the Frobenius theory introduced in [40], we look at the following problem. Given (u, Y) as the solution of

$$\begin{cases} du_t &= \frac{d}{dx} u_t \, dt + \sum_{i=1}^d \Sigma_i(u_t, Y_t) \, d\beta_t^i \\ dY_t &= a(Y_t) \, dt + \sum_{i=1}^d b_i(Y_t) \, d\beta_t^i \\ (u_0, Y_0) &\in H \times \mathbb{R}^n, \end{cases} \tag{0.2}$$

under which conditions on $\Sigma_1, ..., \Sigma_d$, $a, b_1, ..., b_d$, can we find for every $u_0 \in U$, where U is some open subset of H, a parameterization G_{u_0} and a finite dimensional factor process Z such that

$$(u_t, Y_t) = G_{u_0}(Z_t), \quad t \geq 0, \tag{0.3}$$

holds true? Notice that in this case u_0 can take values in an open subset of H, whereas with the consistent factor-models the initial curve must lie in the a-priori chosen submanifold $\mathcal{M}$ (in [71] conditions for the existence of affine factor models are investigated that share this latter property and can thus be seen as lying conceptually in between both approaches).

We will look here at two kinds of models. The first kind will represent a straight forward generalization of (affine) spot-variance models in which the vector fields $\Sigma_1, ..., \Sigma_d$ will be of the constant direction type (see [10]). The second kind will correspond to a generalization of Bergomi's model (see [7]) in that the vector fields $\Sigma_1, ..., \Sigma_d$ will be linear in the first argument. These models too will after a simple transformation be of the first kind. We will find that these two kinds will correspond to (generic versions) of affine and exponentially affine parameterizations. We will also look at combinations of both models and discuss the connection of the latter models to the recently introduced *fractional Bergomi model* (see [44]).

Weak Taylor Expansions

Regarding the usefulness of consistent factor models in the context of variance curve models, Hans Bühler writes in [15, Page 11] (which was in December 2008):

> The use of consistent models is justified if the market is not very liquid and a mathematically sound interpolation scheme plus consistent dynamics is required for risk management. However, by now variance swap quotes on major indices are so deep that this approach is no longer justified: the priority must now be to fit the market quotes of variance swaps first and then superimpose them with consistent dynamics.

Accordingly he proposes in [15, Chapter 4] to extend a spot-variance process given by a finite dimensional diffusion process by taking the entire observed forward variance curve as an additional *parameter* to arrive at an exact fit (this situation is comparable to Hull-White extensions to be discussed in Section II.3.1). Indeed, given a real-valued process $\tilde{\xi}$ as a functional of a (possibly time-inhomogeneous) diffusion process and the observed variance curve $\hat{v}_0$ he proposes to consider the *fitted* spot-variance model given by

$$\xi_t := m(t)\tilde{\xi}_t, \quad m(t) := \frac{\hat{v}_0(t)}{\mathbb{E}[\tilde{\xi}_t]}$$

as in this case the model provides a perfect fit to the observed term structure of variance swap rates. He further shows that in certain examples the model remains analytically tractable. A problem with such models is that the spot-variance process now depends implicitly on the inception date of the variance swap contract. For variance swap contracts with overlapping life-times this might lead to inconsistencies. Further it is not clear how to simulate time-series of the variance curves.

In a very similar context (but for interest rates) in [66] and [50] a solution to this problem is provided which leads to the *consistent re-calibration* (CRC) models. Such models usually do not admit generic finite dimensional realizations but are inherently infinite-dimensional (from the Markovian perspective) and allow for the re-calibration of variance curves that can be chosen from an open subset of an infinite dimensional

Hilbert space. A key property of this models is that they correspond infinitesimally to models that admit generic (affine) finite dimensional realizations. By a suitable discretization, which corresponds to an exponential splitting scheme (see [48]), the model allows for a very tractable simulation algorithm that utilizes the infinitesimally (affine) finite dimensional structure.

Unfortunately, with this discretization scheme the order of weak convergence is quite low and it is in general not possible to utilize possibly existing closed form solutions for derivatives prices. In this thesis we will in the context of variance curve models provide a small-diffusion (or small-parameter) expansion of arbitrary order of weak convergence such that in many interesting cases this will be possible or at least will allow for an efficient Monte-Carlo pricing scheme. In particular, this expansions will allow to essentially compute expected values of functions of the variance curves in terms of the corresponding quantities of a variance curve model that admits generic finite dimensional realizations. We will look at parameterized systems of the form

$$\begin{cases} dX_t^\epsilon & = (AX_t^\epsilon + V(\epsilon, X_t^\epsilon))\,dt + \sum_{i=1}^d V_i(\epsilon, X_t^\epsilon)\,d\beta_t^i \\ X_0 & \in H, \end{cases} \tag{0.4}$$

where for each $\epsilon \in \mathbb{R}$, $V(\epsilon,\cdot), V_1(\epsilon,\cdot), ..., V_d(\epsilon,\cdot)$ are sufficiently regular vector-fields on H and assume that X^0 admits a generic finite dimensional realization. Than for suitable linear functionals $l : H \to \mathbb{R}^m$, we will look at conditions such that

$$\left|\mathbb{E}\Big[f(l \circ X_T^\epsilon)\Big] - \sum_{i=0}^n \frac{\epsilon^i}{i!}\frac{\partial^i}{\partial \epsilon^i}\Big|_{\epsilon=0}\mathbb{E}\Big[f(l \circ X_T^\epsilon)\Big]\right| = o(\epsilon^n), \quad \text{as } \epsilon \to 0, \tag{0.5}$$

for functions $f : \mathbb{R}^m \to \mathbb{R}^n$, $n \geq 1$, that are bounded and measurable. The vehicle to achieve this will be the *weak Taylor expansion* introduced in [68] in the context of finite dimensional stochastic differential equations. This expansions rests on iterative applications of Malliavin's *integration by parts* formula and a crucial requirement is the existence of a smooth density of the finite dimensional random variable $l \circ X_T^\epsilon$, which is usually called *Hypoellipticity in infinite dimensions* (cf. [47, Question 1] and [4]). We will recapture the necessary tools from Malliavin calculus required in the context of SPDEs and show that for the system (0.2), if it admits a finite dimensional realization, under suitable conditions on the parameter process Y, this expansion will be possible with weak convergence of arbitrary order. Here the unperturbed system (i.e. $\epsilon = 0$ in (0.4)) will correspond to the case where Y is deterministic and constant. The situation where (0.2) admits a generic finite dimensional realization only for the case where Y is deterministic and constant will correspond to the CRC models. We shall conjecture conditions such that the expansion holds.

Having this conjectured conditions such that (0.5) holds, we still need to compute the *Malliavin weights* π_i, $i = 1, ..., n$, that satisfy

$$\frac{\partial^i}{\partial \epsilon^i}\Big|_{\epsilon=0}\mathbb{E}\Big[f(l \circ X_T^\epsilon)\Big] = \mathbb{E}\Big[f(l \circ X_T^0)\pi_i\Big], \quad i = 1, ..., n.$$

This weights can be computed in terms of Skohorod integrals (using iteratively Malliavin's integration by parts) and the corresponding representations can significantly reduce the complexity of the given problem. In particular the terms $\mathbb{E}\Big[f(l \circ X_T^0)\pi_i\Big]$ can be computed with plain vanilla Monte-Carlo integration using the (Markovian) finite-dimensional structure of X_T^0, whereas X_T^ϵ for $\epsilon > 0$ is in general a strictly infinite-dimensional Markov process where standard Monte-Carlo techniques can not be employed.

However, in general this weights will not be variance optimal (cf. [42]) in the sense that $\pi_j \neq \mathbb{E}[\pi_j | f(l \circ X_T^0)]$. Inspired by [70] and using a certain martingality conditions given in [9] we will in some cases be able to derive explicit formulas for the terms $\frac{\partial^i}{\partial \epsilon^i}\Big|_{\epsilon=0} \mathbb{E}\Big[f(l \circ X_T^\epsilon)\Big]$ and to compute the variance optimal weights by expanding the solution of the corresponding Kolmogorov PDE (which is justified if (0.5) holds, but also under weaker conditions). If this martingality condition does not hold, this PDE perturbation approach leads to recursive PDEs for the terms $\frac{\partial^i}{\partial \epsilon^i}\Big|_{\epsilon=0} \mathbb{E}\Big[f(l \circ X_T^\epsilon)\Big]$ which are nice but would require nested Monte-Carlo simulations and are thus less tractable when compared to the representations given by the Malliavin weights.

Applications

We will look at the system

$$\begin{cases} dX_t^\epsilon = -\frac{1}{2} u_t^\epsilon(0)\, dt + \sqrt{u_t^\epsilon(0)}\, d\beta_t^1 \\ du_t^\epsilon = \frac{d}{dx} u_t^\epsilon\, dt + \epsilon \sum_{i=1}^d \sigma_i(u_t^\epsilon, Y_t^\epsilon)\, d\beta_t^i \\ dY_t^\epsilon = \epsilon^2\, c_0(Y_t^\epsilon)\, dt + \epsilon \sum_{i=1}^d c_i(Y_t^\epsilon)\, d\beta_t^i, \\ (X_0^\epsilon, u_0^\epsilon, Y_0^\epsilon) = (x, u, y) \in \mathbb{R} \times H \times \mathbb{R}^m, \end{cases} \tag{0.6}$$

where X^ϵ, u^ϵ and Y^ϵ correspond to the processes for the log-price, the forward variance and the stochastic volatility of forward variance and apply the weak Taylor PDE expansion on the log-price and functionals of the forward variance. In both cases we will derive explicit formulas that can be used for pricing options on the log-price and functionals of the forward variance (a functional of the forward variance is given for example by the VIX). In the first case, the formula consists of sums of products of partial derivatives of the Black & Scholes price with deterministic functions and in the second case of sums of products of derivatives of the payoff with deterministic functions. The first expansion is very similar to the situation considered in [69] where (finite dimensional) stochastic volatility models are considered and [9] where a very similar variance curve setting is considered, although in a parameterization, where the variance curves are given as solutions to infinite dimensional stochastic differential equations with a time-varying state-space.

Outline of the Thesis

The remainder of this thesis can be divided into three parts. In the first part, consisting of Chapters II and III, we recapture the notion of variance curves in the context of *idealized variance swaps* and give necessary and sufficient conditions for the existence of generic finite dimensional realizations for certain variance curve models. In the second part, consisting of Chapters IV and V, we formulate conditions such that the weak Taylor expansion can be applied to functionals of variance curves. For this we recapture the necessary Malliavin calculus for Hilbert space valued stochastic processes and derive conditions related to Hypoellipticity in infinite dimensions. In the second part, we look at Consistent Re-calibration Models for variance curves and conjecture sufficient conditions for the applicability of the weak Taylor expansion. The final part is given in Chapter VI, where Applications of the weak Taylor PDE expansion are discussed.

Notations

For the notation we refer to the Appendix A

Chapter II

HJM Models for forward variances

1 Variance Swaps

Let S denote the price process of a traded stock and by $X := \log(S)$ the corresponding log-price process. A variance swap with variance strike K^2 initiated at time $t = 0$ with maturity time $T \geq 0$ pays out the variance notional N_{var} times the difference between the annualized discrete quadratic variation of the log-price X over pre-defined business days $0 = t_0 < ... < t_n = T$ and the strike K^2. Formally the payoff at time T, denoted by $VS_K(T,T)$, is

$$VS_K(T,T) := N_{var}\Big(\frac{d}{n}\sum_{i=1}^{n}(X_{t_i} - X_{t_{i-1}})^2 - K^2\Big), \tag{1.1}$$

where d is chosen such that $d/n \approx 1/T$. If the logarithmic returns $(X_{t_i} - X_{t_{i-1}})$, $i = 1, ..., n$, are assumed to have a zero mean, then the payoff is just the difference between the annualized *realized variance* of the logarithmic returns and the strike K^2. As practitioners like to think in volatility terms usually the payoff is specified in terms of vega-notional N_{vega} instead of the variance notional N_{var} in which case the payoff is given by

$$VS_K(T,T) = N_{vega}\Big(\frac{\frac{d}{n}\sum_{i=1}^{n}(X_{t_i} - X_{t_{i-1}})^2 - K^2}{2K}\Big). \tag{1.2}$$

While the variance notional corresponds to the cash profit/loss if at maturity the annualized realized variance exceeds the variance strike by 1 variance unit, the vega notional corresponds to the average of the (absolute values) of the payoffs if the difference is one volatility unit and is justified by the relationship

$$\Big(\frac{(K+x)^2 - K^2}{4K}\Big) - \Big(\frac{(K-x)^2 - K^2}{4K}\Big) = x, \quad \text{for all } x \geq 0. \tag{1.3}$$

In the following we denote the price of the variance swap at time t, with $0 \leq t \leq T$, by $VS_K(t,T)$ and if $K = 0$ we just write $VS(t,T)$. We will mostly assume that $N_{var} = 1$, $K = 0$ and skip the annualization.

2 Idealized Variance Swaps

It has been proven to be fruitful to assume that the payoff of a variance swap is given by the quadratic variation $[X, X]$ of the log-price process X accumulated from inception to maturity. For this we assume that we are given a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P})$, satisfying the usual assumptions, such that on this space the log-price process X satisfies the following conditions. Recall the notion of semimartingale characteristics given in [53, Definition 2.6] and the notion of differential characteristics, that is, semimartingale characteristics that are absolutely continuous (with respect to the Lebesgue measure) given in [55, Section 2].

Assumption 2.1. *The log-price X satisfies both,*

- *it is a real-valued, locally square integrable semimartingale, and*
- *its semimartingale characteristics are absolutely continuous with respect to the Lebesgue measure.*

Remark 2.2. In this case the predictable quadratic variation $\langle X, X\rangle$ is absolutely continuous and there is a predictable process $\partial_t\langle X, X\rangle$ that is unique up to a $d\mathbb{P} \otimes dt$-null set such that

$$\langle X, X\rangle_t = \int_0^t \partial_t\langle X, X\rangle_s \, ds, \quad \forall t \geq 0.$$

Moreover, $[X, X] - \langle X, X\rangle$ is a local martingale and a true martingale whenever X is square integrable.

Proof. This follows from [55, Section 2] and [53, Proposition 4.5]. □

This class covers almost all continuous time models that are considered in mathematical finance. We denote again by $VS(t, T)$ the price of a variance swap at time t issued at time 0 with maturity at T. We also consider another related (but artificial) product, namely the *predictable variance swap* which differs from the usual variance swap in that it pays out the predictable quadratic variation with price at time t denoted by $VS^P(t, T)$. This is summarized in the following Definition.

Definition 2.3. *The* payoff *of the idealized (predictable) variance swap at maturity is given by the (predictable) quadratic variation at T, that is, for $T \geq 0$ fixed but arbitrary, we have*

$$VS(T, T) := [X, X]_T, \quad \textit{and} \quad VS^P(T, T) := \langle X, X\rangle_T.$$

For each fixed maturity T, the price-process *of the Variance swap is given by the process*

$$[0, T] \ni t \mapsto VS(t, T)$$

and at each time t, the term structure of Variance swap prices *is given by the curve*

$$[t, \infty) \ni T \mapsto VS(t, T).$$

The terminology is also applied for predictable variance swaps.

Remark 2.4. To assume that the variance swap pays out the quadratic variation instead of the discrete quadratic variation as in (1.1) is standard (see for example [15] and [54] for discussions of this aspect) and motivated by the higher tractability in continuous time models.

As the predictable variance swap is an artifical instrument, we are free to define its price process. The following will be very convenient.

Definition 2.5. *The price process of the predictable variance swap is given by*

$$VS^P(t,T) := VS(t,T) + \langle X,X\rangle_t - [X,X]_t, \quad t \in [0,T]. \tag{2.1}$$

In accordance with Remark 2.2 we arrive at the following Proposition.

Proposition 2.6. *The variance swap price process is a (local) martingale, if and only if, the predictable variance swap is a (local) martingale.*

This Proposition will be very convenient in the following, as it will be crucial for the HJM-Musiela Theory to assume that the term structure of variance swap prices, that is, $T \mapsto VS(t,T)$ is absolutely continuous with respect to the Lebesgue measure, which would immediately rule out any non-continuous log-price processes X. But as it is sufficient to consider the predictable variance swap price processes a much larger class of log-price processes X is possible, namely those satisfying the conditions of Assumption 2.1. Now assuming that $\mathbb{P}$ is a martingale measure for the market of variance swaps, that is, the price processes of the variance swaps (and due to Proposition 2.6 also the predictable variance swaps) are martingales under $\mathbb{P}$. In this case it follows from [28] and Remark 2.2 that for $t \in [0,T]$

$$VS(t,T) := \mathbb{E}[[X,X]_T|\mathcal{F}_t] = \mathbb{E}[\langle X,X\rangle_T|\mathcal{F}_t] + [X,X]_t - \langle X,X\rangle_t \tag{2.2}$$

and together with (2.1) that

$$\begin{aligned} VS^P(t,T) &= \mathbb{E}[\langle X,X\rangle_T|\mathcal{F}_t] = \langle X,X\rangle_t + \mathbb{E}[\langle X,X\rangle_T - \langle X,X\rangle_t|\mathcal{F}_t] \\ &= \int_0^t \partial_s\langle X,X\rangle_s\,ds + \int_t^T \mathbb{E}[\partial_s\langle X,X\rangle_s|\mathcal{F}_t]\,ds. \end{aligned} \tag{2.3}$$

Finally we introduce the predictable *variance swap rate* $V^P(t,T)$ defined as the second term in (2.3), that is,

$$V^P(t,T) := \mathbb{E}[\langle X,X\rangle_T - \langle X,X\rangle_t|\mathcal{F}_t] = \int_t^T \mathbb{E}[\partial_s\langle X,X\rangle_s|\mathcal{F}_t]\,ds, \tag{2.4}$$

which corresponds to the price of a zero strike predictable variance swap at inception time t and maturity at time T. Let $V(t,T)$ denote the corresponding variance swap rate.

Proposition 2.7. *We have* $V(t,T) = V^P(t,T)$ *for all* t,T *with* $0 \le t \le T$.

Proof. This again follows from Remark 2.2, as

$$V(t,T) := \mathbb{E}[[X,X]_T - [X,X]_t|\mathcal{F}_t] = \mathbb{E}[\langle X,X\rangle_T - \langle X,X\rangle_t|\mathcal{F}_t] = V^P(t,T).$$

□

Thus we see that the variance swap rate is actually the same for the variance swap and the predictable variance swap and at each time t prior to maturity T we have the well-known additive decompositions of the prices given by

$$VS^P(t,T) = \langle X,X\rangle_t + V(t,T), \quad VS(t,T) = [X,X]_t + V(t,T),$$

such that the price of the (predictable) variance swap at time t is the sum of the (predictable) quadratic variation up to time t and the variance swap rate V(t,T), explaining why variance swaps are usually quoted by the variance swap rate. As the term structure of variance swap rates given by (2.4) is absolutely continuous we can introduce the *spot-variance process* and the *forward-variance process* (see also Remark 2.2) as follows.

Definition 2.8. *The spot-variance process is given by* $\xi_t := \partial_t \langle X, X\rangle_t$*, for* $t \geq 0$*, and the forward variance process* $(v(t,T))_{t\in[0,T]}$ *by*

$$v(t,T) := \partial_T V(t,T) = \mathbb{E}[\xi_T|\mathcal{F}_t], \quad t \in [0,T]. \tag{2.5}$$

Notice the similarity of these quantities to the short rate and forward rate in the interest rates theory (see for example [34]). In particular the short-end condition holds true, that is,

$$\xi_t = v(t,t), \quad \forall t \geq 0. \tag{2.6}$$

Finally we arrive at the representation of the predictable variance swaps and variance swap rates in terms of the spot variance process and the forward variance process given by

$$VS^P(t,T) = \int_0^t \xi_s \, ds + \int_t^T v(t,s)\, ds, \tag{2.7}$$

$$V(t,T) = \int_t^T v(t,s)\, ds. \tag{2.8}$$

3 Stochastic Volatility Models

From Representations (2.5) and (2.7) we see that the prices of predictable variance swaps are fully determined by the spot variance process ξ corresponding to the log-price process X. If the spot variance process is assumed to be a non-negative semimartingale, the pair $(X, \sqrt{\xi})$ is usually called *stochastic volatility model.*

Definition 3.1. *Given a log-price* X *as in Assumption 2.1, a* stochastic volatility model *is called any pair* (X, η) *such that* η *is a locally square integrable semimartingale satisfying* $\partial_t\langle X, X\rangle_t = \eta_{t-}^2$ *for all* $t \geq 0$*. In this case, the process* η *is called the stochastic volatility process corresponding to* X*.*

We need also the following definition, which corresponds to the case where only the spot-variance process is explicitly given. Here a corresponding log-price can be constructed, which in general will not be uniquely determined but has to satisfy a certain *compatibility* condition.

Definition 3.2. *Given a locally square integrable semimartingale* η*, a semimartingale* X *will be called* compatible log-price model for the stochastic volatility η*, if* (X, η) *is a stochastic volatility model in the sense of Definition 3.1.*

Remark 3.3. Note that as any semimartingale defined as in [53, Definition 4.21] is adapted and càdlàg, the process $t \mapsto \eta_{t-}$ is adapted and càglàd and in particular predictable (see [53, Proposition 2.6]). Accordingly it qualifies as local characteristic, see Remark 2.2. Moreover, as the processes $(\eta_t)_{t\geq 0}$ and $(\eta_{t-})_{t\geq 0}$ only disagree on a Lebesgue null-set, we have

$$\int_0^t \eta_s \, ds = \int_0^t \eta_{s-} \, ds, \quad \forall t \geq 0$$

and so it is only a minor abuse of notation to set $\partial_t \langle X, X \rangle_t = \eta_t^2$ for all $t \geq 0$. This is justified by the fact that we are ultimately interested in the price process of predictable variance swaps as given in (2.7) where only the integrated versions of the spot-variance and forward variance are used.

In the following we will assume that η is in fact square integrable. In this case the spot-variance process $\xi := \eta^2$ has the canonical decomposition (see [53, Definition I.4.22]) given by

$$\xi = \xi_0 + M + A, \tag{3.1}$$

such that M is a martingale and A is a predictable process of finite variation that is $\mathbb{P}$ integrable. Then the forward variance is given by (see Definition 2.8)

$$v(t,T) = \mathbb{E}[\xi_T | \mathcal{F}_t] = \xi_t + \mathbb{E}[A_T - A_t | \mathcal{F}_t], \quad t \in [0,T] \tag{3.2}$$

and we see that for each fixed time t, with $0 \leq t \leq T$, the shape of the term-structure of the forward variance $T \mapsto v(t,T)$ depends basically only on the finite-variation process A.

3.1 Spot-Variance Processes with Affine Drift

Now we will consider the important case where the finite variation part A is of the (Hull-White extended) affine form, which includes the popular class of (time-inhomogeneous) *affine processes* (see [30] and [33] respectively).

Definition 3.4. *A spot-variance process ξ is said to have an affine drift, if the process A in the canonical decomposition* (3.1) *satisfies*

$$A_t = \alpha t + \kappa \int_0^t \xi_s \, ds, \quad t \geq 0,$$

where α and κ are real constant. If α is a continuous deterministic function, such that the process A satisfies

$$A_t = \int_0^t (\alpha_s + \kappa \xi_s) \, ds, \quad t \geq 0, \tag{3.3}$$

the spot-variance ξ is said to have an Hull-White extended affine drift.

Proposition 3.5. *If the spot-variance process ξ has an affine drift, then for each fixed time t the term structure of the forward variance is (time-homogeneous) affine in the spot-variance with representation given by*

$$v(t,T) = \Phi(T-t) + \Psi(T-t)\xi_t \quad t \in [0,T], \tag{3.4}$$

where $\Phi(x) := \frac{\alpha}{\kappa}(\Psi(x) - 1)$ and $\Psi(x) := e^{\kappa x}$ for all $x \geq 0$. If ξ has an Hull-White extended affine drift, the term structure of the forward variances have at each time t a (time-inhomogeneous) affine representation given by

$$v(t,T) = \tilde{\Phi}(t,T) + \Psi(T-t)\xi_t \quad t \in [0,T], \tag{3.5}$$

where $\tilde{\Phi}(t,T) := \int_t^T \Psi(T-s)\alpha_s\,ds$.

Moreover, the process α *can be used to match any differentiable initial forward variance curve* v_0 *by choosing* $\xi_0 = v_0(0)$ *to satisfy the short-end condition* (2.6) *and by setting*

$$\alpha_t := \frac{d}{dt}v_0(t) - \kappa v_0(t), \quad \forall t \geq 0. \tag{3.6}$$

In this case the forward variance can be represented by

$$v(t,T) = v_0(T) + \Psi(T-t)\big(\xi_t - v_0(t)\big), \quad t \in [0,T]. \tag{3.7}$$

Proof. The Representation (3.5) follows from (3.2) applied to (3.3) which gives

$$\mathbb{E}[\xi_T|\mathcal{F}_t] = \xi_t + \int_t^T (\alpha_s + \kappa\mathbb{E}[\xi_s|\mathcal{F}_t])\,ds, \quad T \geq t,$$

and accordingly for each fixed t, the term structure $T \mapsto v(t,T)$ is a solution of the semi-linear ODE

$$\begin{cases} \frac{d}{dT}f_T = \kappa f_T + \alpha_T, & T \geq t, \\ f_t = \xi_t, \end{cases}$$

which is solved by (3.5) and for constant α this reduces to (3.4). For the second claim, comparing Representations (3.7) and (3.5) we see that the claim follows if for the choice α as in (3.6) we have $\tilde{\Phi}(t,T) = v_0(T) - \Psi(T-t)v_0(t)$, but this follows from integration by parts, as

$$\tilde{\Phi}(t,T) \;=\; \int_t^T e^{\kappa(T-s)}(\frac{d}{dt}v_0(s) - \kappa v_0(s))\,ds = v_0(T) - e^{\kappa(T-t)}v_0(t).$$

In particular, by choosing $\xi_0 = v_0(0)$ we see that $v(0,T) = v_0(T)$ for all $T \geq 0$ and hence the claim. □

Now assume that the spot-variance process ξ has an Hull-White extended affine drift with representation given by (3.3). Then it follows from the Representation (3.5) and Itô's product rule that

$$\begin{aligned} v(t,T) &= v(0,T) + \tilde{\Phi}(t,T) - \tilde{\Phi}(0,T) + \Psi(T-t)\xi_t - \Psi(T)\xi_0 \\ &= v(0,T) - \int_0^t \Psi(T-s)A_s\,ds + \int_0^t \Psi(T-s)\,d\xi_s & (3.8) \\ &= v(0,T) + \int_0^t \Psi(T-s)\,dM_s. & (3.9) \end{aligned}$$

Note that the Representation (3.9) is independent of the drift term A and in particular the same irrespective whether ξ has a plain affine drift or an Hull-White extended affine drift. However, the difference is in the initial curve $v(0,\cdot)$. From the Representations (3.4) and (3.5) we see that the initial curve is given by

$$v(0,T) = \int_0^T e^{\kappa(T-s)}\alpha_s\,ds + \Psi(T)\xi_0,$$

that is, for constant α it is necessarily an element of the parameterized family of curves given by

$$\Big\{\mathbb{R}_+ \ni x \mapsto e^{\kappa x}\big(\frac{\alpha}{\kappa} + \xi_0\big) - \frac{\alpha}{\kappa}\,\Big|\,\alpha,\kappa,\xi_0 \in \mathbb{R}\Big\}, \tag{3.10}$$

where we considered ξ_0 also as a parameter, whereas if α is assumed to be a continuous curve that can be used as a parameter, then any continuously differentiable initial curve is possible by choosing α as in (3.6).

Example 3.6 (Heston Model). Consider the Hull-White extended Heston model (see [52]) given as the solution of

$$\begin{cases} dX_t & = -\frac{1}{2}\xi_t\, dt + \sqrt{\xi_t}\, d\beta_t^1 \\ d\xi_t & = \kappa\left(\theta_t - \xi_t\right) dt + \rho\sqrt{\xi_t}\, d\beta_t^2 \\ X_0 & = x,\ \xi_0 = \xi, \end{cases}$$

where κ and ρ are positive constants and θ is a continuous deterministic process satisfying $2\kappa\theta_t > \rho^2$ for all $t \geq 0$ and β^1 and β^2 are possibly correlated Brownian motions with correlation $\zeta \in [-1, 1]$. Then according to Proposition 3.5 the forward variances have the representation

$$v(t,T) = \int_t^T e^{-\kappa(T-s)}\kappa\theta_s\, ds + e^{-\kappa(T-t)}\xi_t$$

and given any continuously differentiable initial curve v_0 we can set θ such that $\kappa\theta_t = \frac{d}{dt}v_0(t) + \kappa v_0(t)$ holds for all $t \geq 0$. Then if also $\xi_0 = v_0(0)$ we have $v(0,T) = v_0(T)$ for all $T \geq 0$ and

$$\begin{aligned} v(t,T) &= v(0,T) + e^{-\kappa(T-t)}(\xi_t - v_0(t)) \\ &= v(0,T) + \int_0^T e^{-\kappa(T-s)}\rho\sqrt{\xi_s}\, d\beta_s^2, \end{aligned} \tag{3.11}$$

where $v(0,T) = v_0(T)$.

The Representation (3.11) is the first example of a (classic) HJM-model (see [51]), in that for each fixed but arbitrary $T \geq 0$, the forward variance $v(t,T)$ can be represented by an Itô-process. In particular it can start at an arbitrary (continuously differentiable) initial curve $v_0(T)$. Note that the Representations (3.8) and (3.9) are included in more general, semimartingale driven *HJM* models as considered for example in [56]. If we choose $M_t = \int_0^t \sigma_s\, d\beta_s$ we are in the considered (classic) HJM setting, as the forward variances then have the representation

$$v(t,T) = v(0,T) + \int_0^t e^{\kappa(T-s)}\sigma_s\, d\beta_s. \tag{3.12}$$

In the next section we will look at general (classic) HJM models.

4 HJM Theory

In the previous sections we defined the (idealized) payoff of a (predictable) variance swap by the (predictable) quadratic variation of the log-price. Under the Assumption 2.1 we saw that the price of a predictable variance swap as given in Representation (2.7) is determined by the spot-variance, given as the weak time derivative of the predictable quadratic variation. The forward variances are given by the term-structure equation, see Definition 2.8. This setting is very tractable, in particular for spot-variance models with affine drift (see Definition 3.4), but as for short-rate models for the market of zero-coupon bonds, lacks the flexibility to model realistic market movements and in particular to start at an arbitrary initial curve. In the following we will take the HJM perspective (see [51]) by starting with a general model for the forward variances v. Given a model for the forward variances v, the spot-variance ξ is then determined by the short-end condition (see (2.6)) $\xi_t := v(t,t)$, $t \geq 0$, which is well-defined under some mild conditions on the forward variance process. Then we can define (as in (2.7)) for each $T \geq 0$ the variance swap rate and predictable variance swap price processes for each fixed $T > 0$ by

$$V(t,T) := \int_t^T v(t,s)\,ds, \quad t \leq T \tag{4.1}$$

$$VS^P(t,T) := \int_0^t \xi_s\,ds + V(t,T), \quad t \leq T. \tag{4.2}$$

Remark 4.1. Equation (4.2) very closely resembles the corresponding situation in the HJM theory for interest rates as introduced in [51]. In fact, if we assume that v is a model for the forward (interest) rate, then the first term corresponds to the logarithm of the bank account process and the second term to the logarithm of the zero coupon bond price process multiplied by -1.

For the definition of the variance swap price process we need the concept of a compatible log-price (generalizing the Definition 3.2, cf. also [35]).

Definition 4.2. *Let ξ be a predictable process modeling the spot-variance. Then a semimartingale X satisfying the conditions of Assumption 2.1 is said to be a compatible log-price process for the spot-variance ξ, if $\partial_t\langle X,X\rangle_t = \xi_t$ for all $t \geq 0$ (see Remark 2.2).*

If X is a compatible log-price process X we arrive via (2.1) at the representation for the variance swap price process by defining

$$VS(t,T) := [X,X]_t - \langle X,X\rangle_t + VS^P(t,T), \quad t \leq T. \tag{4.3}$$

For the case of a continuous log-price process X, this is basically the same setup as in [15]. Finally, following [28] and using the Representation (4.3) (cf. Remark 2.2), we can make the following Definition regarding the absence of arbitrage.

Definition 4.3. *The market of variance swaps is free of arbitrage, if and only if, there exists a probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$, such that the predictable variance swap price processes $t \mapsto VS^P(t,T)$, $t \in [0,T]$, $T \geq 0$, is a (local) martingales. The measure $\mathbb{Q}$ is then called an equivalent (local) martingale measure (E(L)MM). If under $\mathbb{P}$ the variance swap price processes is a (local) martingales, it will be called (local) martingale measure ((L)MM).*

4.1 Term Structure Movements and Drift Condition

In this subsection we closely follow [34, Chapter 6] and [32, Chapter 4]. We denote by $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P})$ a given complete filtered probability space satisfying the usual conditions and such that β is a d-dimensional $(\mathcal{F}_t)$-Brownian motion, for some $d \in \mathbb{N}$ and $\mathbb{P}$ some given probability measure which not necessarily is a LMM. We assume further that we are given an initial term structure $v(0,\cdot)$ and mappings α and σ that satisfy the following conditions (see Appendix A for the notation).

(HJM1) α is a measurable mapping from $(\Delta \times \Omega, \mathcal{B}(\Delta) \otimes \mathcal{F})$ into $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and σ is a measurable mapping from $(\Delta \times \Omega, \mathcal{B}(\Delta) \otimes \mathcal{F})$ into $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, where $\Delta := \{(t,T) \in \mathbb{R}^2 \,|\, 0 \le t \le T\}$.

(HJM2) For each $T > 0$, $\sigma(\cdot, T) \in \mathcal{L}_T^{loc}(\mathbb{R})$ and $\alpha(\cdot, T)$ is predictable and Lebesgue integrable.

(HJM3) As functions of T, $v(t,T), \alpha(t,T)$ and $\sigma(t,T)$ are a.s. continuous.

Then, for each fixed but arbitrary $T > 0$, we assume that the forward variance process evolves according to the Itô process given by

$$v(t,T) = v(0,T) + \int_0^t \alpha(s,T)\,ds + \sum_{j=1}^d \int_0^t \sigma^j(s,T)\,d\beta_s^j, \quad t \in [0,T]. \tag{4.4}$$

Now we can immediately formulate the well-known HJM-drift condition for forward variance models.

Proposition 4.4. *$\mathbb{P}$ is an LMM in the sense of Definition 4.3, if and only if,*

$$\alpha(t,T) = 0, \quad \textit{for all } T, \quad d\mathbb{P} \otimes dt\textit{-a.s.} \tag{4.5}$$

In this case the predictable variance swap price process evolves according to

$$VS^p(t,T) = VS^p(0,T) + \sum_{j=1}^d \int_0^t \tilde{\sigma}^j(s,T)\,d\beta_s^j,$$

where $VS^p(0,T) := \int_0^T v(0,s)\,ds$ and $\tilde{\sigma}^j(t,T) := \int_t^T \sigma^j(t,u)\,du$ for all $t \le T$. If further $\sigma(\cdot,T) \in \mathcal{L}_T^2(\mathbb{R})$ then $\mathbb{P}$ is a martingale measure.

Proof. As this is a well-known result, we only sketch the proof. Conditions **(HJM1)**-**(HJM3)** enable us to use the ordinary and stochastic Fubini Theorem, see for example [34, Theorem 2.3.2]. By denoting the spot variance process $\xi_t := v(t,t)$, $t \ge 0$, we get

$$\int_t^T v(t,u)\,du \;=\; \int_0^T v(0,u)\,du + \int_0^t \tilde{\alpha}(s,T) - \xi_s\,ds + \sum_{j=1}^d \int_0^t \tilde{\sigma}^j(s,T)\,d\beta_s^j,$$

where $\tilde{\alpha}(t,T) := \int_t^T \alpha(t,u)\,du$ and $\tilde{\sigma}^j(t,T) := \int_t^T \sigma^j(t,u)\,du$ for $t \le T$ and accordingly from (4.2) we arrive at the representation

$$VS^P(t,T) = \int_0^T v(0,u)\,du + \int_0^t \tilde{\alpha}(s,T)\,ds + \sum_{j=1}^d \int_0^t \tilde{\sigma}^j(s,T)\,d\beta_s^j.$$

Thus $\mathbb{P}$ is a *LMM*, if and only if, for all $T \ge 0$, $\int_0^t \tilde{\alpha}(s,T)\,ds$ is indistinguishable from zero, which is equivalent to $\tilde{\alpha}(t,T) = 0$ for all $T \ge 0$, $d\mathbb{P} \otimes dt$-a.s. Thus (4.5) follows from differentiation. □

In this generality HJM models are not very useful for applications. It should rather be seen as a general framework to formulate the drift-condition for forward variance models. In the following we will discuss a very popular class of *HJM* models.

4.2 Linear Models

In the following we assume that $\mathbb{P}$ is a MM (recall Definition 4.3). A very tractable choice for the forward variances is then given by so-called linear models, which are models that for each fixed but arbitrary T are solutions to the *linear* stochastic differential equation given by

$$v(t,T) = v(0,T) + \int_0^t v(s,T)\, dM_s(T), \quad t \in [0,T], \tag{4.6}$$

where for each T, $t \mapsto M_t(T)$ is a continuous local martingale with representation

$$M_t(T) = \sum_{j=1}^{d} \int_0^t \Sigma_j(s,T)\, d\beta_s^j,$$

for $\Sigma_1, ..., \Sigma_d$ satisfying the conditions **(HJM1)-(HJM3)**. Solutions to such models can be represented by (see [65, Proposition IX.2.3])

$$\begin{aligned} v(t,T) &= v(0,T)\mathcal{E}_t(M(T)) \qquad (4.7) \\ &= v(0,T) \exp\Big(\sum_{j=1}^{d} \int_0^t \Sigma_j(s,T)\, d\beta_s^j - \frac{1}{2} \sum_{j=1}^{d} \int_0^t \Sigma_j^2(s,T)\, ds \Big). \end{aligned}$$

The situation is particularly tractable if

$$\Sigma_j(t,T) = \sigma_j(T-t), \quad j = 1, ..., d, \tag{4.8}$$

for smooth deterministic functions $\sigma_1, ..., \sigma_d$, since in this case the corresponding forward variance process in Musiela's parameteriaton (to be introduced in the next section) given by $t \mapsto v(t, t + \cdot)$ becomes a time-homogeneous Markov process.

Example 4.5 (Constant Diffusion). The simplest (non-trivial) choice is $d = 1$ and σ being a constant. In this case the forward variance process and the variance swap rate can be represented as

$$v(t,T) = v(0,T)\mathcal{E}_t(\sigma\beta), \quad V(t,T) = V(0,T)\mathcal{E}_t(\sigma\beta), \quad \forall 0 \le t \le T.$$

The spot-variance process $\xi_t := v(t,t) = v(0,t)\mathcal{E}_t(\sigma\beta)$ can be represented for a strictly positive and differentiable initial curve $v(0,t)$ by

$$\xi_t = v(0,0) \exp\Big(\int_0^t \Big(\frac{1}{v(0,s)} \frac{d}{ds} v(0,s) - \frac{1}{2}\sigma^2 \Big)\, ds + \int_0^t \sigma\, d\beta_s \Big).$$

In particular, if $v(0,T) := ae^{bT}$, then the spot-variance process reduces to the solution of

$$d\xi_t = b\xi_t\, dt + \sigma\xi_t\, d\beta_t, \quad \xi_0 = a, \tag{4.9}$$

and accordingly follows a geometric Brownian motion and in particular has an affine drift in the sense of Section 3.1. Hence Proposition 3.5 applies and we have the interesting special case in which a linear forward variance model admits an affine drift. We can show this also by a direct computation as,

$$\begin{aligned} v(t,T) &= v(0,T) + \int_0^t \sigma v(s,T)\, d\beta_s = v(0,T)\mathcal{E}_t(\sigma\beta) \\ &= a\exp(bT)\exp\Big(-\int_0^t \frac{1}{2}\sigma^2\, ds + \int_0^t \sigma\, d\beta_s \Big) = \exp(b(T-t))\xi_t. \end{aligned}$$

We summarize this in the following Proposition.

Proposition 4.6. *Let v be the solution of the linear SDE given in* (4.6) *with $d = 1$ and $\Sigma(t,T) = \sigma$. If the initial curve $v(0,t)$ is of the form $v(0,t) = a\exp(bt)$ then the forward variance model v admits an affine realization.*

Example 4.7 (The Bergomi model). In the Bergomi model (see [7]) the diffusion coefficients $\sigma_1, ..., \sigma_d$ are given by $\sigma_i(x) = \omega e^{-\kappa_i x}$ for $x \geq 0$ and for all $i = 1, ..., d$. In this case we have

$$M_t(T) = \sum_{j=1}^{d} \int_0^t \omega e^{-\kappa_j(T-s)}\, d\beta_s^j = \sum_{j=1}^{d} \omega e^{-\kappa_j(T-t)} Z_t^j, \tag{4.10}$$

where each Z^j is a solution to

$$Z_t^j = -\kappa_j Z_t^j\, dt + d\beta_t^j, \quad Z_0 = 0 \tag{4.11}$$

and accordingly (independent) Ornstein-Uhlenbeck processes. From (4.7) the forward variances in the Bergomi model have the representation

$$\frac{v(t,T)}{v(0,T)} = \exp\Big(\sum_{j=1}^{d} e^{-\kappa_j(T-t)} Z_t^j - \sum_{j=1}^{d} \frac{\omega^2}{4\kappa_j}(e^{-2\kappa_j(T-t)} - e^{-2\kappa_j T})\Big) \tag{4.12}$$

$$= \exp\Big(\sum_{j=1}^{d} e^{-\kappa_j(T-t)} Z_t^j - \sum_{j=1}^{d} \frac{\omega^2}{2} e^{-2\kappa_j(T-t)} \mathbb{E}[(Z_t^j)^2]\Big), \tag{4.13}$$

where the second equation corresponds to representation used in [7].

Remark 4.8. Note the contrast to HJM models for forward interest rates in which models with linear diffusion coefficients generally do not exist. That is, by denoting the forward rare at time t for time-of-maturity T by $f(t,T)$, it can be shown (see [34, Section 6.4.1] and the references therein) that the equation corresponding to (5.9) with $M_t(T) = \sigma\beta_t$, for some $\sigma > 0$, given by

$$df(t,T) = \Big(\sigma f(t,T) \int_t^T \sigma f(t,u)\, du\Big)\, dt + \sigma f(t,T)\, d\beta_t,$$

can not admit a finite valued solution. This apparently lead to the development of the LIBOR market model, see again [34, Section 6.4.1] for a discussion of this aspect.

5 HJM Musiela Theory

From now on we will assume that $\mathbb{P}$ is a *LMM* (see Definition 4.3). In this case, if **(HJM1)** – **(HJM3)** hold true, it follows from Proposition 4.4, that for each fixed but arbitrary $T \geq 0$, the forward variance process $v(\cdot, T)$ can be represented as

$$v(t,T) = v(0,T) + \sum_{j=1}^{d} \int_0^t \sigma^j(s,T)\, d\beta_s^j, \quad t \in [0,T]. \tag{5.1}$$

Now we will switch to the HJM Musiela perspective by considering the forward variance process in Musiela's parametrization (see [62])

$$u(t,x) := v(t, t+x), \quad \forall (t,x) \in \mathbb{R}_+ \times \mathbb{R}_+$$

and state conditions (following [40, Section 4]) such that the process $u_t := u(t,\cdot)$ takes values in a Hilbert space H that satisfies the conditions (**H1**) and (**H2**) in [40], namely

(H1) H is a separable Hilbert space continuously embedded in $C(\mathbb{R}_+, \mathbb{R})$ (that is, for every $x \in \mathbb{R}_+$, the pointwise evaluation $ev_x : u \mapsto u(x)$ is a continuous linear functional on H) and $1 \in H$ (the constant function 1).

(H2) The family $\{S_t \,|\, t \in \mathbb{R}_+\}$ of right-shifts $S_t u = u(t + \cdot)$ forms a strongly continuous semi-group S in H with infinitesimal generator denoted by $\frac{d}{dx}$.

Remark 5.1. Unless otherwise stated, we shall think of the Hilbert space H that was introduced in [32, Chapter 5], which consists of absolutely continuous functions h on $\mathbb{R}_+$ equipped with the norm

$$\|h\|_H^2 := |h(0)|^2 + \int_{\mathbb{R}_+} |h'(x)|^2 w(x)\, dx, \tag{5.2}$$

where $w : \mathbb{R}_+ \to [1, \infty)$ is a non-decreasing weighting function. Notice that we do not need an integrability condition on the weighting function as in [32, Condition (5.1)] as the additional condition **(H3)** necessary for the regularity of the HJM-drift for forward interest rates is not required here (cf. [32, Remark 5.1.1]). Hence, possible choices include the constant function $w(x) = 1$ for all $x \in \mathbb{R}_+$ and $w(x) = e^{\alpha x}$ for some $\alpha > 0$. According to [32, Theorem 5.1.1] and [32, Corollary 5.1.1.] the space satisfies **(H1)** and **(H2)**, the elements are bounded and it holds that

$$D(d/dx) = \{h \in H \,|\, h' \in H\}.$$

If the weighting function w needs to be emphasized, we will write H_w.

In the following we will consider a slightly more general situation than necessary, by looking at models given by

$$f(t,T) = f(0,T) + \int_0^t \alpha(s,T)\, ds + \sum_{j=1}^d \int_0^t \sigma^j(s,T)\, d\beta_s^j, \quad t \in [0,T]. \tag{5.3}$$

Lemma 5.2. *If $r^* := f(0,\cdot)$, $\alpha_t(\omega) := \alpha(t,\omega,\cdot + t)$ and $\sigma_t^j(\omega) := \sigma^j(t,\omega,\cdot + t)$, $j = 1, ..., d$, satisfy the conditions*

(C1) *$r^* \in H$, and*

(C2) *α and σ^j, $j = 1, ..., d$ are H-valued predictable processes satisfying*

$$\mathbb{P}\Big[\int_0^t \|\alpha_s\|_H\, ds < \infty\Big] = 1, \quad \forall t \geq 0,$$

and $\sigma^j \in \mathcal{L}^{loc}(H)$, for $j = 1, ..., d$.

Then for each $t \geq 0$, $r_t := f(t, t + \cdot)$ can be represented as

$$r_t = S(t)r^* + \int_0^t S(t-s)\alpha_s\, ds + \sum_{j=1}^d \int_0^t S(t-s)\sigma_s^j\, d\beta_s^j,$$

and is a mild solution in H of

$$\begin{cases} dr_t = (\frac{d}{dx} r_t + \alpha_t)\, dt + \sum_{j=1}^d \sigma_t^j\, d\beta_t^j \\ r_0 = r^*. \end{cases}$$

If also $\sigma_j \in \mathcal{L}^2(H)$, $j = 1, ..., d$, then the mild solution can be chosen to be continuous.

Proof. This is [32, Section 4.2] and [26, Poposition 7.3]. □

Thus, if $u^* := v(0,\cdot)$ and $\sigma_t(\omega) := \sigma(t,\omega,\cdot+t)$ satisfy the conditions **(C1)** and **(C2)**, then for each $t \geq 0$, u satisfies

$$u_t = S(t)u^* + \sum_{j=1}^{d} \int_0^t S(t-s)\sigma_s^j \, d\beta_s^j \tag{5.4}$$

and accordingly is a mild solution of

$$\begin{cases} du_t = \frac{d}{dx}u_t \, dt + \sum_{j=1}^{d} \sigma_t^j \, d\beta_t^j \\ u_0 = u^*. \end{cases} \tag{5.5}$$

Proposition 5.3. *Fix some $T > 0$. If $\sigma_j \in \mathcal{L}_T^2(H)$, $j = 1,...,d$, $\sigma_j \in D(d/dx)$ $d\mathbb{P}\otimes dt$-almost surely on $\Omega \times [0,T]$ and $\frac{d}{dx}\sigma_j \in \mathcal{L}_T^2(H)$, $j = 1,...,d$, then for any $u^* \in D(d/dx)$, (5.4) is also a strong solution of (5.5).*

Proof. The existence of a continuous mild solution follows from the above and that this is also a strong solution follows from [26, Proposition 6.4]. □

5.1 Spot-Variance Models with Affine Drift

We continue the discussion of spot-variance models with affine drift as given in Section 3.1. The following proposition follows immediately from Lemma 5.2. We consider the Hilbert Space H given in Remark 5.1 with weighting function $w(x) = e^{\alpha x}$ for some fixed $\alpha \geq 0$.

Proposition 5.4. *Let $u_t := v(t, t+\cdot)$ and $u^* := v(0,\cdot)$, where v is given by (3.12) with $\sigma \in \mathcal{L}^{loc}(\mathbb{R})$. If $v(0,\cdot) \in H$ and $2\kappa < -\alpha$, then u is a mild solution of*

$$\begin{cases} du_t = \frac{d}{dx}u_t \, dt + \sigma_t e^{\kappa\cdot} \, d\beta_t \\ u_0 = u^*. \end{cases} \tag{5.6}$$

If $u^* \in D(d/dx)$ and $\sigma \in \mathcal{L}^2(\mathbb{R})$ the solution is also a strong solution.

Proof. It follows from (5.2), that $e^{\kappa\cdot} \in D((d/dx)^n) \subset H$ for $n = 0$ (and then also for all $n \geq 0$) if and only if $2\kappa < -\alpha$. In this case the process $\tilde{\sigma} := (\sigma_t e^{\kappa\cdot})_{t\geq 0}$ is in $\mathcal{L}^{loc}(H)$ as $\sigma \in \mathcal{L}^{loc}(\mathbb{R})$ which gives Condition **(C2)** and thus according to Lemma 5.2 the first claim. For the second claim, note that $\tilde{\sigma}$ is in $\mathcal{L}^2(H)$ whenever $\sigma \in \mathcal{L}^2(\mathbb{R})$ and that in this case also $\frac{d}{dx}\tilde{\sigma} = \kappa\tilde{\sigma} \in \mathcal{L}^2(H)$. Hence the second claim follows from Proposition 5.3. □

Let the conditions of Proposition 5.4 be satisfied and denote by u the strong solution of (5.6) for $u_0 = u^* \in D(d/dx)$. Also consider the representation given in (3.7) which in Musiela's parametrization $\tilde{u}_t := v(t, t+\cdot)$ corresponds to

$$\tilde{u}_t = S_t u^* + e^{\kappa\cdot}(\xi_t - u^*(t)), \tag{5.7}$$

where $\tilde{u}_0 := u^*$ and $\xi_t = \tilde{u}_t(0)$ is the spot-variance. Notice that this representation consists of the sum of an deterministic process in H given by $t \mapsto S_t u^*$, where at each time t, $x \mapsto S_t u^*(x) = u^*(t+x)$ and a process given by $t \mapsto e^{\kappa\cdot}(\xi_t - u^*(t))$

where $e^{\kappa\cdot}$ is a constant element in H and $t \mapsto (\xi_t - u^*(t))$ is a real-valued stochastic process. Accordingly at each time t, $e^{\kappa\cdot}(\xi_t - u^*(t))$ takes values in the linear one-dimensional space given by the span of $e^{\kappa\cdot}$ and consequently $\tilde{u}_t$ given by (5.7) stays in the 2-dimensional submanifold with boundary of H given by

$$\mathcal{M}_{u^*} := \{S_t u^* + e^{\kappa\cdot} z | (t,z) \in \mathbb{R}_+ \times \mathbb{R}\} \subset D(d/dx).$$

Proposition 5.5. *Let $u^* \in D(d/dx)$ and $\sigma \in \mathcal{L}^2(\mathbb{R})$, such that ξ is a strong solution to*

$$d\xi_t = (\alpha_t + \kappa\xi_t)\,dt + \sigma_t\,d\beta_t, \quad \xi_0 = u^*(0),$$

with α given by (3.6). *Then $\tilde{u}$ given by* (5.7) *is a strong solution to* (5.6) *and accordingly $\tilde{u} = u$ by uniqueness of strong solutions.*

Proof. From the Definition of α given in (3.6) and Itô's product rule, it follows that $\tilde{\xi}_t := \xi_t - u^*(t)$ is a strong solution to

$$d\tilde{\xi}_t = \kappa\tilde{\xi}_t\,dt + \sigma_t\,d\beta_t, \quad \tilde{\xi}_0 = 0 \tag{5.8}$$

and accordingly from the definition of the semi-flow S_t, the linearity of the (stochastic) integral and the boundedness of $\frac{d}{dx}$ on $\mathcal{M}_{u^*}$, we get

$$\begin{aligned}
\tilde{u}_t &= S_t u^* + e^{\kappa\cdot}\tilde{\xi}_t = u^* + \int_0^t \frac{d}{dx} S_s u^*\,ds + e^{\kappa\cdot}\tilde{\xi}_t \\
&= u^* + \int_0^t \frac{d}{dx} S_s u^*\,ds + \int_0^t \kappa e^{\kappa\cdot}\tilde{\xi}_s\,ds + \int_0^t \sigma_s e^{\kappa\cdot}\,d\beta_s \\
&= u^* + \int_0^t \Big(\frac{d}{dx} S_s u^* + \frac{d}{dx} e^{\kappa\cdot}\tilde{\xi}_s\Big)\,ds + \int_0^t \sigma_s e^{\kappa\cdot}\,d\beta_s \\
&= u^* + \int_0^t \frac{d}{dx}\tilde{u}_s\,ds + \int_0^t \sigma_s e^{\kappa\cdot}\,d\beta_s,
\end{aligned}$$

which gives the claim. □

Thus, the SPDE (5.6) admits for every $u_0 \in D(d/dx)$ a *generic finite dimensional realization* (see next chapter). Such models are also called *generic affine realizations* as for initial curves from the set (3.10) (which corresponds to an affine submanifold of H) the term structure of the model given by (3.5) is affine in the spot-variance.

5.2 Linear models

We continue the discussion of linear models started in Section 4.2. We assume that in (4.6) the diffusion coefficients are of the form (4.8), that is, we look at forward variance processes with representation

$$v(t,T) = v(0,T) + \sum_{j=1}^d \int_0^t v(s,T)\sigma_j(T-s)\,d\beta_s^j. \tag{5.9}$$

As in the last subsection, we want to derive conditions such that this equation can be represented in Musiela's parametrization $u_t := v(t, t+\cdot)$ as a stochastic partial differential equation in the Hilbert space H, for some weighting function w. Before we proceed we introduce the pointwise multiplication operator m which acts on two given functions $f, g : \mathbb{R}_+ \to \mathbb{R}$ by pointwise multiplication, that is,

$$m(f,g) : \mathbb{R}_+ \to \mathbb{R}, \quad x \mapsto m(f,g)(x) = f(x)g(x). \tag{5.10}$$

Lemma 5.6. *Let σ be absolutely continuous. Then the multiplication operator given by* (5.10) *satisfies $m(\sigma, h) \in H$ for all $h \in H$ if and only if $\sigma \in H$. In this case for any fixed $\sigma \in H$ we have $m(\sigma, \cdot) \in L(H, H)$.*

Proof. Necessity of the first claim follows from $m(\sigma, 1) = \sigma$. For the sufficiency, first note that the product of absolutely continuous functions is again absolutely continuous. Further, we have from (5.2)

$$\begin{aligned} \|m(\sigma, h)\|_H^2 &= |\sigma(0)h(0)|^2 + \int_{\mathbb{R}_+} \Big|\sigma(x)'h(x) + \sigma(x)h'(x)\Big|^2 w(x)\, dx \\ &\leq |\sigma(0)h(0)|^2 + 2\int_{\mathbb{R}_+} \Big(\Big|\sigma(x)'h(x)\Big|^2 + \Big|\sigma(x)h'(x)\Big|^2\Big) w(x)\, dx \\ &=: |\sigma(0)h(0)|^2 + I_1 + I_2, \end{aligned}$$

where

$$\begin{aligned} I_1 &= 2\int_{\mathbb{R}_+} \Big|\sigma(x)'\Big(h(0) + \int_0^x h'(y)\, dy\Big)\Big|^2 w(x)\, dx \\ &\leq 4\int_{\mathbb{R}_+} \Big(\Big|\sigma(x)'h(0)\Big|^2 + \Big|\sigma(x)' \int_{\mathbb{R}_+} h'(x)\, dx\Big)\Big|^2\Big) w(x)\, dx \\ &= 4(|h(0)|^2 + \|h'\|_{L^1(\mathbb{R}_+)}^2) \int_{\mathbb{R}_+} |\sigma'(x)|^2 w(x)\, dx \\ &= 4(|h(0)|^2 + \|h'\|_{L^1(\mathbb{R}_+)}^2)(\|\sigma\|_H^2 - |\sigma(0)|^2) < \infty, \end{aligned}$$

as $\|h'\|_{L^1(\mathbb{R}_+)}^2 \leq C_1 \|h\|_H^2$, where C_1 is a constant that depends only on w according to [32, Equation (5.3)]. Similarly, for I_2 we have

$$I_2 = 4(|\sigma(0)|^2 + \|\sigma'\|_{L^1(\mathbb{R}_+)}^2)(\|h\|_H^2 - |h(0)|^2) < \infty,$$

which gives the first claim. Now fix $\sigma \in H$ with $\|\sigma\|_H^2 =: M$, then we see from the above that

$$\|m(\sigma, h)\|_H^2 \leq (9M + 8C_1 M)\|h\|_H^2,$$

whence the second claim. □

Remark 5.7. For a different proof of Lemma 5.6 see [71, Lemma 4.2].

Now assume that $u^* := v(0, \cdot)$ satisfies Condition **(C1)**. Then it is easy to see that the Equation (5.9) corresponds formally in Musiela's parametrization to

$$\begin{cases} du_t = \frac{d}{dx} u_t\, dt + \sum_{j=1}^d m(\sigma^j, u_t)\, d\beta_t^j \\ u_0 = u^*, \end{cases} \tag{5.11}$$

but unlike to the models considered in Section 5.1 we can not apply Lemma 5.2 directly, as the forward variance processes in (5.9) are given by solutions to SDEs and not by Itô-processes as in Section 5. In particular, we can not check Condition **(C2)** without knowing the process u. Therefore we proceed here by showing first that (5.11) admits a unique continuous mild solution. In a second step we will then show that condition **(C2)** is satisfied and apply Lemma 5.2.

Lemma 5.8. *Let $\sigma_j \in H$, $j = 1, ..., d$. Then for any $u^* \in H$, (5.11) has a unique continuous weak solution u such that for any $T > 0$ there exists a positive constant $C := C(T)$ and*

$$\mathbb{E}[\sup_{t\in[0,T]} \|u_t\|_H^2] \le C(1 + \|u^*\|_H^2). \tag{5.12}$$

Proof. The follows immediately from Proposition I.3.2 and Lemma 5.6. □

Corollary 5.9. *If $u^* := v(0, \cdot)$ satisfies Condition **(C1)**, then $u_t := v(t, \cdot)$, where $v(t, T)$ denotes the solution of (5.9), is a continuous mild solution of (5.11).*

Proof. By setting $v(t, T) := u(t, T - t)$ where u denotes the mild solution of (5.11) we get

$$\begin{aligned} v(t,T) &= S_t u^*(T-t) + \sum_{i=1}^d \int_0^t S_{t-s} m(\sigma_j, u_s)(T-t)\, d\beta_s^j \\ &= u_0(T) + \sum_{i=1}^d \int_0^t \sigma_j(T-s) u_s(T-s)\, d\beta_s^j \\ &= v(0,T) + \sum_{i=1}^d \int_0^t \sigma_j(T-s) v(s,T)\, d\beta_s^j, \end{aligned}$$

which is (5.9). Now in the notation of Lemma 5.2 we have pointwise

$$\sigma_t^j(\omega)(x) := \sigma_j(x) v(t, \omega, t + x) = m(\sigma_j, u_t(\omega))(x)$$

and using (5.12) it follows that $u \in \mathcal{L}_T^2(H)$ and hence also $m(\sigma_j, u) \in \mathcal{L}_T^2(H)$ which is **(C2)**. Hence the claim follows from Lemma 5.2. □

Thus we have shown that the solution v of (5.9) is in Musiela's parametrization a continuous mild solution of (5.11). On the other hand, we know from (4.7) that (5.9) is solved by

$$v(t,T) = \exp\Big(\log(v(0,T)) + \sum_{j=1}^d \int_0^t \sigma_j(T-s)\, d\beta_s^j - \frac{1}{2}\sum_{j=1}^d \int_0^t \sigma_j^2(T-s)\, ds\Big).$$

Now if we define $\tilde{J}(t, T) := \log(v(t, T))$ then it follows from the first part of Lemma 5.2 that for each $J^* := \log(v(0, t + \cdot)) \in H$, $J_t := \tilde{J}(t, t + \cdot)$ is a mild solution of

$$\begin{cases} dJ_t = (\frac{d}{dx} J_t - \frac{1}{2}\sum_{j=1}^d m(\sigma_j, \sigma_j))\, dt + \sum_{j=1}^d \sigma_j\, d\beta_t^j \\ J_0 = J^*, \end{cases} \tag{5.13}$$

as $m(\sigma_j, \sigma_j)$ and σ_j satisfy Condition (**C2**), on each bounded interval $[0, T]$ and each $j = 1, ..., d$. The process J can be decomposed as follows.

Lemma 5.10. *Let J^2 be the mild solution to the deterministic Cauchy problem*

$$\begin{cases} dJ_t^2 = (\frac{d}{dx} J_t^2 - \frac{1}{2}\sum_{j=1}^d m(\sigma_j, \sigma_j))\, dt \\ J_0^2 = J^*. \end{cases} \tag{5.14}$$

Then $J^1 := J - J^2$, is a mild solution to

$$\begin{cases} dJ_t^1 = \frac{d}{dx} J_t^1\, dt + \sum_{j=1}^d \sigma_j\, d\beta_t^j \\ J_0^1 = 0, \end{cases} \tag{5.15}$$

Proof. This follows from [26, Proposition A.4]. □

For the second part of Lemma 5.2 we have the following Condition.

Lemma 5.11. *If $\sigma_1, ..., \sigma_d$ satisfies*

$$\int_0^T \|S_t \sigma_j\|_H^2 \, dt < \infty, \quad j = 1, ..., d,$$

then the mild solution J of (5.13) *and J^1 of* (5.15) *can be chosen to be continuous on* $[0, T]$.

Proof. This is [26, Theorem 5.2]. □

For $h \in H$ we consider the pointwise exponential map on H given by

$$x \mapsto \exp(h)(x) := \exp(h(x)). \tag{5.16}$$

By construction it seems to be obvious that the process $t \mapsto \exp(J_t)$ agrees with $t \mapsto u_t$. However, as the process J (given as the mild solution of (5.13)) is not an Itô-process, we can not apply the Itô formula. We will show that the result still holds true by an application of [74, Theorem 2.2], but need first the following Lemmas.

Lemma 5.12. *Let w be an arbitrary weighting function. Then* $\exp(H) \subset H$.

Proof. For $h \in H$ arbitrary we have

$$\begin{aligned}
\|e^h\|_H^2 &= \; = |e^{h(0)}|^2 + \int_{\mathbb{R}_+} \left| h'(x) e^{h(x)} \right|^2 w(x) \, dx \\
&\leq |e^{h(0)}|^2 + \int_{\mathbb{R}_+} \left| h'(x) \Big(e^{h(0)} + \int_0^x (e^{h(y)})' \, dy \Big) \right|^2 w(x) \, dx \\
&\leq |e^{h(0)}|^2 + 2 \int_{\mathbb{R}_+} \left(\left| h'(x) e^{h(0)} \right|^2 + \left| h'(x) \int_0^x (e^{h(y)})' \, dy \right|^2 \right) w(x) \, dx \\
&\leq |e^{h(0)}|^2 (1 + 2\|h\|_H^2) + 2 \int_{\mathbb{R}_+} \left| h'(x) \int_0^x (e^{h(y)})' \, dy \right|^2 w(x) \, dx \\
&\leq |e^{h(0)}|^2 (1 + 2\|h\|_H^2) + 2 \left| \int_0^x (e^{h(y)})' \, dy \right|^2 \|h\|_H^2 \\
&\leq |e^{h(0)}|^2 (1 + 2\|h\|_H^2) + 2 \|(e^{h(\cdot)})'\|_{L^1(\mathbb{R}_+)}^2 \|h\|_H^2 .
\end{aligned}$$

Thus the claims follows upon showing that $\|(e^{h(\cdot)})'\|_{L^1(\mathbb{R}_+)} < \infty$, but this follows from the boundedness of $h \in H$, as

$$\begin{aligned}
\|(e^{h(\cdot)})'\|_{L^1(\mathbb{R}_+)} &= \int_{\mathbb{R}_+} |(e^{h(x)})'| \, dx = \int_{\mathbb{R}_+} e^{h(x)} |h'(x)| \, dx \\
&= \int_{\mathbb{R}_+ \cap h' \geq 0} e^{h(x)} h'(x) \, dx - \int_{\mathbb{R}_+ \cap h' < 0} e^{h(x)} h'(x) \, dx \\
&= \int_{h(\mathbb{R}_+ \cap h' \geq 0)} e^y \, dy - \int_{h(\mathbb{R}_+ \cap h' < 0)} e^y \, dy < \infty.
\end{aligned}$$

□

Lemma 5.13. *Let w be an arbitrary weighting function. Then for all $f, g \in H$ we have*

$$\exp(f + g) = m(\exp(f), \exp(g)).$$

Proof. From Lemmas 5.6 and 5.12 we know that both sides of the equation are well-defined elements in H. Hence, as the claim holds pointwise, we get immediately that $\| \exp(f + g) - m(\exp(f), \exp(g))\|_H^2 = 0$ and thus the claim. □

Lemma 5.14. *The exponential map $\exp : H \to H_+$, where $H_+ := \exp(H)$, is a diffeomorphism.*

Proof. We show first that for every h_0 the Frèchet derivative $D\exp(h_0)$ exists and is given by $m(\exp(h_0), \cdot)$. It suffices to show this for $h_0 = 0$. We have

$$\begin{aligned} \Big\| \frac{\exp(\epsilon h) - 1 - \epsilon h}{\epsilon} \Big\|_H^2 &= \Big| \frac{\exp(\epsilon h(0)) - 1 - \epsilon h(0)}{\epsilon} \Big|^2 \\ &+ \int_{\mathbb{R}_+} \Big| \frac{(\exp(\epsilon h(x)) - 1 - \epsilon h(x))'}{\epsilon} \Big|^2 w(x)\, dx =: I_1 + I_2. \end{aligned}$$

By Taylor's Theorem, the first term I_1 vanishes as $\epsilon \to 0$ for all $h \in B_H$, where B_H denotes the unit-sphere in H. For the second term, it holds true that

$$\begin{aligned} I_2 &= \int_{\mathbb{R}_+} \Big| \frac{\exp(\epsilon h(x))\epsilon h'(x) - \epsilon h'(x)}{\epsilon} \Big|^2 w(x)\, dx \\ &= \int_{\mathbb{R}_+} (h'(x))^2 (\exp(\epsilon h(x)) - 1)^2 w(x)\, dx \\ &\leq \Big(\int_{\mathbb{R}_+} (h'(x))^2 w(x)\, dx \Big) \Big(\exp(\epsilon h(0)) + \int_{\mathbb{R}_+} \frac{d}{dy} \exp(\epsilon h(y))\, dy - 1 \Big)^2 \\ &= (\|h\|_{H_w}^2 - |h(0)|^2) \Big(\exp(\epsilon h(0)) + \int_{\epsilon h(\mathbb{R}_+)} \exp(z)\, dz - 1 \Big)^2 \to 0, \end{aligned}$$

as $\epsilon \to 0$, again uniformly for all $h \in B_H$ and hence the first claim holds. Next, for any $h_0 \in H$, the derivative $D\exp(h_0) = m(\exp(h_0), \cdot)$ is an isomorphism, with inverse given by

$$[D\exp(h_0)]^{-1} = m(\exp(-h_0), \cdot),$$

hence it follows from the Inverse Mapping Theorem (see for example [62, 2.5.2 Inverse Mapping Theorem]) that $\exp : H \to H^+$ is a diffeomorphism. It is obvious that the inverse is given by the pointwise logarithm $\log : H^+ \to H$, $\log(h)(x) := \log(h(x))$ with derivative given according to the Inverse Mapping Theorem by

$$D\log(h_0) = \Big[D\exp(\log(h_0)) \Big]^{-1} = m(\exp(-\log(h_0), \cdot) = m(\frac{1}{h_0}, \cdot),$$

for every $h_0 \in H^+$. □

Now we can finally show the desired result.

Proposition 5.15. *Let u denote the unique continuous mild solution of* (5.11) *and let $\tilde{u}$ be the process given by $t \mapsto \tilde{u}_t := e^{J_t}$. Then $u = \tilde{u}$ on $[0, T]$.*

Proof. We know from Lemma 5.14 that $\exp : H \to H^+$ is diffeomorphism and we have that $\exp(D(d/dx)) = H^+ \cap D(d/dx)$ as $\frac{d}{dx}\exp(h) = m(\exp(h), \frac{d}{dx}h) \in H$ by Lemma 5.6 whenever $h \in D(d/dx)$. Also we have $\exp_*(\frac{d}{dx}) = \frac{d}{dx}$ (where $\exp_*(\frac{d}{dx})$ denotes the push forward of $\frac{d}{dx}$, see [74]) as

$$\begin{aligned}\exp_*\Big(\frac{d}{dx}\Big)(h) &= D\exp(\log(h))\frac{d}{dx}\log(h) = m(h, \frac{d}{dx}\log(h)) \\ &= m(h, m(\frac{1}{h}, \frac{d}{dx}h)) = \frac{d}{dx}h.\end{aligned}$$

Hence the conditions of [74, Theorem 2.2] are satisfied and we have that $\tilde{u} = e^J$ is a continuous mild solution to

$$\begin{cases} d\tilde{u}_t = (\frac{d}{dx}\tilde{u}_t - \frac{1}{2}\sum_{j=1}^d \exp_* m(\sigma_j,\sigma_j)(\tilde{u}_t) + \chi(\tilde{u}_t))\,dt + \sum_{j=1}^d (\exp_*\sigma_j)(\tilde{u}_t)\,d\beta_t^j \\ \tilde{u}_0 = \exp(J_0), \end{cases}$$

where $\exp_*\sigma_j(h) = m(\sigma_j, h)$ and

$$\chi(h) := \frac{1}{2}\sum_{j=1}^d D^2\exp(\log(h))(\sigma_i,\sigma_i) = \frac{1}{2}\sum_{j=1}^d \exp_* m(\sigma_j,\sigma_j)(h).$$

Thus $\tilde{u}$ satisfies (5.11) and the claim follows from uniqueness. □

Example 5.16 (The Bergomi model revisited). In the Bergomi model (see Example 4.7) in Musiela's parametrization the diffusion coefficients are given by $\sigma_j = \omega e^{-\kappa_j \cdot}$ for $j = 1, ..., d$ where $\kappa > 0$ and accordingly the corresponding equation (5.15) is given by

$$\begin{cases} dJ_t^1 = \frac{d}{dx}J_t^1\,dt + \sum_{j=1}^d \omega e^{-\kappa_j \cdot}\,d\beta_t^j \\ J_0^1 = 0, \end{cases} \tag{5.17}$$

which corresponds to the type of equation given in (5.6). It follows from Proposition 5.4 that whenever $e^{-\kappa\cdot} \in H$, (5.17) admits a strong solution on $[0,T]$ for each $T > 0$. Moreover, by setting as in (5.7) and (5.8),

$$\tilde{J}_t^1 := \sum_{j=1}^d e^{-\kappa_j \cdot} Z_t^j \tag{5.18}$$

with each Z^j, $j = 1, ..., d$ being strong solutions of

$$dZ_t^j := -\kappa_j Z_t^j\,dt + \omega\,d\beta_t^j, \tag{5.19}$$

it follows just as in Proposition 5.5 that (5.18) is a strong solution of (5.17) and hence $J^1 = \tilde{J}^1$ on $[0,T]$. And finally we get from Lemmas 5.10, 5.13 and Proposition 5.15 that

$$u_t = m(\exp(J_t^2), \exp(J_t^1)) = m(\exp(J_t^2), \exp(\sum_{j=1}^d e^{-\kappa_j \cdot} Z_t^j)). \tag{5.20}$$

Now it is easy to see, that

$$\{m(e^{J_t^2}, e^{\sum_{j=1}^d e^{-\kappa_j \cdot} z_j}) | (t, z_1, ..., z_d) \in \mathbb{R}_+ \times \mathbb{R}^d\} \subset D(d/dx)$$

and that accordingly u is in this case a strong solution to (5.11) whenever $u^* \in D(d/dx)$. Note that in this case (5.11) admits a *generic finite dimensional realization* to be introduced in the next chapter. For reasons that are similar to those given at the end of Section 5.1 such models will be called *exponential affine realizations.*

Chapter III

Finite dimensional Realizations

At the end of the last chapter we saw that the SPDEs (II.5.6) and (II.5.11) (with the diffusion coefficients given as in Example II.5.16) admit generic *finite dimensional realizations* in that the solutions for any initial curve within an open set of H could be represented by a smooth parametrization of finite dimensional Itô processes as given in (II.5.7) and (II.5.19) respectively. In this chapter we will look at more general versions of the SPDEs (II.5.6) and (II.5.11) and investigate the conditions on the diffusion coefficients such that the SPDEs admit *generic finite dimensional realizations*, that is, finite dimensional realizations for every initial curve in some open set of H. Answers to this question for the HJM-Musiela equation for forward (interest) rate models were first found by Tomas Björk and co-workers in a series of papers including [11] and [13] and later on completely solved by Damir Filipović and Josef Teichmann in a series of papers including [38] and [40]. For applications to forward variances see also [15] where the invariance conditions (see (2.2) and (2.3)) with respect to a *given* finite dimensional submanifold were cited. In the following the main reference will be [40]. In particular we will recapture the existence of affine realizations which is very closely related to the situation for the forward interest rates. We shall also systematically consider the existence of finite dimensional realizations for linear SPDEs, which are generalizations of (II.5.11), beyond the simple case that corresponds to Bergomi's variance curve model (see [7]). The main tool in both cases will be a version of the Frobenius Theorem which was introduced in [38]. For this we will restrict ourselves to Markovian models albeit with stochastic volatility (i.e. forward variance models that have local and stochastic volatility).

1 Markovian HJMM Models

From the preceding results we see that basically every HJM model given by an infinite system (one for each *time-to-maturity* T) of forward variance equations as in (II.5.1) can be transferred into a single equation as in (II.5.5) in an infinite dimensional Hilbert space. For our following considerations the specification of the diffusion coefficient σ as given in the equation (II.5.5) is too general and not very useful for applications. Also so far we have only considered an equation for one given initial curve. On the other hand we do not want to restrict ourselves to the purely time-homogeneous case by sacrificing the possibility of considering a stochastic diffusion coefficient. A reasonable compromise is given by the class of forward variance models with stochastic volatility evolving in the state space $\mathcal{H} := H \times \mathbb{R}^m$ for some $m \in \mathbb{N}$, where H satisfies the

Assumptions (**H1**) and (**H2**) given above (we adopted this setting from [40]). On this space $\mathcal{H}$ and under the standing assumption that $\mathbb{P}$ is a (local) martingale measure for the market of variance swaps, we consider the (general) stochastic equation given by

$$\begin{cases} dh_t = (Ah_t + \alpha(h_t))\,dt + \sum_{j=1}^d \Sigma^j(h_t)\,d\beta_t^j, \\ h_0 \in \mathcal{U} := U \times \mathbb{R}^m, \end{cases} \tag{1.1}$$

where U is some open convex set in H, A is a linear operator on $\mathcal{H}$ with domain $D(A)$ generating a strongly continuous semigroup $\mathcal{S}$ and $\alpha, \Sigma_1, ..., \Sigma_d$ are smooth vector fields on $\mathcal{H}$. If we understand (1.1) as a model for the forward variances with stochastic volatility, we will set

$$A := \begin{pmatrix} \frac{d}{dx} & 0 \\ 0 & 0 \end{pmatrix}, \quad \alpha(h) := \begin{pmatrix} 0 \\ a(h) \end{pmatrix}, \tag{1.2}$$

for some smooth map $a : \mathcal{H} \to \mathbb{R}^m$, such that the drift-condition (see Proposition II.4.4) is satisfied. In this case, the first H-valued coordinate process u of the solutions $h = (u, Y)$ will be interpreted as the forward variance process and the $\mathbb{R}^m$-valued second coordinate Y as the (stochastic) volatility of the forward variance. Note that as d/dx generates the shift-semigroup $\{S_t \,|\, t \in \mathbb{R}_+\}$ by (**H2**), A given in (1.2) generates the strongly continuous semigroup $\{\mathcal{S}_t \,|\, t \in \mathbb{R}_+\}$, with $t \mapsto \mathcal{S}_t(u, Y) = (S_t u, Y)$ on $\mathcal{H}$ and $D(A) = D(d/dx) \times \mathbb{R}^m$. Thus, setting $(\sigma^i, c^i) := (\Sigma^{i1}, \Sigma^{i2})$, for $i = 1, ..., d$, we arrive at the coordinate representation of equation (1.1) given by

$$\begin{cases} du_t = \frac{d}{dx} u_t\,dt + \sum_{j=1}^d \sigma^j(u_t, Y_t)\,d\beta_t^j, \\ dY_t = a(u_t, Y_t)\,dt + \sum_{j=1}^d c^j(u_t, Y_t)\,d\beta_t^j, \\ (u_0, Y_0) \in \mathcal{U} := U \times \mathbb{R}^m. \end{cases} \tag{1.3}$$

2 The Frobenius Theorem

In this section we give a short summary of the results of [40] (where everything can be found in full detail) that will be used in the following. The central result will be the Frobenius Theorem (see Theorem 2.2) that will ultimately yield equivalent conditions for the existence of *generic finite dimensional realizations* for the (general) SPDE (1.1). The analysis will be carried out in the Frèchet space $D(A^\infty)$ as A will generally be unbounded on $\mathcal{H}$, but a bounded linear operator on $D(A^\infty)$, which is a necessary prerequisite for this Frobenius Theorem.

If (1.1) admits such a generic finite dimensional realization, there will be in particular (see [40, Definition 2.5]) for any $h_0 \in \mathcal{U} \cap D(A^\infty)$ an open neighborhood $\mathcal{V}$ around h_0, an open set V in $\mathbb{R}_+^n$ and a C^∞-map $\Gamma : V \times \mathcal{V} \to \mathcal{U} \cap D(A^\infty)$, that is an immersion for each fixed $h \in \mathcal{V}$, such that for any $h^* \in \mathcal{V}$ there exists a V-valued diffusion process Z and a strictly positive stopping time τ such that

$$h_{t\wedge\tau} = \Gamma(Z_{t\wedge\tau}, h^*), \quad \text{for all } t \geq 0, \tag{2.1}$$

is the unique local solution of (1.1) with $h_0 = h^*$. The existence of a generic finite dimensional realization around any $h_0 \in \mathcal{U} \cap D(A^\infty)$ is basically equivalent (see Theorem 2.2) to the existence of a (weak) foliation of finite dimensional submanifolds $(\mathcal{M}_h)_{h\in\mathcal{V}}$ with boundary of $\mathcal{H}$ such that for any $h^* \in \mathcal{V}$ the solution of the stochastic equation (1.1) with initial value h^* satisfies

$$h_{t\wedge\tau} \in \mathcal{M}_{h^*}, \quad \text{for all } t \geq 0.$$

In fact, in this case each leaf $\mathcal{M}_{h^*}$ can be defined by the parametrization $\Gamma(\cdot, h^*)$, which gives necessity. Sufficiency follows under some mild conditions from the Frobenius Theorem, which is cited in Theorem 2.2.

For a *given* finite dimensional submanifold $\mathcal{M} \subset \mathcal{U} \cap \mathcal{D}(A)$ with boundary of $\mathcal{H}$ and an initial curve $h^* \in \mathcal{M}$, necessary and sufficient conditions for the invariance of $\mathcal{M}$ with respect to the (weak) solution of (1.1) with initial value $h^* \in \mathcal{M}$, are given by the well-known consistency conditions formulated for the drift and diffusion coefficient of (1.1) in the Stratonovich representation

$$\Xi(h) := Ah + \alpha(h) - \frac{1}{2}\sum_{j=1}^{d} D\Sigma_j(h)\Sigma_j(h) \in T_h\mathcal{M}, \tag{2.2}$$

$$\Sigma_j(h) \in T_h\mathcal{M}, \quad j = 1, ..., d, \tag{2.3}$$

for all $h \in \mathcal{M}$, where for the boundary elements $h \in \partial\mathcal{M}$, $\Xi(h)$ is inward pointing and the Σ_j, $j = 1, ..., d$ are parallel to the boundary, see [40, Theorem 1.2] and the references therein.

Notice that (2.2) and (2.3) are purely geometric conditions on the Stratonovich coefficients of (1.1) and we have

$$D_h := \langle \Xi(h), \Sigma_1(h), ..., \Sigma_d(h) \rangle \subset T_h\mathcal{M}, \quad \forall h \in \mathcal{M}.$$

Accordingly, in the language of classical Frobenius Theory, $D := (D_h)_{h\in\mathcal{M}}$ is a *tangent distribution* of $\mathcal{M}$ and if there exists an *integral manifold* $\mathcal{N} \subset \mathcal{M}$ of the distribution D, then the conditions (2.2) and (2.3) will be satisfied for every $h \in \mathcal{N}$ and accordingly $\mathcal{N}$ will be left invariant by (1.1) for any $h_0 \in \mathcal{N}$. Thus, at least formally, it is clear how to proceed if there is a-priori no candidate foliation $(\mathcal{M}_h)_{h\in\mathcal{V}}$ of submanifolds with boundary of $\mathcal{H}$ and the question is, whether the stochastic equation (1.1) admits a finite dimensional realization.

If the operator A is bounded in $\mathcal{H}$ the classical Frobenius Theorem can be used as in [13]. However this poses some strong restrictions on the considered Hilbert space and hence on the diffusion coefficients of (1.1). Damir Filipović and Josef Teichmann (see [38, Theorem 3.9]) solved this problem by introducing a Frobenius Theorem on the Frèchet space $D(A^\infty)$ by using the boundedness of A on $D(A^\infty)$. By doing so, the problem of finding a generic finite dimensional realization is shifted to the Frèchet space $D(A^\infty)$, which however under some mild conditions means no loss of generality (see [39, Theorem 3.2]). Before citing this version of the Frobenius Theorem we need some concepts regarding the analysis on Frèchet spaces, which we summarize in the following Definition (which is [40, Definition 3.11], a full account can be found in [38, Section 2] and references therein).

Definition 2.1. *Let E be a Frèchet space, U an open subset. A* distribution *on U is a collection of vector subspaces $D = \{D_f\}_{f\in U}$ of E. A vector field X on U is said to take values in D if $X(f) \in D(f)$ for $f \in U$. A distribution D on U is said to be* involutive *if, for any two locally given vector fields X, Y with values in D, the Lie bracket $[X, Y]$, given by*

$$[X, Y](f) = DX(f)Y(f) - DY(f)X(f), \quad f \in U$$

takes values in D.

Theorem 2.2. *Let $X_1, ..., X_n$ be smooth vectorfields that are pointwise linearly independent on an open subset $\mathcal{U}$ of $D(A^\infty)$, such that $X_1, ..., X_{n-1}$ admit local flows and X_n a local semiflow and let D be the distribution on $\mathcal{U}$ generated by $X_1, ..., X_n$.*

If

$$\langle \Xi(f), \Sigma_1(f), ..., \Sigma_d(f) \rangle \subset D(f), \quad \forall f \in \mathcal{U}, \tag{2.4}$$

then (1.1) *admits a generic finite dimensional realization around any $h^* \in \mathcal{U}$ if D is involutive. In this case the map in* (2.1) *is given by*

$$V \times \mathcal{V} \ni (u, h^*) \mapsto \Gamma(u, h^*) := Fl_{u_1}^{X_1} \circ ... \circ Fl_{u_n}^{X_n}(h^*), \tag{2.5}$$

where V and $\mathcal{V}$ are open sets in $\mathbb{R}_+^n$ and $D(A^\infty)$ respectively.

Proof. This follows from [40, Theorem 3.14]. □

If the conditions of the last Theorem are satisfied, then for every initial curve $h^* \in \mathcal{U}$ the underlying coordinate diffusion process Z can be found as in [34, Section 6.4] where the invariant submanifold with boundary $\mathcal{M}_{h^*}$ of $\mathcal{H}$ is given by the parametrization $\Gamma(\cdot, h^*)$.

Definition 2.3. *In the setting of Theorem 2.2, we say that* (1.1) *admits a* minimal finite dimensional realization*, if the (constant) dimension of the distribution D agrees on $\mathcal{U}$ with the dimension of the distribution $(\langle \Xi(h), \Sigma_1(h), ..., \Sigma_d(h) \rangle)_{h \in \mathcal{U}}$.*

In the following we will investigate the existence of generic finite dimensional realizations for generalized versions of the SPDEs given in (II.5.6) and (II.5.11) which are both of the form (1.1).

3 Affine realizations

In Section II.5.1 we have looked at some forward variance models implied by spot-variance models with Hull-White extended affine drift. If the spot-variance process ξ is a Hull-White extended diffusion with representation

$$\xi_t = \xi_0 + \int_0^t (\alpha_s + \kappa \xi_s)\, ds + \int_0^t \sigma(\xi_s)\, d\beta_s,$$

where σ is a smooth map (this is a special case of the models we considered in Section II.5.1 and corresponds to the case where $\sigma_t = \sigma(\xi_t)$) , the forward variance process u is a solution to the SPDE (II.5.6) with σ_t replaced by $\sigma(u_t(0))$, that is,

$$\begin{cases} du_t = \frac{d}{dx} u_t\, dt + \sigma(u_t(0)) e^{\kappa \cdot}\, d\beta_t \\ u_0 = h, \end{cases} \tag{3.1}$$

and will be solved for any initial curve $h \in D(d/dx)$ by

$$u_t = S_t u_0 + e^{\kappa \cdot} \tilde{\xi}_t, \tag{3.2}$$

where $\tilde{\xi}_t := \xi_t - h(t)$ is a strong solution to

$$d\tilde{\xi}_t = \kappa \tilde{\xi}_t\, dt + \sigma(\tilde{\xi}_t + h(t))\, d\beta_t, \quad \tilde{\xi}_0 = 0.$$

Note that the diffusion coefficient $\Sigma(u) := \sigma(u(0)) e^{\kappa \cdot}$ is sensitive with respect to the forward curve only trough the linear evaluation map $ev_0(u) = u(0)$. We will now look at generalizations of this situation by considering the Equation (1.1) and assuming that [40, (A1)-(A3)] are satisfied:

(A1) The diffusion coefficients $\Sigma_1, ..., \Sigma_d$ are pointwise linearly independent and of the form

$$\Sigma_i(u, Y) = \phi_i(l(u), Y), \quad 1 \leq i \leq d, \tag{3.3}$$

where $l \in L(H, \mathbb{R}^p)$, for some $p \in \mathbb{N}$, $\phi_i : \mathbb{R}^{p+m} \to D(A^\infty)$, for $i = 1, ..., d$, are smooth and pointwise linearly independent maps. Moreover,

$$a(u, Y) = \phi_0(l(u), Y), \tag{3.4}$$

where $\phi_0 : \mathbb{R}^{p+m} \to \mathbb{R}^m$ is smooth.

(A2) For every $q \geq 0$, the map

$$\Big(l, l \circ \Big(\frac{d}{dx}\Big), ..., l \circ \Big(\frac{d}{dx}\Big)^q\Big) : D\Big(\Big(\frac{d}{dx}\Big)\Big) \to \mathbb{R}^{p(q+1)}$$

is open.

(A3) A is unbounded on $\mathcal{H}$.

Remark 3.1. It follows from Assumption **(A1)** that basically every weak solution of (1.1) is also a strong solution, which follows from [26, Proposition 6.4]

Lemma 3.2. *Given Assumptions **(A1)-(A3)**, the vector fields $\Sigma_1, ..., \Sigma_d$, $[\Xi, \Sigma_j]$, $j = 1, ..., d$, as well as all multiple Lie Brackets on $\mathcal{U}$ are Banach maps and in particular admit local flows. Ξ is not a Banach map but generates a local semi-flow.*

Proof. This is [40, Lemma 3.17, Lemma 3.20 and Theorem 3.19] □

Under these assumptions, Damir Filipović and Josef Teichmann found in [40] the striking result, that if the SPDE (1.1) with $m = 0$ admits a generic finite dimensional realization, then it is necessarily affine:

Theorem 3.3. *Let $\mathcal{O}$ be an open subset of $\mathcal{U} \cap D(A^\infty)$ such that on $\mathcal{O}$ the conditions of Theorem 2.2 are satisfied, then there exist pointwise linearly independent vector fields $\lambda_1, ..., \lambda_{N_{LA}-1} \in C^\infty(\mathcal{O}, D(A^\infty))$ such that*

$$D_{LA}(u, Y) = \langle \Xi(u, Y), \lambda_1(Y), ..., \lambda_{N_{LA}-1}(Y) \rangle$$

and

$$\Sigma_j(u, Y) \in \langle \lambda_1(Y), ..., \lambda_{N_{LA}-1}(Y) \rangle, \quad \textit{for all } (u, Y) \in \mathcal{O}. \tag{3.5}$$

Proof. This follows from [40, Theorem 4.5]. □

In particular if $m = 0$, then the vector fields $\lambda_1, ..., \lambda_{N_{LA}-1}$ are necessarily constant and thus lead by Theorem 2.2 to an affine realization. In the following we will assume that (3.5) holds for constant vector fields $\lambda_1, ..., \lambda_d$ as in the case of a minimal realization (i.e., $N_{LA} = d+1$, see Definition 2.3). In the next chapters we will look at more general models.

(A4) There are linearly independent vectors $\lambda_1, ..., \lambda_d \in D(A^\infty)$ such that the diffusion coefficients $\Sigma_1, ..., \Sigma_d$ satisfy

$$\Sigma_j(u, Y) \in \langle \lambda_j \rangle, \quad \text{for all } (u, Y) \in \mathcal{U} \cap D(A^\infty),\ j = 1, ..., d. \tag{3.6}$$

Combined with **(A1)** this implies that the diffusion coefficients satisfy

$$\Sigma_j(u,Y) = \phi_j(l(u),Y)\lambda_j, \quad \forall (u,Y) \in \mathcal{V},\ j = 1,...,d, \tag{3.7}$$

where l is as in **(A1)** and each ϕ_j is a scalar field on $\mathbb{R}^{m+p}$. Under these assumptions we arrive at the following coordinate representation corresponding to (1.3)

$$\begin{cases} du_t = \frac{d}{dx}u_t\,dt + \sum_{j=1}^d \sigma^j(u_t,Y_t)\,d\beta_t^j, \\ dY_t = a(u_t,Y_t)\,dt + \sum_{j=1}^d c^j(u_t,Y_t)\,d\beta_t^j, \\ (u_0,Y_0) \in \mathcal{U} := U \times \mathbb{R}^m, \end{cases} \tag{3.8}$$

where

$$\Sigma_j(u,Y) = \begin{pmatrix} \sigma^j(u,Y) \\ c^j(u,Y) \end{pmatrix} = \phi_j(l(u),Y)\begin{pmatrix} \lambda_j^1 \\ \lambda_j^2 \end{pmatrix}, \quad j = 1,...,d \tag{3.9}$$

and a as defined in (3.4). Now, in order to apply Theorem 2.2 to the distribution generated by the vector fields $\Xi, \lambda_1, ..., \lambda_d$ on an open subset $\mathcal{V}$ of $\mathcal{U} \cap D(A^\infty)$ we have to ensure that the vector fields are pointwise linearly independent and that the distribution is involutive. We have the following lemma (cf. [40, Remark 5.4]) that ensures that we can replace Ξ with $\pi := A + \alpha$, where A and α were defined in (1.2).

Lemma 3.4. *We have*

$$\langle \Xi(u,Y), \lambda_1, ..., \lambda_d \rangle = \langle \pi(u,Y), \lambda_1, ..., \lambda_d \rangle, \quad \text{for all } (u,Y) \in \mathcal{U} \cap D(A^\infty), \tag{3.10}$$

where

$$\pi := A + \alpha. \tag{3.11}$$

Proof. The claim follows upon showing that

$$\Xi(u,Y) - \pi(u,Y) = -\frac{1}{2}\sum_{j=1}^d D\Sigma_j(u,Y)\Sigma_j(u,Y) \in \langle \lambda_1, ..., \lambda_d \rangle.$$

But this follows from the form of the diffusion coefficients specified in (3.7) as

$$D\Sigma_j(u,Y)\Sigma_j(u,Y) \quad = \quad \zeta_j(u,Y)\lambda_j,$$

where $\zeta_j(u,Y) \in \mathbb{R}$ is given by

$$\zeta_j(u,Y) = D_u\phi_j(l(u),Y) \cdot \Sigma_j^1(u,Y) + D_Y\phi_j(l(u),Y) \cdot \Sigma_j^2(u,Y).$$

□

Hence we have a candidate distribution D_{LA} given by

$$D_{LA}(h) := \langle \pi(h), \Sigma_1, ..., \Sigma_d \rangle, \quad h \in \mathcal{U} \cap D(A^\infty), \tag{3.12}$$

satisfying (2.4) in Theorem 2.2. If the remaining conditions of Theorem 2.2 are satisfied then it follows from (2.5) that the solution of (3.8) for every $h^* \in \mathcal{U}$ stays in the submanifold with boundary of $\mathcal{H}$ given by the parametrization

$$\Gamma(z,h^*) = Fl_{z^0}(h^*)^\pi + \sum_{i=1}^d z_i\lambda_i.$$

We will now use Lemma 3.4 in accordance with (2.2) to derive necessary conditions on the drift a in (3.8) and the vector fields $\lambda_1, ..., \lambda_d$ in (3.9) such that the distribution D_{LA} is involutive.

Lemma 3.5. *The distribution $D_{LA}(h) = \langle \pi(h), \lambda_1, ..., \lambda_d \rangle$ on some open subset $\mathcal{V}$ of $\mathcal{U} \cap D(A^\infty)$ is involutive if and only if the following two conditions are satisfied:*

1. *The drift coefficient $a : H \times \mathbb{R}^m \to \mathbb{R}^m$ in (3.8) is constant in the first argument and affine in the second, i.e. there is a linear map $\kappa \in L(H \times \mathbb{R}^m, \mathbb{R}^m)$ and $\bar{a} \in \mathbb{R}^m$ such that*
$$a(h) = \bar{a} + \kappa h. \tag{3.13}$$

2. *There is a matrix $B \in \mathbb{R}^{d \times d}$, such that*
$$\begin{pmatrix} \frac{d}{dx}\lambda_i^1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \kappa\lambda_i \end{pmatrix} = \sum_{j=1}^{d} B^{ij}\lambda_j, \quad i = 1, ..., d. \tag{3.14}$$

Proof. Recalling Definition 2.1, it follows that D_{LA} is involutive if and only if $[\pi, \lambda_i](h) \in \langle \pi(h), \lambda_1, ..., \lambda_d \rangle$ for all $h \in \mathcal{V}$ and $i = 1, ..., d$. This is equivalent to the existence of $d+1$ coordinate mappings $B^{i0}, ..., B^{id}$ such that

$$[\pi, \lambda_i](h) = \begin{pmatrix} \frac{d}{dx}\lambda_i^1 \\ Da(h)\lambda_i \end{pmatrix} = B^{i0}(h) \begin{pmatrix} \frac{d}{dx}h^1 \\ a(h) \end{pmatrix} + \sum_{j=1}^{d} B^{ij}(h) \begin{pmatrix} \lambda_i^1 \\ \lambda_i^2 \end{pmatrix}. \tag{3.15}$$

Now, as this has to hold on the open set $\mathcal{V}$, we can conclude from Assumption **(A3)** as in [40, Lemma 4.2] that $B^{i0}(h) = 0$ for all $h \in \mathcal{V}$ as otherwise this would imply that $\frac{d}{dx}$ is a Banach map (which it is not by Assumption **(A3)** and [40, Lemma 3.2]). Thus by looking at the first coordinate of the Equation (3.15)

$$\frac{d}{dx}\lambda_i^1 = \sum_{j=1}^{d} B^{ij}(h)\lambda_i^1,$$

we conclude that B^{ij} are in fact constants. Which in turn gives that

$$Da(h)\lambda_i = \sum_{j=1}^{d} B^{ij}\lambda_i^2$$

and hence a is necessarily an affine map. This gives the first claim. The second claim follows now from (3.15) applied to a given by (3.13) and recalling that $B^{i0}(h) = 0$ for all $h \in \mathcal{V}$ and that B^{ij} are constants. □

Remark 3.6. If the conditions of Lemma 3.5 are satisfied, $\lambda_1^1, ..., \lambda_d^1$ are necessarily *quasi-exponentials*, that is,
$$\lambda^1 = e^{B\cdot}\lambda^1(0), \tag{3.16}$$
where B is the matrix given in (3.14) and $\lambda^1 := (\lambda_1^1, ..., \lambda_d^1)$ (understood as a column vector). The importance of quasi-exponentials for finite dimensional realizations is well-known for forward-interest rate models and first recognized in [13], see for example [13, Remark 5.1, Proposition 6.1].

To apply Theorem 2.2 the dimension of the distribution must be (finite and) constant on $\mathcal{V}$. If we assume that (3.14) holds, then this is equivalent to π being pointwise linearly independent of $\lambda_1, ..., \lambda_d$ on $\mathcal{V}$. Hence we have the following Lemma.

Lemma 3.7. *Assuming that the conditions of Lemma 3.5 hold, the dimension of the distribution D_{LA} on $\mathcal{V}$ is constant if and only if $\mathcal{V}$ does not intersect with the* singular set $\mathfrak{S}$ *given by*

$$\mathfrak{S} := \tilde{A}^{-1}(\langle \lambda_1, ..., \lambda_d \rangle) - \tilde{A}^{-1}(\begin{pmatrix} 0 \\ \bar{a} \end{pmatrix}). \tag{3.17}$$

In particular, $\mathfrak{S} = \emptyset$ *if and only if*

$$\kappa^{-1}\bar{a} \notin \Big\{ \Big(\mathcal{U} \cap D(A^{\infty}) \Big) \cap \Big(\frac{d}{dx}^{-1}(\{0\}) \times \mathbb{R}^m \Big) \Big\}. \tag{3.18}$$

Here $\tilde{A}$ *is the linear vector field on* $\mathcal{V}$ *given by*

$$\tilde{A}h := \Big(A + \tilde{\kappa} \Big) h := Ah + \begin{pmatrix} 0 \\ \kappa h \end{pmatrix}, \tag{3.19}$$

with A *and* κ *given in* (1.2) *and* (3.13) *respectively.*

Proof. As argued above the claim is equivalent to

$$\pi h \notin \langle \lambda_1, ..., \lambda_d \rangle \quad \text{for all } h \in \mathcal{V}.$$

Accordingly the singular set is given by

$$\mathfrak{S} = \{ h \in \mathcal{U} \cap D(A^{\infty}) \,|\, \pi(h) \in \langle \lambda_1, ..., \lambda_d \rangle \}, \tag{3.20}$$

which by recalling (3.11) gives (3.17). The second claim follows from Lemma 3.5 as it shows that

$$\langle \lambda_1, ..., \lambda_d \rangle \subset \tilde{A}^{-1}(\langle \lambda_1, ..., \lambda_d \rangle)$$

and accordingly $\mathfrak{S} = \emptyset$ if and only if $\tilde{A}^{-1}(\begin{pmatrix} 0 \\ a \end{pmatrix}) = \emptyset$. Thus the claim follows from noting that

$$\tilde{A}^{-1}(\begin{pmatrix} 0 \\ \bar{a} \end{pmatrix}) \quad = \quad \{ h \in \mathcal{U} \cap D(A^{\infty}) \,|\, \frac{d}{dx} h^1 = 0 \ \text{ and } \ \kappa h = \bar{a} \}.$$

□

Lemma 3.8. *If* $\mathfrak{S} \neq \emptyset$*, then for all* $h \in \mathfrak{S}$*,* $\mathcal{M}_h := h + \langle \lambda_1, ..., \lambda_d \rangle$ *satisfies the invariance conditions of* (2.2) *and* (2.3)*.*

Proof. Let $h \in \mathfrak{S}$. Then $\mathcal{M}_h \subset \mathfrak{S}$ since it follows from Lemma 3.5 that $\langle \lambda_1, ..., \lambda_d \rangle \subset \tilde{A}^{-1}(\langle \lambda_1, ..., \lambda_d \rangle)$. Further, by Lemma 3.4 and (3.7) the validity of (2.2) and (2.3) follows if for all $f \in \mathcal{M}_h$ we have $\pi f \in \langle \lambda_1, ..., \lambda_d \rangle$ but this is just the defining condition of $\mathfrak{S}$ given in (3.20). □

Theorem 3.9. *If* $\mathcal{V} = \mathcal{U} \cap D(A^{\infty}) \setminus \mathfrak{S}$ *and the conditions of Lemma 3.5 are satisfied then Equation* (3.8) *with diffusion vector fields given by* (3.9) *admits a generic finite dimensional realizations around any* $h^* \in \mathcal{V}$*. In particular, there is a* $\mathbb{R}^d$*-valued time-inhomogeneous diffusion process* Z *such that*

$$h_t = \Gamma((t, Z_t), h^*) = Fl_t^{\pi}(h^*) + \sum_{j=1}^{d} Z_t^j \lambda_j, \tag{3.21}$$

is the unique continuous local solution to (3.8) *with* $h_0 = h^*$. *The process* Z *satisfies*

$$dZ_t = B^T Z_t \, dt + \sum_{j=1}^{d} \rho_j(t, Z_t) e_j \, d\beta_t^j, \quad Z_0 = 0,$$

where B *was given in* (3.14), $e_1, ..., e_d$ *are the standard basis vectors of* $\mathbb{R}^d$ *and*

$$\rho_j(t, z) = \phi_j(l(\Gamma^1((t, z), h^*)), \Gamma^2((t, z), h^*)), \quad j = 1, ..., d.$$

If $\mathcal{V} = \mathfrak{S}$, *the solution* h_t *given in* (3.21) *can be represented by*

$$h_t = \tilde{\Gamma}(\tilde{Z}_t, h^*) = h^* + \sum_{j=1}^{d} \tilde{Z}_t^j \lambda_j, \tag{3.22}$$

where $\tilde{Z}$ *is a* $\mathbb{R}^d$*-valued time-homogeneous diffusion process satisfying*

$$d\tilde{Z}_t = (B^T \tilde{Z}_t + \Pi(h^*)) \, dt + \sum_{j=1}^{d} \tilde{\rho}_j(\tilde{Z}_t) e_j \, d\beta_t^j, \quad \tilde{Z}_0 = 0,$$

where $\Pi(h^*)$ *is the column vector comprised of the coordinates of* $\pi(h^*)$ *with respect to the basis* $\lambda_1, ..., \lambda_d$ *and*

$$\tilde{\rho}_j(z) = \phi_j(l(\Gamma^1(z, h^*)), \Gamma^2(z, h^*)), \quad j = 1, ..., d.$$

Proof. We consider first the case $\mathcal{V} = \mathcal{U} \cap D(A^\infty) \setminus \mathfrak{S}$. Then from the Assumptions **(A1)-(A4)** and Lemma 3.5 and 3.7 it follows that the conditions of Theorem 2.2 are satisfied. As for (3.21) we proceed as in [40, Theorem 5.3] and [34, Section 6.4]. For $h^* \in \mathcal{V}$ we can construct a submanifold $\mathcal{M}_{h^*}$ by the parametrization $\Gamma(\cdot, h^*)$ (cf. (2.5)) given by

$$\Gamma(z, h^*) = Fl_{z_0}^{\pi}(h^*) + \sum_{i=1}^{d} z_i \lambda_i, \quad z = (z_0, ..., z_d) \in [0, \epsilon) \times V,$$

for some $\epsilon > 0$ and some open neighborhood V of zero in $\mathbb{R}^d$. Now, for this initial curve h^* we are in the setting of [34] as conditions (2.2) and (2.3) are satisfied and accordingly can proceed as in [34, Section 6.4] to construct the coordinate process Z. There are mappings $\tilde{b}$ and $\tilde{\rho}$ from $\mathbb{R}^{d+1}$ into $\mathbb{R}^{d+1}$ and $\mathbb{R}^{d \times d+1}$, respectively, that are uniquely determined through

$$\pi(\Gamma(z, h^*)) = D_z\Gamma(z, h^*) \cdot \tilde{b}(z), \tag{3.23}$$

$$\Sigma^j(\Gamma(z, h^*)) = D_z\Gamma(z, h^*) \cdot \tilde{\rho}^j(z), \quad j = 1, ..., d, \tag{3.24}$$

for all $z \in [0, \epsilon) \times V$. Here (3.23) is equivalent to

$$\pi(Fl_{z_0}^{\pi}(h^*)) + \sum_{i=1}^{d} z_i \tilde{A} \lambda_i = D_{z_0} Fl_{z_0}^{\pi}(h^*) \tilde{b}^0(z) + \sum_{i=1}^{d} \tilde{b}^i(z) \lambda_i,$$

where $\tilde{A}$ is given in (3.19). Thus using the Representation (3.14) gives

$$\tilde{b}^0(z) \equiv 1, \quad \tilde{b}^j(z) = \sum_{i=1}^{d} z_i B^{ij}, \quad z \in [0, \epsilon) \times V \quad j = 1, ..., d.$$

In the same way, by considering the Representation (3.7), we get from (3.24) that

$$\tilde{\rho}^{j,0}(z) \equiv 0, \quad \tilde{\rho}^{ji}(z) = \delta_{ij}\phi_j(l(\Gamma^1(z,h^*)), \Gamma^2(z,h^*)), \quad i,j = 1,...,d,$$

where $\delta_{ij} = 1$ if $i = j$ and else zero. We claim that the $V \subset \mathbb{R}^d$-valued diffusion Z is given by $Z_0 = 0$ and

$$dZ_t^j = b^j(Z_t)\,dt + \rho^j(t, Z_t)\,d\beta_t^j, \quad j = 1,...,d,$$

where the mappings b from $\mathbb{R}^d$ into $\mathbb{R}^d$ and ρ from $\mathbb{R}^{d+1}$ into $\mathbb{R}^d$ are defined by

$$\begin{aligned} b^j(z) &:= \tilde{b}^j(z) = \sum_{i=1}^d z_i B^{ji}, \\ \rho^j(t,z) &:= \tilde{\rho}^{jj}(t,z) = \phi_j(l(\Gamma^1((t,z),h^*)), \Gamma^2((t,z),h^*)) \end{aligned} \tag{3.25}$$

for $(t,z) \in [0,\epsilon) \times V$ and $j = 1,...,d$ is such that (3.21) is a local solution to (3.8). Indeed, by linearity of (stochastic) integrals, setting $\Gamma_t := \Gamma((t,Z_t),h^*) = Fl_t^\pi(h^*) + \sum_{j=1}^d Z_t^j \lambda_j$, we have,

$$\begin{aligned} \Gamma_t &= Fl_t^\pi(h^*) + \int_0^t \sum_{j=1}^d b^i(Z_s)\lambda^i\,ds + \sum_{j=1}^d \int_0^t \rho^j(s,Z_s)\lambda_j\,d\beta_s^j \\ &= h^* + \int_0^t \Big(\pi(Fl_s^\pi(h^*)) + \tilde{A}\sum_{i=1}^d Z_s^i \lambda^i\Big)\,ds + \sum_{j=1}^d \int_0^t \phi_j(l(\Gamma_s^1), \Gamma_s^2)\lambda_j\,d\beta_s^j \\ &= h^* + \int_0^t \Big(A(\Gamma_s) + \alpha(\Gamma_s)\Big)\,ds + \sum_{j=1}^d \int_0^t \Sigma_j(\Gamma_s)\,d\beta_s^j \end{aligned}$$

and thus the first claim follows from uniqueness of strong solutions to SPDEs. The second claim follows from (local) uniqueness by using the Lemma 3.8 and applying the same analysis as above for the parametrization $\tilde{\Gamma}((z_1,...,z_d),h^*) = h^* + \sum_{j=1}^d z_j\lambda_j$ and recalling that for $h^* \in \mathfrak{S}$ we have $\pi(h^*) \in \mathfrak{S}$. □

Remark 3.10. The Representation given in (3.21) is called *generic affine realization* as by the second part of Theorem 3.9 it can be seen as a generalization of the *affine realizations* given in (3.22) that occur for initial curves h^* that lie in the singular set $\mathfrak{S}$.

The coordinate representation of (3.21) for $h^* = (u^*, Y^*)$ is given by

$$\begin{cases} u_t &= (Fl_t^\pi)^1(u^*) + \sum_{j=1}^d Z_t^j \lambda_j^1 \\ Y_t &= (Fl_t^\pi)^2(u^*,Y^*) + \sum_{j=1}^d Z_t^j \lambda_j^2, \end{cases} \tag{3.26}$$

where

$$Fl_t^\pi(u^*,Y^*) = \begin{pmatrix} (Fl_t^\pi)^1(u^*) \\ (Fl_t^\pi)^2(u^*,Y^*) \end{pmatrix} = \begin{pmatrix} S_t u^* \\ \int_0^t (\kappa Fl_s^\pi(u^*,Y^*) + \bar{a})\,ds \end{pmatrix}.$$

Accordingly we have in particular constructed simultaneously generic finite dimensional realizations for u_t and Y_t such that both have the same d-dimensional coordinate process Z, which was possible due to the special compatibility condition of the drift coefficient Y_t and the choice of the $\lambda_1,...,\lambda_d$ given in Lemma 3.5 by the conditions (3.13) and (3.14) respectively. These conditions can be considerably relaxed when the forward variance process u and the parameter process Y are assumed to be driven by independent Brownian motions. Compare this situation to forward interest rate models with a stochastic volatility as discussed in [12], in which the stochastic volatility process Y is assumed to be an autonomous process. We see that under the given Assumptions **(A1)-(A4)** we have strengthened [12, Proposition 5.2]) into two directions. First, the stochastic volatility process Y can have drift and diffusion coefficients that depend on the state of the forward variance, and second, the requirement that the vectors $\lambda_1, ..., \lambda_d$ are quasi-exponentials is not only sufficient but also necessary.

However, we saw in Lemma 3.5 that in the present case the diffusion coefficient a of Y is necessarily of the affine form given in (3.13). In the next section we will see that this condition can be relaxed when considering an independent parameter process.

3.1 Independent Parameter process

In the previous subsection we have constructed a generic finite dimensional realization for the joint process $h_t = (u_t, Y_t)$ given in (3.8) such that the drift coefficient a of Y_t and the vectors $\lambda_1, ..., \lambda_d$ satisfied the conditions of Lemma 3.5. Under the conditions of Theorem 3.9 we arrived at the coordinate Representations (3.26) for u_t and Y_t. In particular, we see that the m-dimensional process Y_t could be represented by

$$Y_t = (Fl_t^{\pi})^2(u^*, Y^*) + \sum_{j=1}^{d} Z_t^j \lambda_j^2$$

and thus leaves the set $\mathcal{M}^Y(u^*, Y^*)$ given by

$$\mathcal{M}^Y(u^*, Y^*) := \Big\{ (Fl_t^{\pi})^2(u^*, Y^*) + \sum_{j=1}^{d} z^j \lambda_j^2 \,\Big|\, t \in \mathbb{R}_+, z \in \mathbb{R}^d \Big\}$$

invariant. This restriction on the parameter process is a result of the assumption that both process are driven by the same Brownian motions and of the choice of the diffusion coefficients given in (3.7). Having this in mind, we will in the following look at the system

$$\begin{cases} du_t = \frac{d}{dx} u_t \, dt + \sum_{j=1}^{d} \phi_j(l(u_t), Y_t) \lambda_j^1 \, d\beta_t^j, \\ dY_t = a(u_t, Y_t) \, dt + \sum_{j=d+1}^{d+m} c^j(u_t, Y_t) \, d\beta_t^j, \\ (u_0, Y_0) \in \mathcal{U} := U \times \mathbb{R}^m. \end{cases} \tag{3.27}$$

We will now repeat the analysis of the previous subsection for the candidate distribution D_{LA} given by

$$D_{LA}(h) := \Big\langle \Xi(h), \begin{pmatrix} \lambda_1^1 \\ 0 \end{pmatrix}, ..., \begin{pmatrix} \lambda_d^1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ e_1 \end{pmatrix}, ..., \begin{pmatrix} 0 \\ e_m \end{pmatrix} \Big\rangle, \quad h \in \mathcal{U} \cap D(A^\infty), \tag{3.28}$$

where $e_1, ..., e_m$ are the standard basis vectors of $\mathbb{R}^m$. It is evident that D_{LA} satisfies (2.4). We have the following result corresponding to Lemma 3.4. Recall the definition of the linear map A given in (1.2).

Lemma 3.11. *We have*

$$D_{LA}(h) = \Big\langle Ah, \begin{pmatrix} \lambda_1^1 \\ 0 \end{pmatrix}, ..., \begin{pmatrix} \lambda_d^1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ e_1 \end{pmatrix}, ..., \begin{pmatrix} 0 \\ e_m \end{pmatrix} \Big\rangle, \quad h \in \mathcal{U} \cap D(A^\infty). \tag{3.29}$$

Proof. It follows immediately as in Lemma 3.4 that

$$D_{LA}(h) = \left\langle \pi(h), \begin{pmatrix} \lambda_1^1 \\ 0 \end{pmatrix}, ..., \begin{pmatrix} \lambda_d^1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ e_1 \end{pmatrix}, ..., \begin{pmatrix} 0 \\ e_m \end{pmatrix} \right\rangle, \quad h \in \mathcal{U} \cap D(A^\infty),$$

with π given as in (3.11) by

$$\pi(h) = Ah + \begin{pmatrix} 0 \\ a(h) \end{pmatrix}$$

and thus the claim follows from $a(h) \in \mathbb{R}^m$ for all $h \in \mathcal{U} \cap D(A^\infty)$. □

Thus we are now in a setting where we can look at the processes u_t and Y_t separately. It follows in particular, that we can write (3.29) as

$$D_{LA}(h) = \langle \frac{d}{dx} h^1, \lambda_1^1, ..., \lambda_d^1 \rangle \times \mathbb{R}^m, \quad h \in \mathcal{U} \cap D(A^\infty), \tag{3.30}$$

confirming the intuition that now the system (3.27) admits a generic finite dimensional realization if and only if the forward variance process does so for each fixed value of $Y_t = y$ (this corresponds to the *parametrized model* considered in [12, Definition 2.3]).

Lemma 3.12. *The distribution on D_{LA} on $\mathcal{U} \cap D(A^\infty)$ given by* (3.29) *is involutive if and only if D_{LA}^1 is involutive on $U \cap D((d/dx)^\infty)$, where*

$$D_{LA}^1(u) = \langle \frac{d}{dx} u, \lambda_1^1, ..., \lambda_d^1 \rangle, \tag{3.31}$$

which in turn is equivalent to the existence of a matrix $B \in \mathbb{R}^{d \times d}$ such that

$$\frac{d}{dx} \lambda_i^1 = \sum_{j=1}^d B^{ij} \lambda_j^1, \quad i = 1, ..., d. \tag{3.32}$$

Proof. As in Lemma 3.5 the equivalent condition for D_{LA} being involutive on $\mathcal{U} \cap D(A^\infty)$ is given by

$$\left[\begin{pmatrix} \frac{d}{dx} & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \lambda_i^1 \\ 0 \end{pmatrix} \right] h = \begin{pmatrix} \frac{d}{dx} \lambda_i^1 \\ 0 \end{pmatrix} \in \left\langle \begin{pmatrix} \lambda_1^1 \\ 0 \end{pmatrix}, ..., \begin{pmatrix} \lambda_d^1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ e_1 \end{pmatrix}, ..., \begin{pmatrix} 0 \\ e_m \end{pmatrix} \right\rangle$$

from which immediately the first claim follows and the second claim follows just as (3.15) in Lemma 3.5. □

Regarding the singular set $\mathfrak{S}$ it follows immediately from Lemma 3.11 (cf. Lemma 3.7) that

$$\begin{aligned} \mathfrak{S} &= \left\{ h \in \mathcal{U} \cap D(A^\infty) \,\middle|\, Ah \in \left\langle \begin{pmatrix} \lambda_1^1 \\ 0 \end{pmatrix}, ..., \begin{pmatrix} \lambda_d^1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ e_1 \end{pmatrix}, ..., \begin{pmatrix} 0 \\ e_m \end{pmatrix} \right\rangle \right\} \\ &= \{ h \in \mathcal{U} \cap D(A^\infty) \,|\, \frac{d}{dx} h^1 \in \langle \lambda_1^1, ..., \lambda_d^1 \rangle \} \\ &= \{ h \in \mathcal{U} \cap D(A^\infty) \,|\, h^1 \in \langle 1, \lambda_1^1, ..., \lambda_d^1 \rangle \} \end{aligned} \tag{3.33}$$

as $\frac{d}{dx}^{-1}(\langle \lambda_1^1, ..., \lambda_d^1 \rangle)$ is a $d+1$ dimensional linear subspace of $D((d/dx)^\infty)$ since the kernel of $\frac{d}{dx}$ consists of the 1 dimensional subspace given by the constant functions and includes by Lemma 3.12 $\langle \lambda_1^1, ..., \lambda_d^1 \rangle$. The following can be shown just as in Lemma 3.8.

Lemma 3.13. *For all $h \in \mathfrak{S}$,*

$$\mathcal{M}_h := h + \left\langle \begin{pmatrix} \lambda_1^1 \\ 0 \end{pmatrix}, ..., \begin{pmatrix} \lambda_d^1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ e_1 \end{pmatrix}, ..., \begin{pmatrix} 0 \\ e_m \end{pmatrix} \right\rangle$$

satisfies the invariance conditions of (2.2) *and* (2.3).

Finally we can conclude from Lemma 3.11, 3.12 and 3.13 the following special case of Theorem 3.9 which we formulate as a corollary.

Corollary 3.14. *In the setting of Theorem 3.9, for $h^* \in \mathcal{U} \cap D(A^\infty) \setminus \mathfrak{S}$ the (local) solution of* (3.27) *can be represented by*

$$h_t = \Gamma((t, Z_{t,1}, Z_{t,2}), h^*) = Fl_t^A(h^*) + \sum_{j=1}^{d} Z_{t,1}^j \begin{pmatrix} \lambda_j^1 \\ 0 \end{pmatrix} + \sum_{j=1}^{m} Z_{t,2}^j \begin{pmatrix} 0 \\ e_j \end{pmatrix}, \tag{3.34}$$

where the process $Z_{t,1}$ and $Z_{t,2}$ satisfies

$$\begin{aligned} dZ_{t,1} &= B^T Z_{t,1}\, dt + \sum_{j=1}^{d} \rho_j(t, Z_{t,1}, Z_{t,2}) e_j\, d\beta_t^j, \quad Z_{0,1} = 0, \\ dZ_{t,2} &= a(\Gamma((t, Z_{t,1}, Z_{t,2}), h^*))\, dt + \sum_{j=d+1}^{d+m} c_j(\Gamma((t, Z_{t,1}, Z_{t,2}), h^*))\, d\beta_t^j, \quad Z_{0,2} = 0, \end{aligned}$$

with B given in (3.32)*, $e_1, ..., e_d$ are the standard basis vectors of $\mathbb{R}^d$ and*

$$\rho_j(t, z_1, z_2) = \phi_j(l(\Gamma^1((t, z_1, z_2), h^*)), \Gamma^2((t, z_1, z_2), h^*)), \quad j = 1, ..., d.$$

If $h^* \in \mathfrak{S}$ then (3.34) *can be represented by*

$$h_t = \tilde{\Gamma}((\tilde{Z}_{t,1}, Z_{t,2}, h^*) = h^* + \sum_{j=1}^{d} \tilde{Z}_{t,1}^j \begin{pmatrix} \lambda_j^1 \\ 0 \end{pmatrix} + \sum_{j=1}^{m} Z_{t,2}^j \begin{pmatrix} 0 \\ e_j \end{pmatrix}, \tag{3.35}$$

where the process $\tilde{Z}_1$ satisfies

$$d\tilde{Z}_{1,t} = (\Pi(h^*) + B^T \tilde{Z}_{1,t})\, dt + \sum_{j=1}^{d} \rho_j(\tilde{Z}_{t,1}, Z_{t,2}) e_j\, d\beta_t^j, \quad \tilde{Z}_{0,1} = 0,$$

where $\Pi(h^*)$ is the column vector comprised of the coordinates of $\frac{d}{dx} h^*$ with respect to the basis $\lambda_1^1, ..., \lambda_d^1$ and

$$\tilde{\rho}_j(z_1, z_2) = \phi_j(l(\tilde{\Gamma}^1((z_1, z_2), h^*)), \tilde{\Gamma}^2((z_1, z_2), h^*)), \quad j = 1, ..., d.$$

Proof. We arrive at the Representations (3.34) and (3.35) just as we arrived at (3.21) and (3.22) in Theorem 3.9. We just show that (3.34) is indeed a solution of (3.27). Let $(u^*, Y^*) := h^*$ and recalling that

$$Fl_t^A(h^*) = \begin{pmatrix} S_t u^* \\ Y^* \end{pmatrix}$$

if follows that h^1_t satisfies (recall (3.25) and (3.32))

$$\begin{aligned} h^1_t &= S_t u^* + \sum_{j=1}^d Z^j_{t,1}\lambda^1_j \\ &= u^* + \int_0^t \Big(\frac{d}{dx}S_s u^* + \sum_{j=1}^d\sum_{i=1}^d Z^i_{s,1}B^{ji}\lambda^1_j\Big)\,ds + \sum_{j=1}^d \int_0^t \rho_j(s, Z_{s,1}, Z_{s,2})\lambda^1_j\,d\beta^j_s \\ &= u^* + \int_0^t \frac{d}{dx}\Big(S_s u^* + \sum_{i=1}^d Z^i_{s,1}\lambda^1_i\Big)\,ds + \sum_{j=1}^d \int_0^t \rho_j(s, Z_{s,1}, Z_{s,2})\lambda^1_j\,d\beta^j_s \\ &= u^* + \int_0^t \frac{d}{dx}h^1_s\,ds + \sum_{j=1}^d \int_0^t \phi_j(l(h^1_s), h^2_s)\lambda^1_j\,d\beta^j_s \end{aligned}$$

and as it is evident that $h^2_t = Y_t$ the claim is proved. □

Example 3.15 (The generic double mean reverting model)*.* A particularly popular model that is included in the current setup it the *double mean reverting model*, which was introduced (to the best of our knowledge) for forward variance models in [15, Example 3.5]. A specific form of this model, called *double CEV dynamics* is then further investigated in [43]. In its general form it is given in terms of a paramterization by

$$\Gamma(z) := z_3 + (z_1 - z_3)e^{-\kappa\cdot} + (z_2 - z_3)\begin{cases} \frac{\kappa}{\kappa - c}(e^{-c\cdot} - e^{-\kappa\cdot}), & \text{if } \kappa \neq c \\ \kappa\, m(Id(\cdot), e^{-\kappa\cdot}), & \text{if } \kappa = c \end{cases}, \tag{3.36}$$

where the *consistent* (consistent refers here to the validity of the drift-condition) parameter process Z is given by (for notational simplicity we assume a special form of the volatility coefficients, whilst in [15, Example 3.5] the specification is slightly more general)

$$\begin{aligned} dZ^1_t &= \kappa(Z^2_t - Z^1_t)\,dt + \sigma_1(Z_t)\,d\beta^1_t, \\ dZ^2_t &= c(Z^3_t - Z^2_t)\,dt + \sigma_2(Z_t)\,d\beta^2_t, \\ dZ^3_t &= \sigma_3(Z_t)\,d\beta^3_t. \end{aligned}$$

We look first at the case $\kappa \neq c$. By defining $\lambda_1 := e^{-\kappa}$, $\lambda_2 := \frac{\kappa}{\kappa - c}(e^{-c} - \lambda_1)$ and $\lambda_3 := 1 - \lambda_1 - \lambda_2$ it follows that

$$\Gamma(z) = \sum_{i=1}^3 z^i\lambda_i, \quad \frac{d}{dx}\Gamma(z) = \kappa(z^2 - z^1)\lambda_1 + c(z^3 - z^2)\lambda_2. \tag{3.37}$$

Thus $u_t := \Gamma(Z_t)$ satisfies (from the linearity of (stochastic) integrals)

$$\begin{aligned} u_t &= u_0 + \sum_{i=1}^3 (Z^i_t - Z^i_0)\lambda_i \\ &= u_0 + \int_0^t \Big(\kappa(Z^2_s - Z^1_s)\lambda_1 + c(Z^3_s - Z^2_s)\lambda_2\Big)\,ds + \sum_{i=1}^3 \int_0^t \sigma_i(Z_s)\lambda_i\,d\beta^i_s \end{aligned}$$

and using (3.37) we get

$$u_t = u_0 + \int_0^t \frac{d}{dx} u_s \, ds + \sum_{i=1}^{3} \int_0^t \sigma_i(Z_s)\lambda_i \, d\beta_s^i. \tag{3.38}$$

We see that $u_t = \Gamma(Z_t)$ satisfies the drift condition (cf. (II.5.5)) which confirms the above mentioned consistency. Practically and conceptionally this models have two drawbacks. The practical drawback comes from the fact that the possible initial curves must lie in the set

$$\Big\{ \sum_{i=1}^{3} z^i \lambda_i \,|\, z^1 \geq 0, z^2, z^3, \kappa, c > 0 \Big\},$$

where κ and c denote the *parameters* of the directions $\lambda_1, ..., \lambda_3$. The conceptional drawback is that u is not a Markov process in its own filtration. In the following we will generalize this model to overcome both drawbacks. To overcome the second one, we consider now the model

$$u_t = u_0 + \int_0^t \frac{d}{dx} u_s \, ds + \sum_{i=1}^{3} \int_0^t \sigma_i(l_i(u_s))\lambda_i \, d\beta_s^i, \tag{3.39}$$

where $l_1, ..., l_3$ are linear maps as given in Assumption **(A1)**. But now we are in the situation of Corollary 3.14 with

$$m = 0, \quad B^T = \begin{pmatrix} -\kappa & \kappa & 0 \\ 0 & -c & c \\ 0 & 0 & 0 \end{pmatrix}.$$

Accordingly for every u^* in some open subset of $D((d/dx)^\infty)$ we can have

$$u_t = Fl_t^{d/dx} u^* + \sum_{j=1}^{d} \tilde{Z}_s^j \lambda_j, \tag{3.40}$$

where

$$\begin{aligned} d\tilde{Z}_t^1 &= \kappa(\tilde{Z}_t^2 - \tilde{Z}_t^1)\, dt + \sigma_1\Big(l_1(Fl_t^{d/dx} u^* + \sum_{j=1}^{d} \tilde{Z}_s^j \lambda_j)\Big)\, d\beta_t^1, \\ d\tilde{Z}_t^2 &= c(\tilde{Z}_t^3 - \tilde{Z}_t^2)\, dt + \sigma_2\Big(l_2(Fl_t^{d/dx} u^* + \sum_{j=1}^{d} \tilde{Z}_s^j \lambda_j)\Big)\, d\beta_t^2, \\ d\tilde{Z}_t^3 &= \sigma_3\Big(l_3(Fl_t^{d/dx} u^* + \sum_{j=1}^{d} \tilde{Z}_s^j \lambda_j)\Big)\, d\beta_t^3, \end{aligned}$$

with $\tilde{Z}_0^1, \tilde{Z}_0^2, \tilde{Z}_0^3 = 0$.

The case $\kappa = c$ can be treated in just the same way by considering the directions $\lambda_1 := e^{-\kappa}$, $\lambda_2 := \kappa m(Id, e^{-\kappa})$ and $\lambda_3 := 1 - \lambda_1 - \lambda_2$. Notice that the distribution generated by $\frac{d}{dx}, \lambda_1, \lambda_2, \lambda_3$ is involutive as $\frac{d}{dx}\lambda_2 = \kappa\lambda_1 - \kappa\lambda_2$.

4 Exponentially-Affine Realizations

We repeat now the analysis of the previous section by looking at linear models (i.e. linear in the first component) that are generalizations of (II.5.11), and given by

$$\begin{cases} du_t = \frac{d}{dx}u_t\,dt + \sum_{j=1}^d m(\sigma_j(Y_t), u_t)\,d\beta_t^j \\ dY_t = a(u_t, Y_t)\,dt + \sum_{j=1}^d c^j(u_t, Y_t)\,d\beta_t^j, \\ (u_0, Y_0) \in \mathcal{U} := U \times \mathbb{R}^m, \end{cases} \tag{4.1}$$

where U is an open convex set in $H^+ \subset H$ (cf. Lemma II.5.14). Here for each $j = 1, ..., d$, σ_j is assumed to be a smooth map from $\mathbb{R}^m$ into H. From Section II.5.2 it follows that such models correspond in the *time-of-maturity* parametrization $v(t,T) := u_t(T-t)$ to solutions of

$$\begin{cases} dv(t,T) = \sum_{j=1}^d \sigma_j(Y_t)(T-t)v(s,T)\,d\beta_t^j \\ dY_t = a(u_t, Y_t)\,dt + \sum_{j=1}^d c^j(u_t, Y_t)\,d\beta_t^j, \\ (v(0,\cdot), Y_0) \in \mathcal{U} := U \times \mathbb{R}^m, \end{cases} \tag{4.2}$$

and as such can be represented by

$$v(t,T) = v(0,T)\exp\Big(\sum_{j=1}^d \int_0^t \sigma_j(Y_s)(T-s)\,d\beta_s^j - \frac{1}{2}\sum_{j=1}^d \int_0^t \sigma_j^2(Y_s)(T-s)\,ds\Big),$$

which motivates the following slight generalization of Proposition II.5.15. Consider the system

$$\begin{cases} dJ_t = (\frac{d}{dx}J_t - \frac{1}{2}\sum_{j=1}^d m(\sigma_j(\tilde{Y}_t), \sigma_j(\tilde{Y}_t))\,dt + \sum_{j=1}^d \sigma_j(\tilde{Y}_t)\,d\beta_t^j \\ d\tilde{Y}_t = \tilde{a}(J_t, \tilde{Y}_t)\,dt + \sum_{j=1}^d \tilde{c}^j(J_t, \tilde{Y}_t)\,d\beta_t^j, \\ (J_0, \tilde{Y}_0) \in \mathcal{U}^l := U^l \times \mathbb{R}^m, \end{cases} \tag{4.3}$$

where $\tilde{a}(J,Y) := a(\exp(J), Y)$, $\tilde{c}^j(J,Y) := c^j(\exp(J), Y)$ for $j = 1, ..., d$ and $U^l := \log(U)$, which is well-defined as $U \subset H^+$ (cf. Lemma II.5.14).

Corollary 4.1. *Denote by* $(J_t, \tilde{Y}_t)$ *the mild solution of* (4.3) *with initial values given by* $(\log(u^*), y)$*. Then the process* $(\exp(J_t), Y_t)$ *is a mild solution of* (4.1) *for initial values* (u^*, y)*. In particular,* (4.1) *admits a generic finite dimensional realization around any* $(u^*, Y^*) \in \mathcal{U}$ *if and only if* (4.3) *does so around any* $(J^*, \tilde{Y}^*) \in \mathcal{U}^l$*.*

Proof. This first claim follows immediately from Proposition II.5.15 by noting that the map $(\exp, Id_{\mathbb{R}^m}) : \mathcal{H} \to \mathcal{H}^+$ is a diffeomorphism, where $\mathcal{H}^+ := H^+ \times \mathbb{R}^m$ and that (recall (1.2))

$$\begin{pmatrix} \exp \\ Id_{\mathbb{R}^m} \end{pmatrix}_* A = A.$$

The second claim is obvious. □

Thus without loss of generality we can in the sequel look at the somewhat simpler system given in (4.3), which corresponds to models that were considered in the previous section but with the addition of a non-vanishing drift. If we assume that $\tilde{a}$ and $\tilde{c}$ depend on J only through some linear map $l : H \to \mathbb{R}^p$ for $p \in \mathbb{N}$ then the system (4.3) already satisfies (3.3) and (3.4) and accordingly it is only a minor loss of generality to assume that **(A1)-(A3)** holds, which we will do from here on.

Lemma 4.2. *If Assumptions **(A1)-(A3)** hold for* (4.3) *then* $\tilde{c}^j$*, for* $j = 1, ..., d$*, is necessarily constant in the first argument.*

Proof. As the conditions of Theorem 3.3 are satisfied it follows from (3.5) that necessarily

$$\Sigma_j(J,Y) = \begin{pmatrix} \sigma_j(Y) \\ \tilde{c}^j(J,Y) \end{pmatrix} \in \langle \lambda_1(Y), ..., \lambda_d(Y) \rangle, \tag{4.4}$$

for some pointwise linearly independent vector fields $\lambda_1, ..., \lambda_d$ on some open subset $\mathcal{V}$ of $\mathcal{U}^l \cap D(A^\infty)$. Thus the claim follows as $\sigma_j(Y)$ is constant in J. □

As in the previous section we assume that the vector fields $\lambda_1, ..., \lambda_d$ are in fact constant linearly independent vectors in $\mathcal{V}$ and for notational convenience we assume also that **(A4)** holds, that is

$$\Sigma_j(J,Y) = \begin{pmatrix} \sigma^j(Y) \\ \tilde{c}^j(Y) \end{pmatrix} = \phi_j(Y) \begin{pmatrix} \lambda_j^1 \\ \lambda_j^2 \end{pmatrix}, \quad j = 1, ..., d, \tag{4.5}$$

where $\phi_1, ..., \phi_d$ are scalar fields in $\mathbb{R}^m$.

Remark 4.3. Notice that we have so far only assumed that $\lambda_1, ..., \lambda_d$ are linearly independent but left unspecified whether $\lambda_1^1, ..., \lambda_d^1$ or $\lambda_1^2, ..., \lambda_d^2$ share this property.

We know from Example II.5.16 (which corresponds to the Bergomi Model) that for the choice $m = 0$ (i.e. in the absence of stochastic volatility process Y) and $\sigma_j = e^{-\kappa_j \cdot}$, where κ_j is a positive constant, (4.1) admits around any $u^* \in U \cap D((d/dx)^*)$ a generic finite dimensional realization given by (cf. (II.5.20))

$$u_t = \Gamma((t, Z_t), u^*) = m(\exp(Fl_t^{J^2}(\log(u^*)), \exp(\sum_{j=1}^d e^{-\kappa_j \cdot} Z_t^j)), \tag{4.6}$$

where J^2 and Z_t^j, $j = 1, ..., d$ were given in (II.5.14) and (II.5.19) respectively. We look now to the more general case in which the diffusion vector fields are allowed to be of the form (4.5) and consider the following candidate distribution

$$D_{LA}(h) := \langle \Xi(h), \lambda_1, ..., \lambda_d \rangle, \quad h \in \mathcal{V},$$

where $\mathcal{V}$ is some open subset of $\mathcal{U}^l \cap D(A^\infty)$. We can conclude as in Lemma 3.4 that in this case we have

$$D_{LA}(h) = \langle \pi(h), \lambda_1, ..., \lambda_d \rangle, \quad h \in \mathcal{V}, \tag{4.7}$$

where

$$\pi(h) := Ah + \alpha(h), \quad \alpha(h) := \begin{pmatrix} -\frac{1}{2} \sum_{j=1}^d \phi_j^2(h^2) m(\lambda_j^1, \lambda_j^1) \\ \tilde{a}(h^1, h^2) \end{pmatrix}, \tag{4.8}$$

where the superscript in ϕ_j^2 indicates the exponent while in the remaining cases it indicates the coordinates.

Proposition 4.4. *Let* $\lambda_1^1, ..., \lambda_d^1$ *be linearly independent and* $p \in \{0, ..., d\}$ *be such that*

$$\begin{cases} m(\lambda_j^1, \lambda_j^1) \in \langle \lambda_1^1, ..., \lambda_d^1 \rangle, & \textit{for } j = 1, ..., p, \\ m(\lambda_j^1, \lambda_j^1) \notin \langle \lambda_1^1, ..., \lambda_d^1 \rangle, & \textit{for } j = p+1, ..., d, \end{cases} \tag{4.9}$$

where the cases $p = 0$ and $p = d$ are understood as either $j = 1, ..., p$ or $j = p+1, ..., d$ being the empty set. Let $M \in \mathbb{R}^{p\times d}$ be such that

$$m(\lambda_j^1, \lambda_j^1) = \sum_{l=1}^{d} M^{jl} \lambda_l^1, \quad \text{for } j = 1, ..., p. \tag{4.10}$$

Further let $\mathcal{V} =: \mathcal{V}^1 \times \mathcal{V}^2 \subset (U^l \times \mathbb{R}^m) \cap D(A^\infty)$ and $\zeta_{ij} : \mathcal{V}^2 \to \mathbb{R}$ for $(i,j) \in \{1, ..., d\}^2$ be the map given by

$$\zeta_{ij}(g) := -\phi_j(g) D\phi_j(g) \cdot \lambda_i^2. \tag{4.11}$$

Then the distribution D_{LA} given in (4.7) is involutive on $\mathcal{V}$ if and only if each of the following holds true:

1 *There is a matrix $\tilde{\zeta} \in \mathbb{R}^{d\times(d-p-1)}$ such that for $(i,j) \in \{1, ..., d\} \times \{p+1, ..., d\}$ $\zeta_{ij}(h^2) = \tilde{\zeta}_{ij}$ for all $h \in \mathcal{V}$.*

2 *There is a matrix $\tilde{B} \in \mathbb{R}^{d\times d}$ such that*

$$\frac{d}{dx}\lambda_i^1 = \sum_{j=1}^{d} \tilde{B}^{ij} \lambda_j^1 - \sum_{j=p+1}^{d} \tilde{\zeta}_{ij} m(\lambda_j^1, \lambda_j^1), \quad i = 1, ..., d. \tag{4.12}$$

3 *There is a map $\hat{a} : \mathcal{V}^2 \to \mathbb{R}^m$ such that the drift coefficient $\tilde{a}$ in (4.8) satisfies $\tilde{a}(h) = \hat{a}(h^2)$ for all $h \in \mathcal{V}$ and*

$$D\hat{a}(h^2) \cdot \lambda_i^2 = \sum_{j=1}^{d} \Big(\tilde{B}^{ij} + \sum_{l=1}^{p} \zeta_{il}(h^2) M^{lj} \Big) \lambda_j^2 \tag{4.13}$$

for all $i = 1, ..., d$ and $h \in \mathcal{V}$.

Proof. D_{LA} is involutive on $\mathcal{V}$ if and only if $[\pi, \lambda_i](h) \in \langle \lambda_1, ..., \lambda_d \rangle$ for all $i = 1, ..., d$ and $h \in \mathcal{V}$ (cf. Lemma 3.5). As we have (cf. (4.8))

$$[\pi, \lambda_i](h) \quad = \quad \begin{pmatrix} \frac{d}{dx}\lambda_i^1 \\ 0 \end{pmatrix} + \begin{pmatrix} \sum_{j=1}^{d} \zeta_{ij}(h^2) m(\lambda_j^1, \lambda_j^1) \\ D\tilde{a}(h)\lambda_i \end{pmatrix}, \tag{4.14}$$

where ζ_{ij} is given by (4.11), this is equivalent to the existence of a coordinate mapping $B : \mathcal{H} \to \mathbb{R}^{d\times d}$ such that

$$[\pi, \lambda_i](h) = \sum_{j=1}^{d} B^{ij}(h) \lambda_j \tag{4.15}$$

for all $i = 1, ..., d$ and $h \in \mathcal{V}$. As the first coordinate of (4.14) is constant in the first argument h^1 of $h = (h^1, h^2)$ the same must be true for the map B which in turn yields by looking at the second coordinate of (4.14) the same condition for $\tilde{a}$. Thus there is a mapping $\hat{a} : \mathcal{V}^2 \to \mathbb{R}^m$ such that necessarily $\hat{a}(h^2) := \tilde{a}(h^1, h^2)$ and

$$D\hat{a}(h^2) \cdot \lambda_i^2 = \sum_{j=1}^{d} B^{ij}(h^2) \lambda_j^2 \tag{4.16}$$

for all $h \in \mathcal{V}$. Using (4.10) we can write

$$\sum_{j=1}^{d} \zeta_{ij}(h^2) m(\lambda_j^1, \lambda_j^1) = \sum_{l=1}^{d} \Big(\sum_{j=1}^{p} \zeta_{ij}(h^2) M^{jl} \Big) \lambda_l^1 + \sum_{j=p+1}^{d} \zeta_{ij}(h^2) m(\lambda_j^1, \lambda_j^1).$$

Now using this representation in the first coordinate of (4.14) we see that a necessary condition for (4.15) to hold is that

$$\frac{d}{dx}\lambda_i^1 = \sum_{l=1}^{d}\Big(B^{il}(h^2) - \Big(\sum_{j=1}^{p}\zeta_{ij}(h^2)M^{jl}\Big)\Big)\lambda_l^1 - \sum_{j=p+1}^{d}\zeta_{ij}(h^2)m(\lambda_j^1,\lambda_j^1) \tag{4.17}$$

for all $i = 1, ..., d$ and $h \in \mathcal{V}$ and thus again by (4.9) it follows that necessarily $\Big(B^{il}(h^2) - \Big(\sum_{j=1}^{p}\zeta_{ij}(h^2)M^{jl}\Big)\Big)$ for $(i,l) \in \{1,...,d\}^2$ and $\zeta_{ij}(h^2)$ for $(i,j) \in \{1,...,d\}\times\{p+1,...,d\}$ is constant in h^2. Thus letting

$$\tilde{B}^{il} := \Big(B^{il}(h^2) - \Big(\sum_{j=1}^{p}\zeta_{ij}(h^2)M^{jl}\Big)\Big), \quad \text{for } (i,l) \in \{1,...,d\}^2 \tag{4.18}$$

and $\tilde{\zeta}_{ij} := \zeta_{ij}(h^2)$ for $(i,j) \in \{1,...,d\}\times\{p+1,...,d\}$ shows necessity of the claims **1** and **2** and by plugging this into (4.17) gives also necessity of claim **3**. On the other hand, applying conditions **1-3** on (4.14) yields (4.15) and thus sufficiency holds as well. □

Remark 4.5. An example where we will have the situation (4.9) for $p \in \{1, ..., d-1\}$ is for $\lambda_1 = e^{a_1}, \lambda_2 = e^{a_2}, ...$ for distinct $a_1, a_2, ...$ such that at least one pair a_i, a_j exists such that $a_i = 2a_j$.

Remark 4.6. Condition **1** in the last proposition states that $\phi_{p+1}, ..., \phi_d$ given in (4.5) necessarily satisfy the ODEs

$$-\phi_j D\phi_j \cdot \lambda_i^2 = \tilde{\zeta}_{ij}, \quad \text{for all } (i,j) \in \{1,...,d\}\times\{p+1,...,d\}. \tag{4.19}$$

Solutions $\phi_j : \mathcal{V}^2 \to \mathbb{R}$ for suitable $\mathcal{V}^2 \subset \mathbb{R}^m$ can be constructed as follows. For each $j \in \{p+1, ..., d\}$ let $\gamma_j \in \mathbb{R}$ and $l_j \in L(\mathbb{R}^m, \mathbb{R})$ with $l_j(\mathcal{V}^2) \subset [-\gamma_j, \infty)$ such that for $i = 1, ..., d$, l_j satisfies

$$-\frac{1}{2}l_j(\lambda_i^2) = \tilde{\zeta}_{ij}. \tag{4.20}$$

Then it is easy to see that

$$\phi_j : \mathcal{V}^2 \to \mathbb{R}, \quad \phi_j(y) := \sqrt{\gamma_j + l_j(y)} \tag{4.21}$$

satisfies (4.19). It is also evident that any solution can be represented as in (4.21). In particular, if $p = 0$ then it follows from **3** that $\hat{a}$ is a linear map on $\mathcal{V}^2$. Regarding condition $-\frac{1}{2}l_j(\lambda_i^2) = \tilde{\zeta}_{ij}$ recall from Remark 4.3 that we have not assumed that $\lambda_1^2, ..., \lambda_d^2$ are linearly independent neither have we made any statement whether $m \geq d$ or $m < d$.

Lemma 4.7. *Let $h^* \in \mathcal{V}$ and set*

$$\Gamma(z, h^*) := Fl_{z_0}^{\pi}(h^*) + \sum_{i=1}^{d} z_i\lambda_i. \tag{4.22}$$

Then under the conditions of Proposition 4.4 and recalling Remark 4.6, the following representation holds true

$$\begin{aligned}
&\alpha(\Gamma(z,h^*)) - \alpha(Fl_{z_0}^{\pi}(h^*)) \\
= \; & -\frac{1}{2}\sum_{l=1}^{d}\sum_{j=1}^{p}\Big(\phi_j^2(\Gamma^2(z,h^*) - \phi_j^2(Fl_{z_0}^{\pi,2}(h^*))\Big)M^{jl}\lambda_l \\
= \; & \begin{pmatrix} \sum_{j=p+1}^{d}\sum_{i=1}^{d} z_i\tilde{\zeta}_{ij}m(\lambda_j^1,\lambda_j^1) \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \sum_{i=1}^{d}\sum_{j=1}^{d} z_i\tilde{B}^{ij}\lambda_j^2 \end{pmatrix}.
\end{aligned}$$

Proof. From (4.8) we get

$$\begin{aligned} & \alpha(\Gamma(z,h^*)) - \alpha(Fl^{\pi}_{z_0}(h^*)) \\ = & \begin{pmatrix} -\frac{1}{2}\sum_{j=1}^d \Big(\phi_j^2(\Gamma^2(z,h^*)) - \phi_j^2(Fl^{\pi,2}_{z_0}(h^*))\Big) m(\lambda_j^1,\lambda_j^1) \\ \hat{a}(\Gamma^2(z,h^*)) - \hat{a}(Fl^{\pi,2}_{z_0}(h^*)) \end{pmatrix}. \end{aligned}$$

The first coordinate equals, using (4.10) and Remark 4.6

$$-\frac{1}{2}\sum_{l=1}^d\sum_{j=1}^p \Big(\phi_j^2(\Gamma^2(z,h^*)) - \phi_j^2(Fl^{\pi,2}_{z_0}(h^*))\Big) M^{jl}\lambda_l^1 + \sum_{j=p+1}^d\sum_{i=1}^d z_i\tilde{\zeta}_{ij} m(\lambda_j^1,\lambda_j^1).$$

For the second coordinate, letting $\xi(t) := Fl^{\pi,2}_{z_0} + t\sum_{i=1}^d z_i\lambda_i^2$ and using **2** in Proposition 4.4 we get

$$\begin{aligned} & \hat{a}(\Gamma^2(z,h^*)) - \hat{a}(Fl^{\pi,2}_{z_0}(h^*)) = \int_0^1 \frac{d}{ds}\hat{a}(\xi(s))\,ds = \int_0^1 D\hat{a}(\xi(s))\cdot(\sum_{i=1}^d z_i\lambda_i^2)\,ds \\ = & \sum_{i=1}^d z_i \int_0^1 \sum_{j=1}^d \Big(\tilde{B}^{ij} + \sum_{l=1}^p \zeta_{il}(\xi(s))M^{lj}\Big)\lambda_j^2\,ds \\ = & \sum_{i=1}^d\sum_{j=1}^d z_i\tilde{B}^{ij}\lambda_j^2 + \sum_{j=1}^d\sum_{l=1}^p M^{lj}\lambda_j^2 \int_0^1 \sum_{i=1}^d z_i\zeta_{il}(\xi(s))\,ds. \end{aligned}$$

Now for the integral using (4.11) we get form integration by parts

$$\begin{aligned} & \int_0^1 \sum_{i=1}^d z_i\zeta_{il}(\xi(s))\,ds = -\int_0^1 \sum_{i=1}^d z_i\phi_l(\xi(s))D\phi_l(\xi(s))\cdot\lambda_i^2\,ds \\ = & -\int_0^1 \phi_l(\xi(s))\frac{d}{ds}\phi_l(\xi(s))\,ds = -\frac{1}{2}(\phi_l^2(\Gamma^2(z,h^*)) - \phi_l^2(Fl^{\pi,2}_{z_0}(h^*))) \end{aligned}$$

which gives the claim. □

Remark 4.8. Condition **2** in the last proposition states that $\lambda^{1p} := (\lambda_1^1,...,\lambda_p^1)$ and $\lambda^{1(p+1)} := (\lambda_{p+1}^1,...,\lambda_d^1)$, both understood as column vectors, satisfy the following linear and Riccati ODE respectively

$$\begin{aligned} \frac{d}{dx}\lambda^{1p} &= B_{1p}\lambda^{1p} + C_p(\lambda^{1(p+1)}) \\ \frac{d}{dx}\lambda^{1(p+1)} &= B_{1(p+1)}\lambda^{1(p+1)} + B_{2(p+1)}m(\lambda^{1(p+1)},\lambda^{1(p+1)}) + C_{p+1}(\lambda^{1p}), \end{aligned}$$

where

$$\begin{aligned} & m(\lambda^{1(p+1)},\lambda^{1(p+1)}) := (m(\lambda_{p+1}^1,\lambda_{p+1}^1),...,m(\lambda_d^1,\lambda_d^1)) \\ & B_{1p} := (\tilde{B}^{ij})_{i,j=1,\dots,p} \\ & C_p := (\sum_{j=p+1}^d \tilde{B}^{ij}\lambda_j^1 - \sum_{j=p+1}^d \tilde{\zeta}_{ij}m(\lambda_j^1,\lambda_j^1))_{i=1,\dots,p} \\ & B_{1(p+1)} := (\tilde{B}^{ij})_{i,j=p+1,\dots,d} \\ & B_{2(p+1)} := (-\tilde{\zeta}_{ij})_{i,j=p+1,\dots,d} \\ & C_{p+1} := (\sum_{j=1}^p \tilde{B}^{ij}\lambda_j^1)_{i=p+1,\dots,d}. \end{aligned}$$

In this case λ^{1p} can be represented by using variation of constants as

$$\lambda^{1p} = e^{B_{1p}\cdot}\lambda(0) + e^{B_{1p}\cdot}\int_0^{\cdot} C_p e^{-B_{1p}\eta}\, d\eta.$$

In particular, we see from **2** that $\lambda_1^1, ..., \lambda_d^1$ are quasi-exponentials (cf. [13, Corollary 5.1]) if $\tilde{\zeta}_{ij} = 0$ for $(i,j) \in \{1,...,d\} \times \{p+1,...,d\}$. In this case we are in the setting of Remark 4.6 and $\phi_1, ..., \phi_d$ must satisfy (4.19) which hold by choosing them as in (4.21).

Theorem 4.9. *If the dimension of D_{LA} is constant on $\mathcal{V}$ and the conditions of Proposition 4.4 are satisfied, then the system* (4.3) *has the representation*

$$\begin{cases} dJ_t = (\frac{d}{dx}J_t - \frac{1}{2}\sum_{j=1}^d \phi_j^2(Y_t) m(\lambda_j^1, \lambda_j^1)\, dt + \sum_{j=1}^d \phi_j(Y_t)\lambda_j^1\, d\beta_t^j \\ dY_t = \hat{a}(Y_t)\, dt + \sum_{j=1}^d \phi_j(Y_t)\lambda_j^2\, d\beta_t^j, \\ (J_0, \tilde{Y}_0) \in \mathcal{U}^l := U^l \times \mathbb{R}^m, \end{cases} \tag{4.23}$$

and around any $h^* \in \mathcal{V}$ admits a generic finite dimensional realization given by

$$\Gamma(t, Z_t, h^*) := Fl_t^{\pi}(h^*) + \sum_{l=1}^d Z_t^l \lambda_l, \tag{4.24}$$

where Z is the $\mathbb{R}^d$-valued time-inhomogeneous diffusion process given as the solution of

$$dZ_t^l = b_l(t, Z_t)\, dt + \sum_{i=1}^d \rho_i(t, Z_t)\, d\beta_t^i, \quad Z_0^l = 0, \quad l = 1, ..., d, \tag{4.25}$$

with

$$b_l(t, Z_t) = \sum_{i=1}^d Z_t^i \tilde{B}^{il} - \frac{1}{2}\sum_{j=1}^p \Big(\phi_j^2(\Gamma^2(t, Z_t, h^*) - \phi_j^2(Fl_t^{\pi,2}(h^*))\Big) M^{jl},\ l = 1, ..., d \tag{4.26}$$

with $\tilde{B}$ and M given in **2** *and* (4.10) *of Proposition 4.4 and*

$$\rho_l(t, Z_t) = \phi_l(\Gamma^2(t, Z_t, h^*)), \quad l = 1, ..., d.$$

Furthermore, recalling the notation $\mathcal{V} = \mathcal{V}^1 \times \mathcal{V}^2 \subset (U^l \times \mathbb{R}^m) \cap D(A^\infty)$ (cf. Proposition 4.4), in this case also

$$\begin{cases} du_t = \frac{d}{dx}u_t\, dt + \sum_{j=1}^d \phi_j(Y_t) m(\lambda_j^1, u_t)\, d\beta_t^j \\ dY_t = \hat{a}(Y_t)\, dt + \sum_{j=1}^d \phi^j(Y_t)\lambda_j^2\, d\beta_t^j, \\ (u_0, Y_0) \in \mathcal{U} := U \times \mathbb{R}^m, \end{cases} \tag{4.27}$$

admits a generic finite dimensional realization around any $h^* = (u^*, Y^*) \in \exp(\mathcal{V}^1) \times \mathcal{V}^2$ and is given by

$$\tilde{\Gamma}(t, Z_t, (u^*, Y^*)) \quad := \quad \begin{pmatrix} \exp(\Gamma^1(t, Z_t, (\log(u^*), Y^*))) \\ \Gamma^2(t, Z_t, (\log(u^*), Y^*)) \end{pmatrix} \tag{4.28}$$

Proof. We first show (4.24) and proceed as in the proof of Theorem 3.9. Using (3.23) and (3.24) with π given in (4.8) we will determine the coordinate process Z. For condition (3.23), notice that

$$\pi(\Gamma(z,h^*)) = \pi(Fl^\pi_{z_0}(h^*)) + \sum_{i=1}^d z_i A\lambda_i + \alpha(\Gamma(z,h^*)) - \alpha(Fl^\pi_{z_0}(h^*))$$

and accordingly from **2** in Proposition 4.4 and Lemma 4.7 we get

$$\begin{aligned} &\pi(\Gamma(z,h^*)) = \pi(Fl^\pi_{z_0}(h^*)) \\ +\quad & \sum_{l=1}^d \Big[\sum_{i=1}^d z_i \tilde{B}^{il} - \frac{1}{2}\sum_{j=1}^p \Big(\phi_j^2(\Gamma^2(z,h^*)) - \phi_j^2(Fl^{\pi;2}_{z_0}(h^*))\Big) M^{jl}\Big]\lambda_l. \end{aligned} \tag{4.29}$$

Thus it follows from (3.23) that

$$b_l(z) := \sum_{i=1}^d z_i \tilde{B}^{il} - \frac{1}{2}\sum_{j=1}^p \Big(\phi_j^2(\Gamma^2(z,h^*) - \phi_j^2(Fl^{\pi;2}_{z_0}(h^*))\Big) M^{jl},\ l = 1,...,d.$$

Similarly, using (3.24) it follows that

$$\rho_l(z) := \phi_l(\Gamma^2(z,h^*)), \quad l = 1,...,d.$$

Then by defining the coordinate process as in (4.25) and using (4.29) we see that

$$\begin{aligned} & Fl^\pi_t(h^*) + \sum_{l=1}^d Z^l_t \lambda_l \\ =\quad & h^* + \int_0^t \pi(\Gamma(s,Z_s,h^*))\,ds + \sum_{l=1}^d \int_0^t \phi_l(\Gamma^2(s,Z_s,h^*))\lambda_l\,d\beta^l_s \end{aligned}$$

and thus the first claim. The second claim and the Representation (4.28) follow from Corollary 4.1. □

In particular we see that, if $\lambda^1_1, ..., \lambda^1_d$ are quasi-exponentials it follows from Remark 4.8 that necessarily $\phi_1, ..., \phi_d$ are of the form (4.21) such that $\lambda^2_1, ..., \lambda^2_d$ lie in the kernel of the corresponding linear maps $l_1, ..., l_d$. In this case the drift coefficient (4.26) in (4.25) reduces to

$$b_l(z_0, z) = \sum_{i=1}^d \tilde{B}^{il} - \frac{1}{2}\sum_{j=1}^p \gamma_j M^{jl},\ l = 1,...,d$$

as for $j = 1,...,p$ it holds true that

$$\phi_j^2(\Gamma^2(z_0,z,h^*) - \phi_j^2(Fl^{\pi;2}_{z_0}(h^*)) = \gamma_j + l_j(\sum_{i=1}^d z_i\lambda^2_i) = \gamma_j.$$

By comparing this result to [12, Proposition 5.2] we see that the condition that $\lambda_1, ..., \lambda_d$ are quasi-exponentials is (basically) sufficient for the existence of a generic finite dimensional realization but not necessary. We saw also that in the considered case where

u and Y are driven by the same Brownian motions we necessarily arrived at a situation where Y is the solution of an autonomous SDE and hence the ad hoc chosen restriction in [12] of considering only autonomous stochastic parameter processes Y appears in this setting to be justified. This was not the case in the previously investigated models that lead to generic affine realizations given in (3.9). In the next subsection we will see that this restriction is not necessary as well in the case where both processes are driven by independent Brownian motions (cf. *orthogonal noise models* in [12, Section 4]).

4.1 Independent parameter process

We repeat the analysis of the previous subsection right before Lemma 4.2 by looking at a version of the reduced model (4.3) where the Assumptions **(A1)-(A3)** hold true but the parameter process Y is driven by independent Brownian motions, i.e. we look at

$$\begin{cases} dJ_t = (\frac{d}{dx}J_t - \frac{1}{2}\sum_{j=1}^d m(\sigma_j(\tilde{Y}_t), \sigma_j(\tilde{Y}_t))\,dt + \sum_{j=1}^d \sigma_j(\tilde{Y}_t)\,d\beta_t^j \\ d\tilde{Y}_t = \tilde{a}(l(J_t), \tilde{Y}_t)\,dt + \sum_{j=d+1}^{d+m} \tilde{c}^j(l(J_t), \tilde{Y}_t)\,d\beta_t^j, \\ (J_0, \tilde{Y}_0) \in \mathcal{U}^l := U^l \times \mathbb{R}^m, \end{cases} \tag{4.30}$$

where again $\tilde{a}(l(J), Y) := a(l(\exp(J)), Y)$, $\tilde{c}^j(l(J), Y) := c^j(l(\exp(J)), Y)$ for $j = 1, ..., d$ and $U^l := \log(U)$, which is well-defined as $U \subset H^+$. Then it follows as in Lemma 4.2 from Theorem 3.3 that necessarily there are smooth vector fields $\lambda_1, ..., \lambda_d$ such that

$$\begin{pmatrix} \sigma_j(Y) \\ \tilde{c}^j(l(J), Y) \end{pmatrix} \in \langle \lambda_1(Y), ..., \lambda_d(Y), \lambda_{d+1}(Y), ..., \lambda_{d+m}(Y) \rangle,$$

where now $\sigma_j(Y)$ and $\tilde{c}^j(l(J), Y)$ can have different coordinate mappings and hence $\tilde{c}$ must not be constant in the first argument. As in Section 3.1 we assume that

$$\langle \lambda_1(Y), ..., \lambda_d(Y), \lambda_{d+1}(Y), ..., \lambda_{d+m}(Y) \rangle = \left\langle \begin{pmatrix} \lambda_1^1 \\ 0 \end{pmatrix}, ..., \begin{pmatrix} \lambda_d^1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ e_1 \end{pmatrix}, ..., \begin{pmatrix} 0 \\ e_m \end{pmatrix} \right\rangle$$

and that **A4** holds, that is, there are scalar fields on $\mathbb{R}^m$ denoted by $\phi_1, ..., \phi_d$ and on $U^l \times \mathbb{R}^m$ denoted by $\phi_{d+1}, ..., \phi_{d+1+m}$ such that

$$\sigma_j(Y) = \phi_j(Y)\lambda_j^1, \quad j = 1, ..., d, \tag{4.31}$$

$$\tilde{c}^j(l(J), Y) = \tilde{\phi}_j(J, Y)e_{j-d}, \quad j = d+1, ..., d+m. \tag{4.32}$$

Letting $\mathcal{V}$ be an open subset of $\mathcal{U}^l \cap D(A^\infty)$ we consider the following candidate distribution (cf. (4.7))

$$D_{LA}(h) = \left\langle \pi(h), \begin{pmatrix} \lambda_1^1 \\ 0 \end{pmatrix}, ..., \begin{pmatrix} \lambda_d^1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ e_1 \end{pmatrix}, ..., \begin{pmatrix} 0 \\ e_m \end{pmatrix} \right\rangle, \quad h \in \mathcal{V}, \tag{4.33}$$

where

$$\pi(h) := Ah + \alpha(h), \quad \alpha(h) := \begin{pmatrix} -\frac{1}{2}\sum_{j=1}^d \phi_j^2(h^2)m(\lambda_j^1, \lambda_j^1) \\ \tilde{a}(l(h^1), h^2) \end{pmatrix}. \tag{4.34}$$

We arrive at the following version of Proposition 4.4. Notice that in the current setting $\lambda_1^1, ..., \lambda_d^1$ are necessarily linearly independent.

Proposition 4.10. *Under the conditions of Propositions 4.4 with ζ_{ij} in* (4.11) *replaced by*

$$\zeta_{ij}(g) := -\phi_j(g) D\phi_j(g) \cdot e_i, \quad (i,j) \in \{1,...,m\} \times \{1,...,d\} \tag{4.35}$$

the distribution D_{LA} on $\mathcal{V}$ given in (4.33) *is involutive, if and only if, the following two conditions hold true*

1 *There is a matrix B in $\mathbb{R}^{d\times d}$ such that*

$$\frac{d}{dx}\lambda_i^1 = \sum_{j=1}^{d} B^{ij}\lambda_j^1, \quad i = 1,...,d, \tag{4.36}$$

2 *$\phi_{p+1},...,\phi_d$ in* (4.31) *are constant maps.*

Proof. We proceed as in the proof of Proposition 4.4. By noticing that for all $h \in \mathcal{V}$ $[\pi, \lambda_i](h) \in \langle \lambda_1, ..., \lambda_{m+d}\rangle$ if and only if $([\pi, \lambda_i](h))^1 \in \langle \lambda_1^1, ..., \lambda_d^1\rangle$ it suffices to look at the first coordinate. We have

$$([\pi, \lambda_i](h))^1 = \begin{cases} \frac{d}{dx}\lambda_i^1, & \text{for } i = 1,...,d, \\ \sum_{j=1}^{d} \zeta_{ij}(h^2) m(\lambda_j^1, \lambda_j^1), & \text{for } i = d+1,...,d+m, \end{cases} \tag{4.37}$$

which immediately gives the first claim. For the second, by using (4.10) we can again write

$$\sum_{j=1}^{d} \zeta_{ij}(h^2) m(\lambda_j^1, \lambda_j^1) = \sum_{l=1}^{d} \Big(\sum_{j=1}^{p} \zeta_{ij}(h^2) M^{jl}\Big)\lambda_l^1 + \sum_{j=p+1}^{d} \zeta_{ij}(h^2) m(\lambda_j^1, \lambda_j^1)$$

and conclude from (4.37) and (4.9) that $([\pi, \lambda_i](h))^1 \in \langle \lambda_1^1, ..., \lambda_d^1\rangle$ holds for $i = d+1,...,d+m$ if and only if $\zeta_{ij} = 0$ for $(i,j) \in \{1,...,m\} \times \{p+1,...,d\}$ from which via the condition (4.20) (with $\zeta_{ij} = 0$) in Remark 4.6 the second claim follows. □

Interestingly we see that if the distribution D_{LA} is involutive on $\mathcal{V}$ then the second condition of the previous Proposition states that the (logarithm of the) forward variance process J in (4.30) can have a stochastic volatility component only if the corresponding direction λ_i^1 is such that $m(\lambda_i^1, \lambda_i^1)$ lies in the span of $\lambda_1^1, ..., \lambda_d^1$. This result is stronger as in the previous section as here $e_1, ..., e_m$ spans $\mathbb{R}^m$ as opposed to $\lambda_1^2, ..., \lambda_d^2$ (cf. Remark 4.19). On the other hand, we see that in the present case, there are no restrictions posed on the stochastic volatility process Y. Thus we can state the following version of Theorem 4.9 for the system (4.30).

Corollary 4.11. *If the dimension of D_{LA} is constant on $\mathcal{V}$ and the conditions of Proposition 4.10 are satisfied, then the system* (4.30) *has the representation*

$$\begin{cases} dJ_t = (\frac{d}{dx}J_t - \frac{1}{2}\sum_{j=1}^{d} \phi_j^2(\tilde{Y}_t) m(\lambda_j^1, \lambda_j^1)\, dt + \sum_{j=1}^{d} \phi_j(\tilde{Y}_t)\lambda_j^1\, d\beta_t^j \\ d\tilde{Y}_t = \tilde{a}(l(J_t), \tilde{Y}_t)\, dt + \sum_{j=d+1}^{d+m} \tilde{\phi}_j(J_t, \tilde{Y}_t)\lambda_j^2\, d\beta_t^j, \\ (J_0, \tilde{Y}_0) \in \mathcal{U}^l := U^l \times \mathbb{R}^m, \end{cases} \tag{4.38}$$

and around any $h^* \in \mathcal{V}$ admits a generic finite dimensional realization given by

$$\Gamma(t, Z_t, h^*) := Fl_t^{\pi}(h^*) + \sum_{l=1}^{d+m} Z_t^l \lambda_l,$$

where Z is the $\mathbb{R}^d$-valued time-inhomogeneous diffusion process given as the solution of

$$dZ_t^l = b_l(t, Z_t)\,dt + \sum_{i=1}^{d+m} \rho_i(t, Z_t)\,d\beta_t^i, \quad Z_0^l = 0, \quad l = 1, ..., d, \tag{4.39}$$

with

$$b_l(t, Z_t) := \sum_{i=1}^{d} Z_t^i B^{il} - \frac{1}{2}\sum_{j=1}^{p} \Big(\phi_j^2(\Gamma^2(t, Z_t, h^*)) - \phi_j^2(Fl_t^{\pi,2}(h^*))\Big)M^{jl},$$

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

$$b_l(t, Z_t) := \tilde{a}^l(l(\Gamma^1(t, Z_t, h^*)), \Gamma^2(t, Z_t, h^*)) - \tilde{a}^l(l(Fl_t^{\pi,1}(h^*)), Fl_t^{\pi,2}(h^*)),$$

*for $l = d+1, ..., d+m$, with $\tilde{a}^l$ denoting the coordinate of $\tilde{a}$ with respect e_l and B and M were given in **1** and* (4.10) *of Proposition 4.10 and 4.4 respectively, and*

$$\rho_l(t, Z_t) := \begin{cases} \phi_l(\Gamma^2(t, Z_t, h^*)), & l = 1, ..., d, \\ \phi_l(\Gamma(t, Z_t, h^*)), & l = d+1, ..., d+m. \end{cases} \tag{4.40}$$

Furthermore, recalling the notation $\mathcal{V} = \mathcal{V}^1 \times \mathcal{V}^2 \subset (U^l \times \mathbb{R}^m) \cap D(A^\infty)$ (cf. Proposition 4.4), in this case also

$$\begin{cases} du_t = \frac{d}{dx}u_t\,dt + \sum_{j=1}^{d} \phi_j(Y_t)m(\lambda_j^1, u_t)\,d\beta_t^j \\ dY_t = a(l(u_t), Y_t)\,dt + \sum_{j=d+1}^{d+m} \phi_j(u_t, Y_t)\lambda_j^2\,d\beta_t^j, \\ (u_0, Y_0) \in \mathcal{U} := U \times \mathbb{R}^m, \end{cases}$$

admits a generic finite dimensional realization around any $h^* = (u^*, Y^*) \in \exp(\mathcal{V}^1) \times \mathcal{V}^2$ and is given by

$$\tilde{\Gamma}(t, Z_t, (u^*, Y^*)) \quad := \quad \begin{pmatrix} \exp(\Gamma^1(t, Z_t, (\log(u^*), Y^*))) \\ \Gamma^2(t, Z_t, (\log(u^*), Y^*)) \end{pmatrix}.$$

Proof. We omit the proof as it is essentially the same as in Theorem 4.9 and only note that the representation corresponding to (4.29) is now given by

$$\begin{aligned} &\pi(\Gamma(z, h^*)) = \pi(Fl_{z_0}^{\pi}(h^*)) \\ &+ \sum_{l=1}^{d}\Big[\sum_{i=1}^{d} z_i B^{il} - \frac{1}{2}\sum_{j=1}^{p}\Big(\phi_j^2(\Gamma^2(z, h^*)) - \phi_j^2(Fl_{z_0}^{\pi,2}(h^*))\Big)M^{jl}\Big]\begin{pmatrix}\lambda_l^1 \\ 0\end{pmatrix} \\ &+ \sum_{l=d+1}^{d+m}\Big(\tilde{a}^l(l(\Gamma^1(z, h^*)), \Gamma^2(z, h^*)) - \tilde{a}^l(l(Fl_{z_0}^{\pi,1}(h^*)), Fl_{z_0}^{\pi,2}(h^*))\Big)\begin{pmatrix}0 \\ e_{l-d}\end{pmatrix}. \end{aligned}$$

□

We see that as in the case of an independent parameter process for generic affine realizations discussed in Section 3.1 there are no additional restrictions on the parameter process, which brings us to a more general setup when compared to the situation investigated [12], where only autonomous parameter processes were considered. On the other hand, as opposed to the case where the forward variance and the parameter process are driven by the same Brownian motions, we see that the condition that $\lambda_1^1, ..., \lambda_d^1$ are quasi-exponentials are not only sufficient by also necessary due to reasons that were already discussed above after the proof of the Proposition 4.10.

Example 4.12 (Constant Diffusion). We revisit the Example II.4.5 from Musiela's perspective, by looking at the model (4.1) with $m = 0$ and $d = 1$ given by

$$\begin{cases} du_t = \frac{d}{dx} u_t \, dt + m(\sigma, u_t) \, d\beta_t \\ u_0 \in U. \end{cases} \tag{4.41}$$

We know from Corollary 4.11 that (essentially) this model admits a generic finite dimensional realization if and only if σ is a quasi exponential. In this case we can represent the solution by

$$u_t = \exp(Fl_t^{\pi}(\log(u_0)) + Z_t\sigma) = m(\exp(Fl_t^{\pi}(\log(u_0))), \exp(Z_t\sigma)), \tag{4.42}$$

where $\pi = \frac{d}{dx} - \frac{1}{2} m(\sigma, \sigma)$ and by letting $\frac{d}{dx}\sigma = b\sigma$, Z is given as the solution of

$$dZ_t = bZ_t \, dt + d\beta_t. \tag{4.43}$$

We assume now that σ is a constant vector given by $\sigma := c1$ where c is a real number and 1 is the constant vector that satisfies $m(1, u) = u$ for all $u \in H$. Then necessarily $b = 0$ and

$$Fl_t^{\pi}(\log(u_0)) = \log(S_t u_0) - \frac{1}{2} \int_0^t S_{t-s} c^2 m(1, 1) \, ds = \log(S_t u_0) - \frac{1}{2} c^2 t 1 \tag{4.44}$$

and accordingly

$$u_t = S_t u_0 \exp((-\frac{1}{2} c^2 t + c\beta_t) 1). \tag{4.45}$$

Thus for any initial curve u_0 of the form $ae^{b\cdot}$ it follows that the short-variance $\xi_t := u_t(0)$ follows a geometric Brownian motion given by

$$\xi_t = a + \int_0^t b\xi_s \, ds + \int_0^t c\xi_s \, d\beta_s \tag{4.46}$$

for which we know that the forward variance can be represented as $u_t = e^{b\cdot}\xi_t$ which (again) gives an example of an exponentially affine realization that has also an affine realization.

Example 4.13 (Bergomi Model). From Proposition 4.10 we see that the basic Bergomi model considered in Example II.5.16 corresponds to the case where $m = 0$ and B being a diagonal matrix.

5 Mixed Models

So far we saw that forward variance models that have diffusion coefficients that are given by the constant direction type $\phi(u)\lambda$ or of multiplication type $m(\lambda, u)$ can lead to generic finite dimensional realizations. In this section we will briefly look at mixtures of these types, namely additive and (in some sense) multiplicative mixtures. In both cases we restrict our self to autonomous equations, that is, without stochastic volatility.

5.1 Additive Mixtures

We look here at forward variance processes that are given as solutions of

$$\begin{cases} du_t = \frac{d}{dx} u_t \, dt + \phi(u_t)\lambda_1 \, d\beta_t^1 + m(\lambda_2, u_t) \, d\beta_t^2 \\ u_0 \in U \subset D((d/dx)^\infty). \end{cases} \tag{5.1}$$

The candidate distribution is given by $D(h) = \langle \Xi(h), \lambda_1, m(\lambda_2, h) \rangle$, for $h \in U$, where

$$\Xi(h) := \frac{d}{dx} h - \frac{1}{2} D\phi(h)\lambda_1 \cdot (\phi(h)\lambda_1) - \frac{1}{2} m(m(\lambda_2, \lambda_2), h).$$

As the second term lies in $\langle \lambda_1 \rangle$ it is enough to replace Ξ with π where

$$\pi(h) := \frac{d}{dx} h - \frac{1}{2} m(m(\lambda_2, \lambda_2), h).$$

Thus we are looking at

$$D(h) = \langle \pi(h), \lambda_1, m(\lambda_2, h) \rangle, \quad h \in U. \tag{5.2}$$

The following proposition gives equivalent conditions for the distribution D to be involutive which correspond to necessary conditions for a minimal generic finite dimensional distribution, see Remark 2.3. Recall that we denote by 1 the element in H such that $m(1, h) = h$ for all $h \in H$.

Proposition 5.1. *The distribution D is involutive on U if and only if there are real numbers a, b, c such that $\lambda_2 = c1$ and $\lambda_1 = ae^{b\cdot}$.*

Proof. The distribution is involutive if and only if the following three conditions are satisfied for all $h \in U$

$$[m(\lambda_2, \cdot), \lambda_1](h) \in D(h), \quad [\pi, \lambda_1](h) \in D(h), \quad [\pi, m(\lambda_2, \cdot)](h) \in D(h). \tag{5.3}$$

Assume that λ_2 is not constant. We show first that in this case λ_1 can not be constant neither. The first condition in (5.3) is equivalent to $m(\lambda_2, \lambda_1) = B^1(h)\lambda_1 + m(\lambda_2, B^2(h)h)$ for all $h \in U$, where B^1 and B^2 are the coordinate maps. Now if λ_1 is constant, say $z1$ for some real number z, then this amounts to $z\lambda_2 = B^1(h)z1 + B^2(h)m(\lambda_2, h)$ for all $h \in U$. As there is an open subset $\tilde{U} \subset U$ on which 1 and $m(\lambda_2, h)$ are linearly independent for all $h \in \tilde{U}$ it follows that necessarily $B^1(h)$ and $m(\lambda_2, B^2(h)h)$ are constant in h. Accordingly it must hold that $z\lambda_2 = B^1 z1 + B^2(h)m(\lambda_2, h)$ for all $h \in \tilde{U}$ which is impossible and hence λ_1 can not be constant in this case. But then we can find again an open subset $\tilde{U} \subset U$ on which λ_1 and $m(\lambda_2, h)$ are linearly independent for all $h \in \tilde{U}$ and thus it must hold both for all $h \in \tilde{U}$, $m(\lambda_2, \lambda_1) = B^1\lambda_1 + m(\lambda_2, B^2(h)h)$ and $m(\lambda_2, B^2(h)h)$ being constant in h which again is impossible and hence we conclude by contradiction the claim regarding λ_2. For the claim on λ_1 we use the second condition in (5.3), which is (recall that we just showed $\lambda_2 = c1$) equivalent to $\frac{d}{dx}\lambda_1 - \frac{1}{2}c^2\lambda_1 = \langle \lambda_1, ch \rangle$ which gives the claim on λ_1 and hence necessity. Showing sufficiency is straight forward. □

We will now show that this result regarding the condition on λ_2 corresponded already to the most general situation. Indeed, by considering the Lie-Algebra D_{LA} on U that is generated by the vector fields π, λ_1 and $m(\lambda_2, h)$ (i.e. this is the distribution that is spanned by this vector fields and all multiple Lie brackets). Now we look for more general conditions such that (5.1) admits a generic finite dimensional realization. It follows from Theorem 2.2 that a necessary condition is that the dimension of the Lie Algebra D_{LA} is finite (i.e. the dimension of the distribution $D_{LA}(h)$ is bounded in $h \in U$).

Proposition 5.2. *If the dimension of D_{LA} is finite then necessarily λ_2 is a quasi-exponential, that is,*

$$dim\Big\langle\Big\{\frac{d^n}{dx^n}\lambda_2\,\Big|\,n\in\mathbb{N}\Big\}\Big\rangle<\infty.$$

Proof. By defining recursively $[\pi, m(\lambda_2,\cdot)]^n := [\pi,[\pi, m(\lambda_2,\cdot)]^{n-1}]$ for $n \geq 2$ with initial condition $[\pi, m(\lambda_2,\cdot)]^1 := [\pi, m(\lambda_2,\cdot)]$ it follows by construction that $[\pi, m(\lambda_2,\cdot)]^n(h) \in D_{LA}(h)$ for all $n \geq 1$ and $h \in U$. Now by a straight forward calculation it follows that

$$[\pi, m(\lambda_2,\cdot)]^n(h) = m(\frac{d^n}{dx^n}\lambda_2, h), \quad n \geq 1, \quad h \in U.$$

Now if the pointwise dimension of D_{LA} is bounded in U by $m \geq 1$ then necessarily (notice that the case $n = 0$ is included as $m(\lambda_2, h) \in D_{LA}(h)$)

$$\dim\langle\{m(\frac{d^n}{dx^n}\lambda_2, h)\,|\,n \geq 0\}\rangle \leq m, \qquad \text{for all } h \in U.$$

In this case we can find a natural number $k \geq 1$ (cf. [71, Lemma 3.6]) such that

$$m(\frac{d^k}{dx^k}\lambda_2, h) = \sum_{i=0}^{k-1} a_i(h) m(\frac{d^i}{dx^i}\lambda_2, h),$$

where $a_0, ..., a_{k-1}$ denote the coordinate functions. As the left hand-side is linear in h it follows that the coordinate functions are necessarily constants and accordingly, again by linearity, we have that necessarily

$$m(\frac{d^k}{dx^k}\lambda_2 - \sum_{i=0}^{k-1} a_i \frac{d^i}{dx^i}\lambda_2, h) = 0,$$

for all $h \in U$ and hence the claim. □

In fact, if we assume that $\lambda_1 \neq 0$ we can show that λ_2 must in fact be a constant element of H. To simplify the notation, we introduce the pointwise powers for elements h of H by $h^n := m(h^{n-1}, h)$ for $n \geq 2$ and $h^1 := h$.

Proposition 5.3. *If the dimension of D_{LA} is finite and $\lambda_1 \neq 0$ then necessarily λ_2 is a constant, i.e. $\lambda_2 = c1$ where c is a real constant and 1 is the neutral element in H with respect to pointwise multiplication.*

Proof. As above we define recursively $[m(\lambda_2,\cdot),\lambda_1]^n := [m(\lambda_2,\cdot),[m(\lambda_2,\cdot),\lambda_1]^{n-1}]$ for $n \geq 2$ and $[m(\lambda_2,\cdot),\lambda_1]^1 := [m(\lambda_2,\cdot),\lambda_1]$. Then we have $[m(\lambda_2,\cdot),\lambda_1]^n(h) \in D_{LA}(h)$ for all $n \geq 1$ and $h \in U$. Now as

$$[m(\lambda_2,\cdot),\lambda_1]^n(h) = m(\lambda_2^n, \lambda_1), \quad n \geq 1,$$

by the same reasoning as in the proof of Proposition 5.2 we can find a $k \geq 1$ such that necessarily

$$\lambda_2^k = \sum_{i=1}^{k-1} a_i \lambda_2^i$$

for real numbers $a_1, ..., a_{k-1}$ which gives the claim. □

Thus we can look at the distribution given in (5.2) which now reduces to

$$D(h) = \langle \frac{d}{dx} h, \lambda_1, h \rangle, \quad h \in U, \tag{5.4}$$

as π is now given by $\pi(h) = \frac{d}{dx} h - \frac{1}{2} c^2 h$. Also note that, by denoting Id the linear map on H given by $Id(h) = h$, we have

$$Fl_t^{Id}(h) = e^{tId} h := \sum_{n=0}^{\infty} \frac{t^n Id^n}{n!} h = \sum_{n=0}^{\infty} \frac{t^n}{n!} Id(h) = e^t h. \tag{5.5}$$

Proposition 5.4. *If the conditions of Proposition 5.1 hold true and the dimension of the distribution* (5.4) *is constant on* U*, then* (5.1) *admits a generic finite dimensional realization around any* $u^* \in U$ *such that the solution can be represented by*

$$\Gamma((Z_t, t), u^*) = Z_t^0 \lambda_1 + e^{Z_t^1} S_t u^*, \tag{5.6}$$

where Z_t *is the time-inhomogeneous diffusion with values in* $\mathbb{R}^2$ *given by*

$$dZ_t^0 = bZ_t^0 \, dt + \phi(\Gamma((Z_t, t), u^*)) \, d\beta_t^1 + cZ_t^0 \, d\beta_t^2, \quad Z_0^0 = 0 \tag{5.7}$$

$$dZ_t^1 = -\frac{1}{2} c^2 \, dt + c \, d\beta_t^2, \quad Z_0^1 = 0, \tag{5.8}$$

where b, c *are the real constants given in Proposition 5.1.*

Proof. By (2.5) we have (cf. (5.5))

$$\Gamma(z, u^*) = Fl_{z_0}^{\lambda_1} \circ Fl_{z_1}^{Id} \circ Fl_{z_2}^{\frac{d}{dx}} u^* = z_0 \lambda_1 + e^{z_1} S_{z_2} u^*.$$

We find the coordinate processes in the usual way by following the steps in [34, Section 6.4]. We just show that (5.6) indeed solves (5.1). Let $\tilde{Z}_t^1 := e^{Z_t^1}$ and notice that

$$d\tilde{Z}_t^1 = c\tilde{Z}_t^1 \, d\beta_t^2.$$

By using the linearity of (stochastic) integrals and Itô's product formula, we get

$$\begin{aligned} \Gamma((t, Z_t), u^*) &= Z_t^0 \lambda_1 + \tilde{Z}_t^1 S_t u^* = \int_0^t bZ_s^0 \lambda_1 \, ds + \int_0^t \phi(\Gamma((Z_s, s), u^*)) \lambda_1 \, d\beta_s^1 \\ &+ \int_0^t cZ_s^0 \lambda_1 \, d\beta_s^2 + u^* + \int_0^t S_s u^* \, d\tilde{Z}_s^1 + \int_0^t \tilde{Z}_s^1 \frac{d}{dx} S_s u^* \, ds \\ &= u^* + \int_0^t \frac{d}{dx} (Z_s^0 \lambda_1 + \tilde{Z}_s^1 S_s u^*) \, ds + \int_0^t \phi(\Gamma((Z_s, s), u^*)) \lambda_1 \, d\beta_s^1 \\ &+ \int_0^t c(Z_s^0 \lambda_1 + \tilde{Z}_s^1 S_s u^*) \, d\beta_s^2 \\ &= u^* + \int_0^t \frac{d}{dx} (\Gamma((Z_s, s), u^*)) \, ds + \int_0^t \phi(\Gamma((Z_s, s), u^*)) \lambda_1 \, d\beta_s^1 \\ &+ \int_0^t c\Gamma((Z_s, s), u^*) \, d\beta_s^2 \end{aligned}$$

and hence the claim. □

5.2 Multiplicative Mixtures

Now we look at multiplicative mixtures by considering the following equation

$$\begin{cases} du_t = \frac{d}{dx}u_t\,dt + \sum_{j=1}^d \tilde{\phi}_j(u_t)m(u_t,\lambda_j)\,d\beta_t^j \\ u_0 \in U, \end{cases} \tag{5.9}$$

where U is an open convex set in $H^+ \subset H$ (cf. Lemma II.5.14) and $\tilde{\phi}_1, ..., \tilde{\phi}_d$ are scalar fields on U. It follows just as in Corollary 4.1 that $\exp(J_t)$ is a mild solution of (5.9) for any $J_0 \in U^l$ where J_t is a mild solution of

$$\begin{cases} dJ_t = (\frac{d}{dx}J_t - \frac{1}{2}\sum_{j=1}^d \phi_j^2(J_t)m(\lambda_j,\lambda_j))\,dt + \sum_{j=1}^d \phi_j(J_t)\lambda_j\,d\beta_t^j \\ J_0 \in U^l, \end{cases} \tag{5.10}$$

where $U^l := \log(U)$ (cf. Section 4 for the notation) and $\phi_j(h) = \tilde{\phi}_j(\exp(h))$ for $j = 1, ..., d$. In particular, (5.9) admits a generic finite dimensional realization for every $u^* \in U$ if and only if (5.10) does so for any $J^* \in U^l$. Thus we are looking at conditions such that (5.10) admits a generic finite dimensional realization. As in Section 4.1 it suffices to look at the distribution D on U^l given by

$$D(h) := \langle \pi(h), \lambda_1, ..., \lambda_d \rangle, \quad h \in U^l, \tag{5.11}$$

where

$$\pi(h) := \frac{d}{dx}h - \frac{1}{2}\sum_{j=1}^d \phi_j^2(h)m(\lambda_j,\lambda_j). \tag{5.12}$$

We are now in a very similar situation as in Section 4 and have a result that is analog to that of Proposition 4.4.

Proposition 5.5. *Let $\lambda_1, ..., \lambda_d$ be linearly independent and $p \in \{0, ..., d\}$ be such that*

$$\begin{cases} m(\lambda_j,\lambda_j) \in \langle \lambda_1, ..., \lambda_d \rangle, & \text{for } j = 1, ..., p, \\ m(\lambda_j,\lambda_j) \notin \langle \lambda_1, ..., \lambda_d \rangle, & \text{for } j = p+1, ..., d, \end{cases} \tag{5.13}$$

and $M \in \mathbb{R}^{p\times d}$ such that

$$m(\lambda_j,\lambda_j) = \sum_{l=1}^d M^{jl}\lambda_l, \quad \text{for } j = 1, ..., p. \tag{5.14}$$

Further let $\zeta_{ij} : U^l \to \mathbb{R}$ for for $(i,j) \in \{1, ..., d\}^2$ be the map given by

$$\zeta_{ij}(h) := -\phi_j(h)D\phi_j(h)\cdot\lambda_i. \tag{5.15}$$

Then the distribution (5.11) is involutive on U^l if and only if the following two conditions hold true:

1 *There is a matrix $\tilde{\zeta} \in \mathbb{R}^{d\times(d-p-1)}$ such that for $(i,j) \in \{1, ..., d\} \times \{p+1, ..., d\}$ $\zeta_{ij}(h) = \tilde{\zeta}_{ij}$ for for all $h \in U^l$.*

2 *There is a matrix $\tilde{B} \in \mathbb{R}^{d\times d}$ such that*

$$\frac{d}{dx}\lambda_i^1 = \sum_{j=1}^d \tilde{B}^{ij}\lambda_j - \sum_{j=p+1}^d \tilde{\zeta}_{ij}m(\lambda_j,\lambda_j), \quad i = 1, ..., d. \tag{5.16}$$

Proof. The distribution (5.11) is involutive on U^l if and only if $[\pi, \lambda_i](h) \in \langle \lambda_1, ..., \lambda_d \rangle$ for all $h \in U^l$. It follows that this is equivalent to (cf. Proposition 4.4) to the existence of a coordinate map $B : U^l \to \mathbb{R}^{d\times d}$, such that

$$\frac{d}{dx}\lambda_i = \sum_{l=1}^{d} \Big(B^{il}(h) - \sum_{j=1}^{p} \zeta_{ij}(h) M^{jl} \Big) \lambda_l - \sum_{j=p+1}^{d} \zeta_{ij}(h) m(\lambda_j, \lambda_j),$$

for $i = 1, ..., d$ and all $h \in U^l$. Now from (5.13), arguing as in Proposition 4.4 and setting

$$\tilde{B}^{il} := B^{il}(h) - \sum_{j=1}^{p} \zeta_{ij}(h) M^{jl}, \quad (i,l) \in \{1, ..., d\}^2$$

gives the claim. □

Note that Remark 4.6 applies here as well, that is, if the condition **1** of the last proposition holds true, then $\phi_{p+1}, ..., \phi_d$ satisfy the ODEs

$$-\phi_j D\phi_j \cdot \lambda_i = \tilde{\zeta}_{ij}, \quad \text{for all}(i,j) \in \{1, ..., d\} \times \{p+1, ..., d\}.$$

Solutions are necessarily of the form

$$\phi_j(h) = \sqrt{\gamma_j + l_j(h)}, \quad j = p+1, ..., d,$$

where each $l_j \in L(U^l, \mathbb{R})$ with $l_j(U^l) \subset [-\gamma_j, \infty)$ satisfying $-\frac{1}{2} l_j(\lambda_i) = \tilde{\zeta}_{ij}$. However, unlike the situation of Remark 4.6 the vectors $\lambda_1, ..., \lambda_d$ are now assumed to be linearly independent. Hence, if for one $j \in \{p+1, ..., d\}$, it is assumed that $\tilde{\zeta}_{1j}, ..., \tilde{\zeta}_{dj}$ are all zero, then by the Rank Nullity Theorem this l_j is necessarily constant and equaling zero as well and accordingly ϕ_j must be constants. The case where $\tilde{\zeta}_{ij} = 0$ for all $(i,j) \in \{1, ..., d\} \times \{p+1, ..., d\}$, which corresponds by **2** to the case where $\lambda_1, ..., \lambda_d$ are quasi-exponentials, we have hence necessarily that $\phi_{p+1}, ..., \phi_d$ are constants. Hence if we assume that $p = 0$ in (5.13) we end up in the situation where all $\phi_1, ..., \phi_d$ are constants and the equation (5.9) reduces to the equation (4.27) with $m = 0$.

Proposition 5.6. *If the conditions of Proposition 5.5 and the dimension of the distribution* (5.11) *is constant on* U^l*, then* (5.10) *admits a generic finite dimensional realization around every* $J^* \in U^l$ *and solutions can be represented by*

$$\Gamma(t, Z_t, J^*) = Fl_t^{\pi}(J^*) + \sum_{l=1}^{d} Z_t^l \lambda_l \tag{5.17}$$

where Z *is the* $\mathbb{R}^d$*-valued time-inhomogeneous diffusion process given as the solution of*

$$dZ_t^l = b_l(t, Z_t)\, dt + \sum_{j=1} \rho_j(t, Z_t)\, d\beta_t^j, \quad Z_0^l = 0, \quad l = 1, ..., d,$$

where

$$b_l(t,z) = \sum_{i=1}^{d} z_i \tilde{B}^{il} - \frac{1}{2} \sum_{j=1}^{p} \Big(\phi_j^2(\Gamma(t,z,J^*)) - \phi_j^2(Fl_t^{\pi}(J^*)) \Big) M^{jl}, \quad l = 1, ..., d$$

*with $\tilde{B}$ and M given in **2** and* (5.14) *and* $\rho_j(t,z) = \phi_j(\Gamma(t,z,J^*))$, *for* $j = 1, ..., d$. *In this case also* (5.9) *admits a generic finite dimensional realization around any* $u^* \in U$ *with representation*

$$\begin{aligned}\tilde{\Gamma}(t, Z_t, u^*) &:= \exp\Big(\Gamma(t, Z_t, \log(u^*))\Big) \\ &= m\Big(Fl_t^{\tilde{\pi}}(u^*), \exp(\sum_{l=1}^{d} Z_t^l \lambda_l)\Big),\end{aligned} \tag{5.18}$$

where $\tilde{\pi}$ *is the vector field on* U *given by*

$$\tilde{\pi}(h) = m(h, \pi(\log(h))) \tag{5.19}$$

Proof. The claim follows again from Theorem 2.2. Regarding the Representation (5.17) we proceed as in the proof of Theorem 4.9 and skip most details as it is essentially the same reasoning. It follows from (2.5) that

$$\Gamma(z, J^*) = Fl_{z_0}^{\pi}(J^*) + \sum_{j=1}^{d} z_j \lambda_j. \tag{5.20}$$

By using the **1** and **2** in Proposition 5.5 and the discussion preceding this Proposition, it follows just as in (4.29) in the proof of Theorem 4.9 that

$$\begin{aligned}&\pi(\Gamma(z, J^*)) - \pi(Fl_{z_0}^{\pi}(J^*)) \\ &\sum_{l=1}^{d} \Big[\sum_{i=1}^{d} z_i \tilde{B}^{il} - \frac{1}{2} \sum_{j=1}^{p} \Big(\phi_j^2(\Gamma(z, J^*)) - \phi_j^2(Fl_{z_0}^{\pi}(J^*))\Big) M^{jl}\Big] \lambda_l\end{aligned}$$

thus by choosing b and ρ as suggested the first claim follows. For the second claim, it remains to show that the Representation (5.18) holds true with $\tilde{\pi}$ given in (5.19). Indeed, we have

$$\exp\Big(\Gamma(t, Z_t, \log(u^*))\Big) = m\Big[\exp\Big(Fl_t^{\pi}(\log(u^*)\Big), \exp\Big(\sum_{l=1}^{d} Z_t^l \lambda_l\Big)\Big].$$

Thus by defining $Fl_t^{\tilde{\pi}}(u^*) := \exp\Big(Fl_t^{\pi}(\log(u^*)\Big)$ the claim follows from

$$\frac{d}{dt} Fl_t^{\tilde{\pi}}(u^*) = m\Big[Fl_t^{\tilde{\pi}}(u^*), \pi\Big(Fl_t^{\pi}(\log(u^*))\Big)\Big].$$

□

Example 5.7. Consider the equation (5.9) with $d = 2$ and $\lambda_1 = \exp(-b\cdot)$ and $\lambda_2 = \exp(-2b\cdot)$. Then in the setting of Proposition 5.5 we are in the situation where $p = 1$, as $m(\exp(-b\cdot), \exp(-b\cdot)) = \exp(-2b\cdot)$ with $M^{11} = 1$ and $M^{12} = 0$. Further according to **2** we have necessarily $\tilde{\zeta}_{12} = \tilde{\zeta}_{22} = 0$ and by the discussion after the Proposition 5.5 we have that ϕ_2 is necessarily a constant. Looking again at **2** we see that $\tilde{B}$ is a diagonal matrix with entries $-b$ and $-2b$. Thus it follows from (and under the assumptions of) the last Proposition that the equation

$$\begin{cases} du_t = \frac{d}{dx} u_t \, dt + \tilde{\phi}_1(u_t) m(u_t, \exp(-b\cdot)) \, d\beta_t^1 + \tilde{\phi}_2 m(u_t, \exp(-2b\cdot)) \, d\beta_t^2 \\ u_0 \in U, \end{cases}$$

is solved for any initial curve $u^* \in U$ by

$$\tilde{\Gamma}(t, Z_t) = m\Big[Fl_t^{\tilde{\pi}}(u^*), \exp\Big(Z_t^1 e^{-b\cdot} + Z_t^2 e^{-2b\cdot}\Big)\Big],$$

where

$$\tilde{\pi}(h) = m(h, \pi(\log(h)))$$

and

$$\pi(h) = \frac{d}{dx}h - \frac{1}{2}\tilde{\phi}^2(\exp(h))m(\lambda_1, \lambda_1) - \frac{1}{2}\tilde{\phi}_2^2 m(\lambda_2, \lambda_2).$$

The coordinate processes are given by

$$\begin{cases} dZ_t^1 &= \Big[-bZ_t^1 - \frac{1}{2}\Big(\tilde{\phi}_1^2\big(\tilde{\Gamma}((t, Z_t^1, Z_t^2), \log(u^*))\big) - \tilde{\phi}_1^2\big(Fl_t^{\tilde{\pi}}(u^*)\big)\Big)\Big]\,dt \\ & +\tilde{\phi}_1\big(\tilde{\Gamma}((t, Z_t^1, Z_t^2), \log(u^*))\big)\,d\beta_t^1, \quad Z_0^1 = 0 \end{cases} \tag{5.21}$$

and

$$dZ_t^2 = -2bZ_t^2\,dt + \tilde{\phi}_2\,d\beta_t^2, \quad Z_0^2 = 0.$$

The situation becomes particularly tractable by choosing (cf. Remark 4.6)

$$\tilde{\phi}_1(h) = \sqrt{(\gamma + l(\log(h)))}, \quad \gamma \in \mathbb{R}, \quad l \in L(U; \mathbb{R}), \tag{5.22}$$

such that l satisfies $l(\langle e^{-b\cdot}, e^{-2b\cdot}\rangle) \subset [-\gamma, \infty)$. In this case Z_t^1 reduces to

$$\begin{aligned} dZ_t^1 &= \Big[-bZ_t^1 - \frac{1}{2}Z_t^1 l(e^{-b\cdot}) - \frac{1}{2}Z_t^2 l(e^{-2b\cdot})\Big]\,dt \\ &+ \sqrt{\gamma + l\big(Fl_t^{\pi}(\log(u^*)) + Z_t^1 e^{-b\cdot} + Z_t^2 e^{-2b\cdot}\big)}\,d\beta_t^1, \quad Z_0^1 = 0 \end{aligned}$$

and π becomes an affine vector field. The *time-of-maturity* forward variance $v(t, T) := u_t(T - t)$ has the representation

$$v(t, T) = \exp\Big(Fl_t^{\pi}(\log(u^*))(T - t) + Z_t^1 e^{-b(T-t)} + Z_t^2 e^{-2b(T-t)}\Big).$$

Notice that this looks very much like the Bergomi model (II.5.20) but here the first coordinate Z_t^1 must not follow an Ornstein-Uhlenbeck process but can be chosen to be of the form (5.21). Also the Flow $Fl_t^{\pi}(\log(u^*))(T - t)$ differs as Lemma II.5.10 does not hold in the present case.

We see that both mixtures, additive and multiplicative, of the basic diffusion coefficients given by $\phi(\cdot)\lambda_1$ and $m(\lambda_2, \cdot)$, lead in some cases to generic finite dimensional realization. While in the additive case the restrictions are quite strong, that is λ_2 must be a constant element in H, the restrictions in the multiplicative case are very much case comparable to the case of a stochastic parameter as considered in Section 4.

6 Relation to the Fractional Bergomi Model

This section is a short outlook to future research. In a series of recent papers, including [5], Jim Gatheral and co-workers suggest that the logarithm of the spot-variance (under the physical measure) should be modeled as a *fractional Brownian motion* which is justified by the empirical evidence found in [44]. Starting from this assumption and under a change of measure with a deterministic Girsanov density (see [5, Section 3]) they arrive at the *Rough Bergomi (rBergormi)* model with spot-variance process (in the special case of $t = 0$, see [5, Section 4]) given by

$$u_t(0) = u_0(t)\exp\Big(\omega\beta_t^H - \frac{1}{2}\omega^2\mathbb{E}[(\beta_t^H)^2]\Big), \tag{6.1}$$

where $\omega > 0$ is a parameter and

$$\beta_t^H := \int_0^t \frac{d\beta_s}{(t-s)^\gamma}, \tag{6.2}$$

corresponds to a truncated version of the fractional Browian motion in the representation suggested by Mandelbrot and Van Ness (see [59]) with Hurst parameter $H = \frac{1}{2} - \gamma$, which shares most properties with it (cf. [60, Section 3] and [24, Definition 1]). They further note in [5, Page 11]

> Specifically, this rBergomi model is non-Markovian in the instantaneous variance ... but is Markovian in the (infinite-dimensional) state vector ... $\xi_t(u)$.

where in our notation $\xi_t(u)$ corresponds to $u_t(u-t)$. In the following we will recapture the Representation (6.1) and show this last point from the point of view of understanding variance curve models u as mild solution of the SPDE (4.1) in a suitable Hilbert space H. Subsequently we will discuss a method to approximate this process by linear-combinations of Bergomi-type processes.

6.1 The Markov property of the rBergomi Model

Under the conditions of Corollary 4.1 mild solutions of (4.1) can be represented as $u_t = \exp(J_t)$ where J is the mild solution of (4.3) (in both cases we assume here $m = 0$ and $d = 1$). We assume that this conditions are met for the choice $\sigma(x) = \omega\frac{1}{x^\gamma}$, for some $\omega > 0$, on a suitable Hilbert space H. In this case we can write J (i.e. the logarithm of the forward variance) as $J = J^1 + J^2$ (cf. Lemma II.5.10) where J^1 and J^2 are given in the mild representations by

$$\begin{aligned} J_t^1 &= \omega\int_0^t S_{t-s}\sigma\, d\beta_s, \quad \text{and} \\ J_t^2 &= S_t J_0 - \frac{1}{2}\omega^2\int_0^t S_{t-s} m(\sigma,\sigma)\, ds, \end{aligned} \tag{6.3}$$

where J^1 is an infinite dimensional Ornstein-Uhlenbeck process (see [25]) and J^2 is the mild solution of a deterministic PDE. Hence J and also J^1 correspond under some mild conditions (see [26, Chapter 9]) to an infinite dimensional Markov process in H. By noting that $\frac{J_t^1(0)}{\omega}$ corresponds to the (truncated) fractional Brownian motion given in (6.2) (recall the choice $\sigma(x) = \frac{1}{x^\gamma}$), that is,

$$J_t^1(0) = \omega\int_0^t S_{t-s}\sigma(0)\, d\beta_s = \omega\int_0^t \frac{1}{(t-s)^\gamma}\, d\beta_s = \omega\beta_t^H,$$

we find that the spot variance is given by

$$\begin{aligned} u_t(0) &= m(\exp(J_t^1), \exp(J_t^2))(0) = \exp(J_t^1(0) + J_t^2(0)) \\ &= u_0(t)\exp\Big(\omega\beta_t^H - \frac{1}{2}\int_0^t \omega^2\frac{1}{(t-s)^{2\gamma}}\, ds\Big), \end{aligned} \tag{6.4}$$

where $u_0(t) := \exp(J_0(t))$ equals the Representation given in (6.1) (notice that the integral converges for $\gamma < \frac{1}{2}$, i.e. for positive Hurst parameters) with the property that the forward variance curve is Markov with respect to the filtration generated by J^1.

6.2 Approximation of the rBergomi Model

Finally we show (formally) how to approximate this processes with real valued Ornstein Uhlenbeck processes by following the ideas given in [16] and [17] (see also the recent article [49]). By defining $\mu(d\eta) := \eta^{\gamma-1}\,d\eta$ and denoting by $\mathcal{L}_\mu$ the Laplace transform with respect to the measure μ, we can represent $\sigma(x) = \frac{1}{x^\gamma}$ by

$$\sigma(x) = \mathcal{L}_\mu \mathcal{L}_\mu^{-1}[\sigma](x) = \frac{1}{\Gamma(\gamma)} \int_{\mathbb{R}_+} e^{-x\eta}\,\mu(d\eta),$$

and by applying Fubini's Theorem we accordingly get for (6.3)

$$J_t^1 = \tilde{\omega} \int_0^t S_{t-s}\Big(\int_{\mathbb{R}_+} e^{-\eta\cdot}\,\mu(d\eta)\Big)\,d\beta_s = \tilde{\omega} \int_{\mathbb{R}_+} \Big(\int_0^t S_{t-s} e^{-\eta\cdot}\,d\beta_s\Big)\,\mu(d\eta), \tag{6.5}$$

where $\tilde{\omega} := \frac{\omega}{\Gamma(\gamma)}$. Notice that the expression in the brackets corresponds for each $\eta > 0$ to the mild solution of

$$dX_t^\eta = \frac{d}{dx} X_t^\eta\,dt + e^{-\eta\cdot}\,d\beta_t, \quad X_0^\eta = 0,$$

for which we know (cf. Theorem 3.9) that it can be represented by $X_t^\eta = Z_t^\eta e^{-\eta\cdot}$, where Z^η is the Ornstein-Uhlenbeck process satisfying

$$dZ_t^\eta = -\eta Z_t^\eta\,dt + d\beta_t, \quad Z_0 = 0. \tag{6.6}$$

Thus we can formally write J^1 as

$$J_t^1 = \tilde{\omega} \int_{\mathbb{R}_+} Z_t^\eta e^{-\eta\cdot}\,\mu(d\eta). \tag{6.7}$$

Using (6.7) we can represent (6.4) as (letting $u_0 = u$)

$$u_t \;=\; S_t u \,\exp\Big(\tilde{\omega} \int_{\mathbb{R}_+} Z_t^\eta e^{-\eta\cdot}\,\mu(d\eta) - \frac{1}{2}\int_0^t \tilde{\omega}^2 \frac{1}{(t-s)^{2\gamma}}\,ds\Big). \tag{6.8}$$

Remark 6.1. The Represetation (6.7) was found in [16] for the fractional Brownian motion which corresponds to the process $J_t^1(0)$, as in this case the expression in the brackets of (6.5) is given by $\int_0^t e^{-\eta(t-s)}\,d\beta_s$ which can immediately be recognized as the solution of (6.6). The underlying infinite dimensional Markov process is here given by $\mathcal{Y}_t := (Z_t^\eta; \eta > 0)$ (cf. [16, Proposition 1]).

Similarly, by starting with the representation of the mild solution of (4.1) and with the same reasoning as above we find

$$\begin{aligned} u_t &= S_t u + \omega \int_0^t S_{t-s} m(\sigma, u_s)\,d\beta_s \\ &= \int_{\mathbb{R}_+} S_t \mathcal{L}_\mu^{-1}[u](\eta) e^{-\eta\cdot}\,\mu(d\eta) + \tilde{\omega}\int_0^t S_{t-s} m\Big(\int_{\mathbb{R}_+} e^{-\eta\cdot}\,\mu(d\eta), u_s\Big)\,d\beta_s \\ &= \int_{\mathbb{R}_+} \Big(S_t \mathcal{L}_\mu^{-1}[u](\eta) e^{-\eta\cdot} + \int_0^t S_{t-s} m(\tilde{\omega} e^{-\eta\cdot}, u_s)\,d\beta_s\Big)\mu(d\eta), \end{aligned} \tag{6.9}$$

where for each $\eta > 0$ the expression in the brackets of (6.9) corresponds to the mild solution of

$$\begin{cases} du_t^\eta &= \frac{d}{dx} u_t^\eta \, dt + m(\tilde{\omega} e^{-\eta \cdot}, u_t^\eta) \, d\beta_t \\ u_0^\eta &= \mathcal{L}_\mu^{-1}[u](\eta) e^{-\eta \cdot}, \end{cases}$$

and hence according to Theorem 4.9 admits a finite dimensional realization and we can write

$$u_t \quad = \quad \int_{\mathbb{R}_+} \exp\Big(Fl_t^{\pi^\eta}(\log(\mathcal{L}_\mu^{-1}[u](\eta) e^{-\eta \cdot})) + Z_t^\eta \tilde{\omega} e^{-\eta \cdot} \Big) \mu(d\eta), \tag{6.10}$$

where Z_t^η agrees with the solution of (6.6) and $Fl_t^{\pi^\eta}(\mathcal{L}_\mu^{-1}[u](\eta) e^{-\eta \cdot})$ denotes the flow of (cf. (4.34))

$$\pi^\eta(h) := \frac{d}{dx} h - \frac{1}{2} \tilde{\omega}^2 m(e^{-\eta \cdot}, e^{-\eta \cdot})$$

at $\log(\mathcal{L}_\mu^{-1}[u](\eta) e^{-\eta \cdot})$. If we now formally approximate the measure $\mu(d\eta)$ by a sum of Dirac measures $\mu^N(d\eta) := \sum_{i=1}^N c_i \delta_{d_i}(d\eta)$ as in [17] (see also [16, Page 99] for the approximation of fractional Brownian motion) for suitably chosen $(c_1, d_1), ..., (c_N, d_N)$ we can approximate the process (6.7) by $\tilde{\omega} \sum_{i=1}^N c_i Z_t^{d_i} e^{-d_i \cdot}$ which plugged into (6.8) gives a process of Bergomi-type and hence admits a finite dimensional representation. Similarly, we can approximate (6.10) by

$$\sum_{i=1}^N c_i \exp\Big(Fl_t^{\pi^{d_i}}(\log(\mathcal{L}_\mu^{-1}[u](d_i) e^{-d_i \cdot})) + Z_t^{d_i} \tilde{\omega} e^{-d_i \cdot} \Big),$$

which is just a linear combination of conventional Bergomi processes. The proof of (weak) convergence of this approximation as $N \to \infty$ is work in progress and will be published soon.

Chapter IV

Weak Taylor Expansions for SPDEs

1 Introduction

Let H be a suitable Hilbert space and consider the following parameterized family of SPDEs

$$\begin{cases} dX_t^\epsilon & = (AX_t^\epsilon + V(\epsilon, X_t^\epsilon))\, dt + \sum_{i=1}^d V_i(\epsilon, X_t^\epsilon)\, d\beta_t^i \\ X_0 & \in H, \end{cases} \tag{1.1}$$

where for each $\epsilon \in \mathbb{R}$, $V(\epsilon, \cdot), V_1(\epsilon, \cdot), ..., V_d(\epsilon, \cdot)$ are sufficiently regular vector-fields on H. For $l \in L(H, \mathbb{R}^N)$ for some $N \in \mathbb{N}$ we will be interested in the weak approximation of $l \circ X_T^\epsilon$ as $\epsilon \to 0$. This approximation will be particularly useful if X^ϵ admits a finite dimensional realization only for $\epsilon = 0$. In [68, Theorem 2.3] a *weak Taylor approximation* is introduced that provides sufficient conditions for the process $l \circ X_T^\epsilon$ such that this approximation converges at arbitrary order. The main conditions are that $l \circ X_T^\epsilon$ is smooth in the Malliavin sense and that the corresponding *Malliavin-Covariance Matrix* is invertible with a p-integrable inverse for every $p \geq 1$. As we are here in an infinite dimensional setting, especially the latter condition is far from trivial. As we could not find a suitable source for conditions on the vector fields of an SPDE such that its solution is smooth in the Malliavin sense (for Malliavin differentiability of mild solutions of SPDEs see for example [19] and [18] (see here the Remark 2.10) and [67]) we recapture the corresponding notions. The theory is similar the finite dimensional case (see [63], from which also most of the notation is adopted).

2 The Malliavin Derivative

We will mainly follow [58] which in turn is mainly based on [46]. For the finite-dimensional case, all results can be found in [63]. Let $T > 0$ be some finite real number and $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ a complete filtered probability space acommodating a d-dimensional Brownian motion $(W_t)_{t\in[0,T]}$ such that the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ is generated by the Brownian motion. For $\phi \in L^2([0,T]; \mathbb{R}^d)$ we introduce the *isonormal Gaussian process*

$$W(\phi) := \sum_{i=1}^d \int_0^T \phi^i(t)\, dW_t^i. \tag{2.1}$$

Let H be a separable Hilbert space. By $\mathcal{S}(H)$ we define the set of smooth cylindrical H-valued random variables F

$$F = \sum_{j=1}^{n} f_j(W(\phi_1), ..., W(\phi_m))h_j,$$

where $m, n \in \mathbb{N}$, $\phi_i \in L^2([0,T];\mathbb{R}^d)$ for $i = 1, ..., m$ and $W(\phi_i)$ is given by (2.1), $h_j \in H$, $f_j \in C_p^\infty(\mathbb{R}^m)$, for $j = 1, ..., n$. We have $\mathcal{S}(H) \subset L^p(\Omega; H)$ for all $p \geq 1$. The (Malliavin) derivative of $F \in \mathcal{S}(H)$ is defined by

$$D_t F := \sum_{j=1}^{n} \sum_{i=1}^{m} \partial_i f_j(W(\phi_1), ..., W(\phi_m))h_j \otimes \phi_i(t), \quad t \in [0,T],$$

where $h_j \otimes \phi_i(t)$ denotes the algebraic tensor product of $h_j \in H$ and $\phi(t) \in \mathbb{R}^d$, which can be identified with an element of $L_2^0 := L_2(\mathbb{R}^d; H)$, the space of Hilbert-Schmidt operators, by

$$(h_j \otimes \phi_i(t))(x) = (\phi_i(t)(x))h_j = (\sum_{l=1}^{d} \phi_i^l(t)x^l)h_j \in H, \quad \forall x \in \mathbb{R}^d$$

and in particular $(h_j \otimes \phi_i(t))(e_l) = \phi_i^l(t)h_j$, where $e_1, ..., e_d$ denote the standard basis vectors of $\mathbb{R}^d$. This way we can, like in the finite dimensional case (i.e. $H = \mathbb{R}$, cf. [63]), introduce the *partial derivatives* $D_t^l F$, defined by $D_t^l F := D_t F e_l$, $l = 1, ..., d$, that is

$$D_t^l F := \sum_{j=1}^{n} \sum_{i=1}^{m} \partial_i f_j(W(\phi_1), ..., W(\phi_m))h_j \phi_i^l(t), \quad t \in [0,T]. \tag{2.2}$$

This way we identify $H \otimes \mathbb{R}^d$ with H^d and understand the derivative operator as the map

$$D : \mathcal{S}(H) \subset L^p(\Omega; H) \to L^p(\Omega; L^2([0,T]; H^d)),$$

that is, for each $l = 1, ..., d$, we consider the partial derivative operator

$$D^l : \mathcal{S}(H) \subset L^p(\Omega; H) \to L^p(\Omega; L^2([0,T]; H))$$

with $D_t^l F$ given by (2.2). Higher derivatives are defined by iterating this procedure, that is, for $F \in \mathcal{S}(H)$ and $k \geq 1$ we define recursively

$$D_{t_1,...,t_k}^k F = \sum_{i=1}^{m} \partial_i D_{t_1,...,t_{k-1}}^{k-1} F \otimes \phi_i(t_k).$$

Remark 2.1. As higher derivatives are denoted by superscripts just as the partial derivatives, we will use the letter k for higher derivatives and the letter l for partial derivatives to distinguish between this notions.

Proposition 2.2. *For each $k \geq 1$ and $p > 1$, the operator $D^k : \mathcal{S}(H) \subset L^p(\Omega; H) \to L^p(\Omega; L^2([0,T]^k; H^d))$ is closable. The domain of the (extended) derivative operator D^k is denoted by $\mathbb{D}^{k,p}(H)$ with norm*

$$\|F\|_{\mathbb{D}^{k,p}(H)} := \Big(\mathbb{E}[\|F\|_H^p] + \sum_{j=1}^{k} E[\|D^j F\|_{L^2([0,T]^j; H^d)}^p]\Big)^{\frac{1}{p}} \tag{2.3}$$

Proof. This is [58, page 92]. □

Remark 2.3. We can also take another perspective which will allow us in some cases to directly generalize finite dimensional results (i.e. $\mathbb{H} = \mathbb{R}$), given for example in [63], to the infinite dimensional case. As we have $\mathcal{S}(H) = \mathcal{S} \otimes H$, where $\mathcal{S}$ denotes the set of real-valued smooth cylindrical random variables, we have for any decomposition $F := \tilde{F} \otimes h \in \mathcal{S} \otimes H = \mathcal{S}(H)$ given by

$$F = f(W(\phi_1), ..., W(\phi_m)) \otimes h,$$

with derivative given by

$$DF = D(\tilde{F}) \otimes h.$$

This way, we can in the setting of Proposition 2.2 understand the operator D^k as the map

$$D^k : \mathcal{S} \otimes H \subset L^p(\Omega; \mathbb{R}) \otimes H \to L^p(\Omega; L^2([0,T]^k; \mathbb{R}^d)) \otimes H.$$

Lemma 2.4. *Let $\phi : H \to \tilde{H}$ be a Frechét-differentiable mapping, where $\tilde{H}$ is an arbitrary separable Hilbert space. Assume that there exists a $q \geq 0$ and a constant $C > 0$ such that*

$$\|\phi(h)\|_{\tilde{H}} \leq C(1 + \|h\|_H^{1+q}), \quad \|\phi'(h)\|_{L(H,\tilde{H})} \leq C(1 + \|h\|^q), \quad \text{for all } h \in H.$$

Then, for all $p > 1$ and $F \in \mathbb{D}^{1,(q+1)p}(H)$ it holds that $\phi(F) \in \mathbb{D}^{1,p}(\tilde{H})$ and

$$D\phi(F) = \phi'(F)DF.$$

In particular, if $F \in \mathbb{D}^{1,\infty}(H)$ then $\phi(F) \in \mathbb{D}^{1,\infty}(\tilde{H})$.

Proof. This is [58, Lemma 4.7]. □

Lemma 2.5. *Let $p > 1$ and F_n be a sequence in $\mathbb{D}^{1,p}(H)$ such that $F_n \to F$ in $L^p(\Omega; H)$ and there exists a constant C such that*

$$\sup_n \|DF_n\|_{L^p(\Omega; L^2([0,T]; H^d))} \leq C.$$

Then $F \in \mathbb{D}^{1,p}(H)$ and $\|DF\|_{L^p(\Omega; L^2([0,T]; H^d)} \leq C$. Moreover, there exists a subsequence $(n_k)_{k\geq 1}$ such that weakly $DF_{n_k} \to DF$ in $L^p(\Omega; L^2([0,T]; H^d)$.

Proof. This is [64, Lemma 3.7]. □

The adjoint $\delta : dom(\delta) \subset L^2(\Omega; L^2([0,T]; H^d)) \to L^2(\Omega; H)$ of the Malliavin derivative is called divergence operator. Its domain $dom(\delta)$ consists of all

$$\Psi \in L^2(\Omega; L^2([0,T]; H^d))$$

such that there exists a constant $C = C(\Psi) > 0$ with

$$|E[\langle DF, \Psi \rangle_{L^2([0,T]; H^d)}]| \leq C\|F\|_{L^2(\Omega; H)}, \quad \text{for all } F \in \mathbb{D}^{1,2}(H).$$

In this case, for $\Psi \in dom(\delta)$ it is defined as the unique element $\delta(\Psi)$ in $L^2(\Omega; H)$ satisfying

$$E[\langle DF, \Psi \rangle_{L^2([0,T]; H^d)}] = E[\langle F, \delta(\Psi) \rangle_H], \quad \forall F \in \mathbb{D}^{1,2}(H).$$

Proposition 2.6. *Let $\alpha \in L^2(\Omega; L^2([0,T]; H^d))$ be a predictable stochastic process with values in H^d. Then $\alpha \in dom(\delta)$ and*

$$\delta(\alpha) = \sum_{i=1}^{d} \int_0^T \alpha^i(s)\, dW_s^i,$$

that is, $\delta(\alpha) \in L^2(\Omega; H)$ and it coincides with the stochastic Itô-integral.

Proof. This is [58, Proposition 4.12]. □

For the following we introduce the space $\mathbb{L}^{1,2}(H^d) := \mathbb{D}^{1,2}(L^2([0,T]; H^d))$ which is isomorphic to $L^2([0,T]; \mathbb{D}^{1,2}(H^d))$ and accordingly we have the following characterization which we state as a Lemma.

Lemma 2.7. *We have $\alpha \in \mathbb{L}^{1,2}(H^d)$ if and only if, $\alpha \in L^2([0,T] \times \Omega; H^d)$, $\alpha_t \in \mathbb{D}^{1,2}(H^d)$ for almost all $t \in [0,T]$ and*

$$\sum_{i=1}^{d} \mathbb{E}\Big[\int_0^T \int_0^T \|D_t \alpha_s^i\|_H^2 \, ds\, dt\Big] < \infty \tag{2.4}$$

Proof. This follows from the isomorphy of $\mathbb{L}^{1,2}(H^d)$ to $L^2([0,T]; \mathbb{D}^{1,2}(H^d))$ and the definition of $\|\cdot\|_{\mathbb{D}^{1,2}(H)}$ given in (2.3). □

Proposition 2.8. *Let $f \in \mathbb{L}^{1,2}(H)$ be a predictable process. Then we have*

$$\int_0^T f(s)\, ds \in \mathbb{D}^{1,2}(H)$$

and

$$D_t^l \int_0^T f(s)\, ds = \int_t^T D_t^l f(s)\, ds$$

for $l = 1, ..., d$ and almost all $t \in [0,T]$.

Proof. This is [58, Proposition 4.8]. □

Further we have $\mathbb{L}^{1,2}(H^d) \subset Dom(\delta)$ which follows from [58, Proposition 4.15] and [46, Proposition 3.2]. Now we can state the desired result, which can be found in [58, Proposition 4.16] if α is in $\mathbb{D}^{2,2}(L^2([0,T]; H^d))$ but is not necessarily predictable.

Proposition 2.9. *If $\alpha \in \mathbb{L}^{1,2}(H^d)$ is a predictable process then $\delta(\alpha) \in \mathbb{D}^{1,2}(H)$ and*

$$D_t^l \sum_{i=1}^{d} \int_0^T \alpha^i(s)\, dW_s^i = \alpha^l(t) + \sum_{i=1}^{d} \int_t^T D_t^l \alpha^i(s)\, dW_s^i$$

for $l = 1, ..., d$ and almost all $t \in [0,T]$.

Proof. The claim follows from [46, Proposition 3.4] in the same way as in the finite dimensional case [63, Lemma 1.3.4] follows from [63, Proposition 1.3.8]. In fact, according to [46, Proposition 3.4], the claim holds if for almost all t, $(s \mapsto D_t^l \alpha_s) \in dom(\delta)$ and $(t \mapsto \sum_{i=1}^{d} \int_0^T D_t^l \alpha_s^i \, dW_s^i) \in L^2(\Omega \times [0,T]; H)$ for each $l = 1, ..., d$. Now, as $(s \mapsto D_t^l \alpha_s) \in L^2(\Omega \times [0,T]; H^d)$ is predictable it follows from Proposition 2.6 that

$s \mapsto D_t^l \alpha_s \in dom(\delta)$ for almost all $t \in [0,T]$. For the second claim, it follows from the Itô isometry and property (2.4) that

$$\sum_{i=1}^{d} \mathbb{E}\Big[\int_0^T \Big\| \int_t^T D_t^l \alpha_s^i \, dW_s^i \Big\|_H^2 dt\Big] = \sum_{i=1}^{d} \mathbb{E}\Big[\int_0^T \int_t^T \|D_t^l \alpha_s^i\|_H^2 \, ds \, dt\Big] < \infty.$$

□

Remark 2.10. The previous Proposition is stated in a stronger version in [19, Proposition 5.4] and [18, Proposition 4.7], where besides predictability, the equivalent conditions of Lemma 2.7 excluding the integrability condition (2.4) are claimed to be sufficient. Unfortunately no proof is provided. Also it seems to contradict [63, Lemma 1.3.4] which states that in the finite dimensional case, that is $H = \mathbb{R}$, given the same conditions as in [19, Proposition 5.4] and [18, Proposition 4.7], the conditions of Proposition 2.9 are not only sufficient but also necessary.

Proposition 2.11. *Let F be a random variable in $\mathbb{D}^{k,\alpha}(H)$ with $\alpha > 1$. If $D^i F \in L^p(\Omega; L^2([0,T]^k; H^d))$ for $i = 0, 1, ..., k$ and for some $p > \alpha$, then $F \in \mathbb{D}^{k,p}(H)$.*

Proof. This can be proved as in the finite dimensional case [63, Proposition 1.5.5] by using the Remark 2.3. □

The next result can be found in a similar version in [19, Lemma 5.3] (however, see Remark 2.10) and [67, Theorem 3.1].

Lemma 2.12. *Assume that ξ is deterministic and $V, V_1, , ..., V_d \in C_b^\infty(H; H)$ in* (I.3.2). *Then the conditions of Theorem I.3.2 are satisfied. Then for any $t \in [0,T]$, the mild solution X_t is in $\mathbb{D}^{1,\infty}(H)$ and for $l = 1, ..., d$, the derivative $D_s^l X_t$ for fixed $s \in [0,T]$ satisfies the equation in H given by*

$$\begin{cases} D_s^l X_t & = S_{t-s} V_l(X_s) + \int_s^t S_{t-u} dV(X_u) D_s^l X_u \, du \\ & \quad + \sum_{i=1}^d \int_s^t S_{t-u} dV_i(X_u) D_s^l X_u \, dW_u^i \end{cases} \tag{2.5}$$

for $s \le t$ and $D_s^l X_t = 0$ for $s > t$. For $t \ge s$ it is a continuous mild solution of

$$\begin{cases} dY_t & = (AY_t + dV(X_t) Y_t) \, dt + \sum_{i=1}^d dV_i(X_t) Y_t \, dW_t^i \\ Y_s & = V_l(X_s). \end{cases} \tag{2.6}$$

Finally we have for all $p \ge 2$ and $l = 1, ..., d$ that

$$\sup_{r \in [0,T]} \mathbb{E}[\sup_{s \in [r,T]} \|D_r^l X_s\|_H^p] < \infty. \tag{2.7}$$

Proof. We follow the proof of [63, Theorem 2.2.1]. Let $p > 2$ and $t \in [0,T]$ be arbitrary. By the stated conditions on the vector fields $V, V_1, ..., V_d$ the equation (I.3.2) has a unique continuous mild solution given as the limit of the Picard iterations (I.3.3) in the space $\mathcal{H}^p$. We will show by induction that for each $k \ge 0$

$$X_t^k \in \mathbb{D}^{1,p}(H), \tag{2.8}$$

$$\psi_k(t) := \sup_{0 \le r \le t} \mathbb{E}\Big[\sup_{s \in [r,t]} \|D_r X_s^k\|_H^p\Big] < \infty, \tag{2.9}$$

$$\psi_{k+1}(t) \le c_1 + c_2 \int_0^t \psi_k(s) \, ds. \tag{2.10}$$

The statement is obviously true for $k = 0$. Now assume it holds for some $k > 0$. The next iteration is given by

$$X_t^{k+1} = S_t\xi + \int_0^t S_{t-s}V(X_s^k)\,ds + \sum_{i=1}^d \int_0^t S_{t-s}V_i(X_s^k)\,dW_s^i.$$

Applying Lemma 2.4 twice (with $q = 0$) we find that $S_{t-s}U(X_s^k) \in \mathbb{D}^{1,p}(H)$ for almost all $s \in [0,t]$ and

$$D_r^l S_{t-s}U(X_s^k) = S_{t-s}dU(X_s^k)D_r^l X_s^k,$$

where $U \in \{V, V_1, ..., V_d\}$. To apply Propositions 2.8 and 2.9 we need to show that $[0,t] \ni s \mapsto S_{t-s}U(X_s^k) \in \mathbb{L}^{1,2}(H)$. But this follows from Lemma 2.7 if both

$$\begin{aligned}
&\mathbb{E}[\int_0^t \|S_{t-s}U(X_s^k)\|_H^2\,ds] < \infty, \quad \text{and}\\
&\mathbb{E}[\int_0^t \int_r^t \|S_{t-s}dU(X_s^k)D_r^l X_s^k\|_H^2\,ds\,dr] < \infty
\end{aligned}$$

are satisfied. The first condition immediately follows from the assumed linear growth condition on U, that is, by letting $M := \sup_{s\in[0,t]} \|S_{t-s}\|_H$, we have

$$\mathbb{E}[\int_0^t \|S_{t-s}U(X_s^k)\|_H^2\,ds] \le tM^2C_1^2(1 + \sup_{s\in[0,t]} \mathbb{E}[\|X_s^k\|_H^2]) < \infty,$$

as $X^k \in \mathcal{H}^p \subset \mathcal{H}^2$. Similarly, the second condition follows from (2.9), as $U \in C_b^\infty(H)$, we have

$$\mathbb{E}[\int_0^t \int_r^t \|S_{t-s}dU(X_s^k)D_r^l X_s^k\|_H^2\,ds\,dr] \le M^2C_2^2 \sup_{0\le r\le t} \mathbb{E}[\sup_{s\in[r,t]} \|D_r^l X_s^k\|_H^2] < \infty,$$

where $C_2 > 0$ is given by $\sup_{x\in H} \|dU(x)\|_{L(H;H)} < C_2$. Hence the conditions of Lemma 2.7 are satisfied and it follows from Propositions 2.8 and 2.6 that $X_t^{k+1} \in \mathbb{D}^{1,2}(H)$ and

$$\begin{cases} D_r^l X_t^{k+1} & = S_{t-r}V^l(X_r^k) + \int_r^t S_{t-u}dV_0(X_u^k)D_r^l X_u^k\,du \\ & \quad + \sum_{i=1}^d \int_r^t S_{t-u}dV_i(X_u^k)D_r^l X_u^k\,dW_u^i. \end{cases} \tag{2.11}$$

Next we show that (2.9) and (2.10) hold for $k+1$ and conclude from Proposition 2.11 that $X_t^{k+1} \in \mathbb{D}^{1,p}(H)$. From an application of Hölder's inequality we get

$$\begin{aligned}
& \|D^l X_t^{k+1}\|_{L^p(\Omega;L^2([0,T];H))}^p && (2.12)\\
= \;& \mathbb{E}[\|D^l X_t^{k+1}\|_{L^2([0,T];H)}^p] = \mathbb{E}[\Big(\int_0^t \|D_r^l X_t^{k+1}\|_H^2\,dr\Big)^{\frac{p}{2}}] \\
\le \;& t^{p-2}\int_0^t \mathbb{E}[\|D_r^l X_t^{k+1}\|_H^p]\,dr \le t^{p-1} \sup_{r\in[0,t]} \mathbb{E}[\|D_r^l X_t^{k+1}\|_H^p] \\
\le \;& t^{p-1} \sup_{r\in[0,t]} \mathbb{E}[\sup_{s\in[r,t]} \|D_r^l X_s^{k+1}\|_H^p]. && (2.13)
\end{aligned}$$

By using (2.11) we get

$$
\begin{aligned}
&\mathbb{E}[\sup_{s\in[r,t]} \|D_r^l X_s^{k+1}\|_H^p] \le \mathbb{E}[\sup_{s\in[r,t]} \|S_{s-r}V_l(X_r^k)\|_H^p] \\
+\; &\mathbb{E}\Big[\sup_{s\in[r,t]} \Big\| \int_r^s S_{s-u} dV(X_u^k) D_r^l X_u^k \, du \Big\|_H^p\Big] \\
+\; &\sum_{i=1}^d \mathbb{E}\Big[\sup_{s\in[r,t]} \Big\| \int_r^s S_{s-u} dV_i(X_u^{k-1}) D_r^l X_u^k \, dW_u^i \Big\|_H^p\Big] =: I_1 + I_2 + I_3.
\end{aligned}
$$

Let $M := \sup_{t\in[0,T]} \|S_t\|_{L(H;H)}$. For the first term, using the linear growth condition on V, we deduce

$$
I_1 \le M^p \mathbb{E}[\|V^l(X_r^{k-1})\|_H^p] \le M^p C_1^p (1 + \mathbb{E}[\|X_r^k\|_H^p]) < \infty,
$$

as $X^k \in \mathcal{H}^p$. By denoting $M_0 := \sup_{x\in H} \|dV(x)\|_{L(H;H)}$, which by assumption is finite, we get for the second term by another application of Hölder's inequality that

$$
I_2 \le \mathbb{E}\Big[\Big(\int_r^t \|S_{t-u} dV(X_u^k) D_r^l X_u^k\|_H \, du\Big)^p\Big] \le M^p M_0^p (t-r)^p \mathbb{E}[\int_r^t \|D_r^l X_u^k\|_H^p \, du].
$$

For the third term, using [26, Lemma 7.2] and proceeding as with the second term, by denoting $M_i := \sup_{x\in H} \|dV_i(x)\|_{L(H;H)}$, $i = 1, ..., d$, we get

$$
\begin{aligned}
I_3 &\le C_2 \sum_{i=1}^d \mathbb{E}\Big[\Big(\int_r^t \|S_{t-u} dV_i(X_u^k) D_r^l X_u^k\|_H^2 \, du\Big)^{p/2}\Big] \\
&\le M^p \bar{M} (t-r)^p \mathbb{E}\Big[\int_r^t \|D_r^l X_u^k\|_H^p \, du\Big],
\end{aligned}
$$

where $\bar{M} := C_2 \sum_{i=1}^d M_i^p$. Putting the terms together and taking the supremum we get

$$
\begin{aligned}
\psi_{k+1}(t) &= \sup_{0\le r\le t} \mathbb{E}[\sup_{s\in[r,t]} \|D_r^l X_s^{k+1}\|_H^p] \le C(r,k,p) + \tilde{M} \int_r^t \mathbb{E}[\|D_r^l X_u^k\|_H^p] \, du \\
&\le C(r,k,p) + \tilde{M} \int_0^t \sup_{0\le r\le u} \mathbb{E}[\sup_{s\in[r,u]} \|D_r^l X_s^k\|_H^p] \, du \\
&= C(r,k,p) + \tilde{M} \int_0^t \psi_k(u) \, du \qquad (2.14)
\end{aligned}
$$

where $C(r,k,p) := M^p C_1^p (1 + \sup_{0\le r\le t} \mathbb{E}[\|X_r^k\|_H^p]) < \infty$ as $X^k \in \mathcal{H}^p$ and $\tilde{M} := t^p M^p (\bar{M} + M_0^p)$. Note that by the second induction hypothesis (2.9) we have $\psi_{k+1}(t) < \infty$. Hence we have shown (2.9) and (2.10). As for (2.8), note that it follows from $\psi_{k+1}(t) < \infty$ and (2.13), that

$$
\|D^l X_t^{k+1}\|_{L^p(\Omega; L^2([0,T];H))}^p \le t^{p-1} \psi_{k+1}(t) < \infty,
$$

for $l = 1, ..., d$ and hence as $X^{k+1} \in \mathbb{D}^{1,2}(H)$ as shown above we conclude from Proposition 2.11 that $X_t^{k+1} \in \mathbb{D}^{1,p}(H)$. Hence we have completed the induction and (2.8)-(2.10) hold for all $k \ge 0$. We know that $X^k \to X$ in $\mathcal{H}^p$ and thus in particular that $X_t^k \to X_t$

in $L^p(\Omega; H)$. Thus it follows from Lemma 2.5 that we have $X_t \in \mathbb{D}^{1,p}(H)$ if there is some finite $C > 0$ such that

$$\sup_k \|D^l X_t^k\|_{L^p(\Omega; L^2([0,T];H)} \leq C, \tag{2.15}$$

for $l = 1, ..., d$. As (2.14) holds for all k (as was proved with induction) we can apply Gronwall's Lemma and get

$$\psi_{k+1}(t) \leq C(r,k,p) e^{\tilde{M}t} \tag{2.16}$$

and since $\sup_k C(r,k,p) = M^p C_1^p (1 + \sup_k \sup_{0 \leq r \leq t} \mathbb{E}[\|X_r^k\|_H^p]) \leq C < \infty$ as X^k converges in $\mathcal{H}^p$, we have that $\sup_k \psi_{k+1}(t) \leq Ce^{\tilde{M}t}$ which in particular gives (2.15) and hence we have $X_t \in \mathbb{D}^{1,p}(H)$ for all $t \in [0,T]$ and $p > 2$ and hence we have $X_t \in \mathbb{D}^{1,\infty}(H)$. The Representation (2.5) follows from the application of the derivative operator D_s^l to X_t. Note that $D_s^l X_t$ is a mild solution to (2.6) but at this point there is not necessarily a continuous version. We show first (2.7). By proceeding as above we get just as in (2.14) that

$$\sup_{0 \leq r \leq T} \mathbb{E}[\sup_{s \in [r,T]} \|D_r^l X_s\|_H^p] \leq C(r,p) + \tilde{M} \int_0^T \sup_{0 \leq r \leq u} \mathbb{E}[\sup_{s \in [r,u]} \|D_r^l X_s\|_H^p] \, du,$$

with $C(r,p) := M^p C_1^p (1 + \sup_{0 \leq r \leq t} \mathbb{E}[\|X_r\|_H^p]) < \infty$ as $X \in \mathcal{H}^p$. Hence we can again apply Gronwall's Lemma to arrive at

$$\sup_{0 \leq r \leq T} \mathbb{E}[\sup_{s \in [r,T]} \|D_r^l X_s\|_H^p] \leq C(r,p) e^{\tilde{M}T} < \infty,$$

which gives (2.7). Finally, the existence of a continuous version of this mild solution follows from [26, Proposition 7.3], as it follows now from (2.7) that for any $p > 2$

$$\mathbb{E}\Big[\int_0^T \|dV_i(X_u) D_r^l X_u\|_H^p \, du\Big] \leq C_i^p T \sup_{u \in [0,T]} \mathbb{E}\Big[\|D_r^l X_u\|_H^p\Big] < \infty,$$

which completes the proof. □

In the following we will understand the map $dV : H \to L(H;H)$ as $dV : H \times H \to H$. Then dV is bounded in the first argument, linear in the second and $dV : C^\infty(H \times H; H)$. As in [1, Proposition 2.4.11] we will consider the partial derivatives of an $W \in C^\infty(H \times H; H)$ defined as (we use $\tilde{d}$ for derivatives of functions in $C^\infty(H \times H; H)$ and d for the derivatives of functions $C^\infty(H;H)$ to avoid confusion)

$$\tilde{d}_1 W(f,g)(e_1) := \tilde{d}W(f,g)(e_1, 0), \quad \tilde{d}_2 W(f,g)(e_2) := \tilde{d}W(f,g)(0, e_2)$$

and it holds true that

$$\tilde{d}W(f,g)(e_1,e_2) = \tilde{d}_1 W(f,g)(e_1) + \tilde{d}_2 W(f,g)(e_2).$$

Applied to $dV : C^\infty(H \times H; H)$ this gives

$$\tilde{d}dV(f,g)(e_1,e_2) = d^2V(f)(g,e_1) + dV(f)(e_2). \tag{2.17}$$

We will also need the following notion. We have $F = (F^1, F^2) \in \mathbb{D}^{1,\infty}(H \times H)$ if and only if $F_1 \in \mathbb{D}^{1,\infty}(H)$ and $F^2 \in \mathbb{D}^{1,\infty}(H)$. In this case it follows from the properties of the tensor-product that $DF = (DF^1, 0) + (0, DF^2)$. By iterating this arguments (and notions) we arrive at the following Lemma.

Lemma 2.13. *$V \in C_b^\infty(H;H)$. With the above notations, $d^nV : H^{n+1} \to H$ satisfies the conditions of Lemma 2.4 for $q = 1$ and accordingly for $(X_1, ..., X_{n+1}) \in \mathbb{D}^{1,\infty}(H^{n+1})$ we have $Dd^nV(X_1, ..., X_n) \in \mathbb{D}^{1,\infty}(H)$ and*

$$\begin{cases} Dd^nV(X_1,...,X_{n+1}) = \tilde{d}d^nV(X_1,...,X_{n+1})(DX_1,...,DX_{n+1}) \\ = d^{n+1}V(X_1)(DX_1,X_2,...,X_{n+1}) \\ + \sum_{i=2}^{n+1} d^nV(X_1)(X_2,...,X_{i-1},DX_i,X_{i+1},...,X_{n+1}). \end{cases} \tag{2.18}$$

Proof. It is enough to show this for $n = 1$ as the proof for $n \geq 2$ is just the same. Let

$$\sup_{f\in H} \|dV(f)\|_{L(H;H)} \leq C_1 < \infty, \quad \text{and} \quad \sup_{f\in H} \|d^2V(f)\|_{L(H\times H;H)} \leq C_2 < \infty.$$

The first condition is satisfied for $q = 0$, as

$$\|dV(f,g)\|_H = \|dV(f)g\|_H \leq C_1\|g\|_H$$

for all $(f,g) \in H \times H$ and for the second condition, using (2.17) we have that

$$\begin{aligned} \|\tilde{d}dV(f,g)\|_{L(H\times H;H)} &= \sup_{\|(e_1,e_2)\|_{H\times H}=1} \|\tilde{d}dV(f,g)(e_1,e_2)\|_H \\ &\leq \sup_{\|(e_1,e_2)\|_{H\times H}=1} \Big(\|d^2V(f)(g,e_1)\|_H + \|dV(f)(e_2)\|_H\Big) \\ &\leq \sup_{\|(e_1,e_2)\|_{H\times H}=1} \Big(C_2(\|g\|_H\|e_1\|_H) + C_1\|e_2\|_H\Big) \\ &\leq C(\|g\|_H), \end{aligned}$$

where $C := \max(C_1, C_2)$. Hence we can apply the Lemma 2.4 and the Representation (2.18) follows from (2.17). □

For the next Lemma we introduce the following notation. Let $1 \leq m \leq n$ be natural numbers, $l = \{l_1, ..., l_n\}$ a set with n elements. Denote by $\mathcal{P}(l)$ the power set of l excluding the empty set. Then we denote by $\mathcal{P}_m(l) \subset (\mathcal{P}(l))^m$ the set with elements of the form $p = \{p_1, ..., p_m\}$ with each $p_i \in \mathcal{P}(l)$, $i = 1, ..., m$, such that each $l_i \in l$, $i = 1, ..., n$, is contained in exactly one p_j, $j = 1, ..., m$. That is,

$$\mathcal{P}_m(l) = \Big\{\{p_1,...,p_m\} \in (\mathcal{P}(l))^m \Big| \forall i \in \{1,...,n\}, \exists! j \in \{1,...,m\} \text{ s.t. } l_i \in p_j\Big\}. \tag{2.19}$$

For example $\mathcal{P}_1(l) = \{p_1\}$ with $p_1 = \{l_1, ..., l_n\}$, $\mathcal{P}_n(l) = \{p_1, ..., p_n\}$ with $p_i = l_i$, $i = 1, ..., n$ and $\mathcal{P}_{n-1}(l) = \Big\{\big\{\{l_1, l_2\}, l_3, ..., l_n\big\}, ..., \big\{l_1, l_2, ..., \{l_{n-1}, l_n\}\big\}\Big\}$.

Lemma 2.14. *Let X be a predictable process with $X_t \in \mathbb{D}^{n,\infty}(H)$ for all $t \in [0,T]$. If $V \in C_b^\infty(H)$, then we have $V(X_t) \in \mathbb{D}^{n,\infty}(H)$ for all $t \in [0,T]$ and for any $(l_1, ..., l_n) \in \{1, ..., d\}^n$ and $(s_1, ..., s_n) \in [0,T]^n$ we have the representation*

$$D^{l_1,...,l_n}_{s_1,...,s_n} V(X_t) = \alpha^{l_1,...,l_n}_{s_1,...,s_n}(t) + dV(X_t)D^{l_1,...,l_n}_{s_1,...,s_n} X_t \tag{2.20}$$

for all $t \in [s_1 \vee ... \vee s_n, T]$ and $D^{l_1,...,l_n}_{s_1,...,s_n} V(X_t) = 0$ if $t \in [0, s_1 \vee ... \vee s_n)$, where $\alpha^{l_1,...,l_n}_{s_1,...,s_n}(t)$ is given by

$$\begin{cases} \alpha^{l_1,...,l_n}_{s_1,...,s_n}(t) & := d^nV(X_t)(D^{l_1}_{s_1}X_t, ..., D^{l_n}_{s_n}X_t) \\ + \sum_{(p,q)\in\mathcal{P}_{n-1}(LS)} & d^{n-1}V(X_t)(D^{p_1}_{q_1}X_t, ..., D^{p_{n-1}}_{q_{n-1}}X_t) \\ & \vdots \\ + \sum_{(p,q)\in\mathcal{P}_2(LS)} & d^2V(X_t)(D^{p_1}_{q_1}X_t, D^{p_2}_{q_2}X_t), \end{cases} \tag{2.21}$$

where it is understood that $\alpha_{s_1}^{l_1}(t) = 0$*. Furthermore, for each* $p \geq 1$ *there is a finite constant* C_p *such that*

$$\mathbb{E}[\sup_{\bar{s}^n \leq t \leq T} \|\alpha_{s_1,...,s_n}^{l_1,...,l_n}(t)\|_H^p] \leq C_p \sum_{j=1}^{n} \sum_{(p,q)\in\mathcal{P}_j(LS)} \sum_{i=1}^{j} \mathbb{E}[\sup_{\bar{s}^n \leq t \leq T} \|D_{q_i}^{p_i} X_t\|_H^p], \tag{2.22}$$

where $\bar{s}^n := s_1 \vee ... \vee s_n$ *and* $LS = \{(l_1, s_1), ..., (l_n, s_n)\}$*. Also,* $\alpha_{s_1,...,s_n}^{l_1,...,l_n}(t) \in \mathbb{D}^{1,p}(H)$ *and*

$$D_{s_{n+1}}^{l_{n+1}} \alpha_{s_1,...,s_n}^{l_1,...,l_n}(t) = \alpha_{s_1,...,s_{n+1}}^{l_1,...,l_{n+1}}(t) - d^2V(X_t)(D_{s_{n+1}}^{l_{n+1}} X_t, D_{s_1,...,s_n}^{l_1,...,l_n} X_t). \tag{2.23}$$

Proof. By using Lemma 2.4 and Lemma 2.13 it can be easily shown by using induction that, with the notation introduced above, we have

$$\begin{aligned} D_{s_1,...,s_n}^{l_1,...,l_n} V(X_t) &= d^n V(X_t)(D_{s_1}^{l_1} X_t, ..., D_{s_n}^{l_n} X_t) \\ + \sum_{(p,q)\in\mathcal{P}_{n-1}(LS)} & d^{n-1}V(X_t)(D_{q_1}^{p_1} X_t, ..., D_{q_{n-1}}^{p_{n-1}} X_t) \\ & \vdots \\ + \sum_{(p,q)\in\mathcal{P}_2(LS)} & d^2V(X_t)(D_{q_1}^{p_1} X_t, D_{q_2}^{p_2} X_t) \\ &+ dV(X) D_{s_1,...,s_n}^{l_1,...,l_n} X_t, \end{aligned}$$

whenever $t \in [s_1 \vee ... \vee s_n, T]$ and $D_{s_1,...,s_n}^{l_1,...,l_n} V(X_t) = 0$ if $t \in [0, s_1 \vee ... \vee s_n)$. From this all of the claims directly follow. □

Lemma 2.15. *Let* X *be a predictable process with* $X_t \in \mathbb{D}^{n,\infty}(H)$ *for all* $t \in [0,T]$ *and* $V_1, ..., V_d \in C_b^\infty(H;H)$*. Then for any* $(l_1, ..., l_n) \in \{1, ..., d\}^n$ *and* $s^n := (s_1, ..., s_n) \in [0,T]^n$*,* $\bar{s}^n := \max(s_1, ..., s_n)$ *and* $t \in [0,T]$

$$\beta(\bar{s}^n) := S_{t-\bar{s}^n} \tilde{\beta}(\bar{s}^n) := S_{t-\bar{s}^n} \sum_{i=1}^{n} D_{s_1,...,s_{i-1},s_{i+1},...,s_n}^{l_1,...,l_{i-1},l_{i+1},...,l_n} V_{l_i}(X_{s_i}) \mathbf{1}_{\{s_i = \bar{s}^n\}}, \tag{2.24}$$

is a well-defined, $\bar{s}^n$*-measurable random variable in* $\mathbb{D}^{1,\infty}(H)$ *and*

$$\begin{cases} D_{n+1}^{l_{n+1}} \beta(\bar{s}^n) &= S_{t-\bar{s}^{n+1}} D_{n+1}^{l_{n+1}} \tilde{\beta}(\bar{s}^n) \\ &= S_{t-\bar{s}^{n+1}} \sum_{i=1}^{n} D_{s_1,...,s_{i-1},s_{i+1},...,s_{n+1}}^{l_1,...,l_{i-1},l_{i+1},...,l_{n+1}} V_{l_i}(X_{s_i}) \mathbf{1}_{\{s_i = \bar{s}^{n+1}\}}. \end{cases} \tag{2.25}$$

Proof. This follows from Lemma 2.4. □

Lemma 2.16. *If* $\sum_{i=1}^d \int_0^t S_{t-s} V_i(X_s)\, dW_s^i \in \mathbb{D}^{n,\infty}(H)$*, then*

$$D_{s_1,...,s_n}^{l_1,...,l_n} \sum_{i=1}^{d} \int_0^t S_{t-s} V_i(X_s)\, dW_s^i = \beta(\bar{s}^n) + \int_{\bar{s}^n}^t D_{s_1,...,s_n}^{l_1,...,l_n} S_{t-s} V_i(X_s)\, dW_s^i. \tag{2.26}$$

Theorem 2.17. *In the setting of Lemma 2.12 and with the notation from Lemma 2.14 and 2.15, for every* $t \in [0,T]$*, the mild solution* X_t *of* (I.3.2) *is in* $\mathbb{D}^\infty(H)$ *and for any* $n \geq 1$*,* $l^n := (l_1, ..., l_n) \in \{1, ..., d\}^n$ *and* $s^n := (s_1, ..., s_n) \in [0,T]^n$*, the process* $Y_t := D_{s^n}^{l^n} X_t$ *is a continuous mild solution of the equation in* H *for* $t \geq \bar{s}^n$ *given by*

$$\begin{cases} dY_t &= (AY_t + \alpha_{s^n}^{l^n}(t) + dV(X_t)Y_t)\, dt + \sum_{i=1}^d (\alpha_{s^n}^{i,l^n}(t) + dV_i(X_t)Y_t)\, dW_t^i \\ Y_{\bar{s}^n} &= \tilde{\beta}(\bar{s}^n) \end{cases} \tag{2.27}$$

and $Y_t = 0$ *for* $t < \bar{s}^n$*.*

Proof. We will show by induction that for all $n \geq 1$, we have

I1 $X_t \in \mathbb{D}^{n,p}(H)$ for all $t \in [0,T]$,

I2 $D^{l^n}_{s^n} X_t$ is for $t \geq \bar{s}^n$ a continuous mild solution of (2.27), and

I3 $\sup_{s^n \in [0,T]^n} \mathbb{E}\Big[\sup_{\bar{s}^n \leq t \leq T} \|D^{l^n}_{s^n} X_t\|^p_H\Big] < \infty$.

We know from Lemma 2.12 that all three claims are satisfied for $n = 1$. So assume it holds for some $n > 1$. Then $Y_t := D^{l^n}_{s^n} X_t$ has the representation

$$\begin{cases} Y_t & = S_{t-\bar{s}^n}\tilde{\beta}(\bar{s}^n) + \int_{\bar{s}^n}^t S_{t-u}\Big(\alpha^{l^n}_{s^n}(u) + dV(X_u)Y_u\Big)\,du \\ & \quad + \sum_{i=1}^d \int_{\bar{s}^n}^t S_{t-u}\Big(\alpha^{l^n}_{i,s^n}(u) + dV_i(X_u)Y_u\Big)\,dW^i_u. \end{cases} \tag{2.28}$$

We show first that $Y_t \in \mathbb{D}^{1,\infty}(H)$ by a very similar method as in Lemma 2.12, however, we can not apply Theorem I.3.2 directly to (2.27) as the condition **SI2** is not evident. Indeed, let $U \in \{V, V_1, ..., V_d\}$, $C_U := \sup_{x\in H} \|dU(x)\|_{L(H;H)}$ and $\gamma \in \{\alpha, \alpha_1, ...\alpha_d\}$ then we have

$$\begin{aligned} &\|\gamma^{l^n}_{s^n}(t,\omega) + dU(X_t(\omega))(x)\|^2_H \leq \|\gamma^{l^n}_{s^n}(t,\omega)\|^2_H + C_U^2\|x\|^2_H \\ \leq\;& C_2 \sum_{j=1}^n \sum_{(p,q)\in\mathcal{P}_j(LS)} \sum_{i=1}^j \|D^{p_i}_{q_i} X_t(\omega)\|^2_H + C_U^2\|x\|^2_H, \end{aligned}$$

so unless $\|D^{p_i}_{q_i} X_t(\omega)\|^2_H$ is uniformly bounded in (ω, t) it is not evident that this condition holds. We proceed similarly to [63, Lemma 2.2.2]. Let

$$\alpha(\bar{s}^n, t) := S_{t-\bar{s}^n}\tilde{\beta}(\bar{s}^n) + \int_{\bar{s}^n}^t S_{t-u}\alpha^{l^n}_{s^n}(u)\,du + \sum_{i=1}^d \int_{\bar{s}^n}^t S_{t-u}\alpha^{l^n}_{i,s^n}(u)\,dW^i_u \tag{2.29}$$

and consider the sequence for $k \geq 0$ defined by

$$\begin{cases} Y^0_t & := \alpha(\bar{s}^n, t) \\ Y^{k+1}_t & := \alpha(\bar{s}^n, t) + \int_{\bar{s}^n}^t S_{t-u} dV(X_u) Y^k_u\,du + \sum_{i=1}^d \int_{\bar{s}^n}^t S_{t-u} dV_i(X_u)Y^k_u\,dW^i_u. \end{cases} \tag{2.30}$$

As in the proof of Lemma 2.12, we show with induction that Y^k_t satisfies the corresponding conditions given in (2.8)-(2.10) and additionally the condition

$$\sup_{0\leq \bar{s}^n \leq T} \mathbb{E}[\sup_{\bar{s}^n \leq t \leq T} \|Y^k_t\|^p_H] < \infty, \tag{2.31}$$

however, we skip some details as the proof is very similar. Let $k = 0$. We start with showing condition (2.31). By proceeding as above we see that

$$\begin{aligned} \mathbb{E}[\sup_{\bar{s}^n \leq t \leq T} \|Y^0_t\|^p_H] &\leq M(T,p)\Big(\mathbb{E}[\|\tilde{\beta}(\bar{s}^n)\|^p_H] + \int_{\bar{s}^n}^T \mathbb{E}[\|\alpha^{l^n}_{s^n}(u)\|^p_H]\,du \\ &+ \sum_{i=1}^d \int_{\bar{s}^n}^T \mathbb{E}[\|\alpha^{l^n}_{i,s^n}(u)\|^p_H]\,du\Big) =: M(T,p)(I_1 + I_2 + I_3). \end{aligned} \tag{2.32}$$

For the first term, using (2.24) and (2.20) we get

$$\begin{aligned} I_1 &\leq m_p \sum_{i=1}^{n} \mathbb{E}[\|D^{l_1,\dots,l_{i-1},l_{i+1},\dots,l_n}_{s_1,\dots,s_{i-1},s_{i+1},\dots,s_n} V_{l_i}(X_{s_i})\|_H^p] \mathbf{1}_{\{s_i=\bar{s}^n\}} \\ &\leq \tilde{m}_p \sum_{i=1}^{n} \mathbb{E}[\|\alpha^{l_1,\dots,l_{i-1},l_{i+1},\dots,l_n}_{s_1,\dots,s_{i-1},s_{i+1},\dots,s_n}(t) + dV(X_t) D^{l_1,\dots,l_{i-1},l_{i+1},\dots,l_n}_{s_1,\dots,s_{i-1},s_{i+1},\dots,s_n} X_t\|_H^p] \\ &\leq c_p \tilde{m}_p \Big(\sum_{i=1}^{n} \mathbb{E}[\|\alpha^{l_1,\dots,l_{i-1},l_{i+1},\dots,l_n}_{s_1,\dots,s_{i-1},s_{i+1},\dots,s_n}(t)\|_H^p] \\ &+ C_p \mathbb{E}[\|D^{l_1,\dots,l_{i-1},l_{i+1},\dots,l_n}_{s_1,\dots,s_{i-1},s_{i+1},\dots,s_n} X_t\|_H^p] \Big) < \infty \end{aligned}$$

uniformly in $s^n \in [0,T]^n$ as both terms consist of Malliavin derivatives up to order $n-2$ (cf. (2.21)) and $n-1$ respectively, this follows from the induction hypothesis **I3**. The same applies to the terms I_2 and I_3 by using again (2.21) and (2.22). Hence we have shown (2.31) for $k=0$. Next, for (2.8), we see that from (2.24) and the induction hypothesis and Lemma 2.4 it follows that $S_{t-\bar{s}^n}\tilde{\beta}(\bar{s}^n) \in \mathbb{D}^{1,p}(H)$. Let $\gamma \in \{\alpha, \alpha_1, \dots.\alpha_d\}$. We claim that $t \mapsto \gamma^{l^n}_{s^n}(t) \in \mathbb{L}^{1,2}(H)$. Indeed, the equivalent conditions of Lemma 2.7 directly follow from (2.21) and (2.22) and the induction hypothesis. Hence Propositions 2.8 and 2.9 apply and $Y_t^0 \in \mathbb{D}^{1,2}(H)$ for almost all $t \in [0,T]$ which gives (2.8) for $k=0$ and it holds that

$$\begin{cases} D^{l_{n+1}}_{s_{n+1}} Y_t^0 & = D^{l_{n+1}}_{s_{n+1}} S_{t-\bar{s}^n}\tilde{\beta}(\bar{s}^n) + \int_{\bar{s}^{n+1}}^t S_{t-u} D^{l_{n+1}}_{s_{n+1}} \alpha^{l^n}_{s^n}(u)\,du \\ & +S_{t-s_{n+1}} \alpha^{l_n}_{l_{n+1},s^n}(s_{n+1}) + \sum_{i=1}^d \int_{\bar{s}^{n+1}}^t S_{t-u} D^{l_{n+1}}_{s_{n+1}} \alpha^{l^n}_{i,s^n}(u)\,dW_u^i, \end{cases} \tag{2.33}$$

with $D^{l_{n+1}}_{s_{n+1}} S_{t-\bar{s}^n}\tilde{\beta}(\bar{s}^n)$ and $D^{l_{n+1}}_{s_{n+1}}\alpha^{l^n}_{i,s^n}$ given in (2.25) and (2.23), respectively. For (2.9) we need to show that

$$\psi_0(t) := \sup_{0\leq r\leq t} \mathbb{E}[\sup_{s\in[r,t]} \|D_r^l Y_s^0\|_H^p] < \infty. \tag{2.34}$$

First note that, using (2.25) and (2.20) we get

$$\begin{aligned} & D^{l_{n+1}}_{n+1}\tilde{\beta}(\bar{s}^n) \\ = & \sum_{i=1}^{n} \Big(\alpha^{l_1,\dots,l_{i-1},l_{i+1},\dots,l_{n+1}}_{s_1,\dots,s_{i-1},s_{i+1},\dots,s_{n+1}}(s_i) + dV_{l_i}(X_{s_i}) D^{l_1,\dots,l_{i-1},l_{i+1},\dots,l_{n+1}}_{s_1,\dots,s_{i-1},s_{i+1},\dots,s_{n+1}} X_{s_i} \Big) \mathbf{1}_{\{s_i=\bar{s}^{n+1}\}}. \end{aligned}$$

and accordingly from the induction hypothesis **I3**, (2.22) and the boundedness of $dV_{l_i}(X_{s_i})$ it follows that $\mathbb{E}[\|D^{l_{n+1}}_{s_{n+1}} S_{t-\bar{s}^n}\tilde{\beta}(\bar{s}^n)\|_H^p] < \infty$ and again from (2.22) we get $\mathbb{E}[\|S_{t-s_{n+1}}\alpha^{l_n}_{l_{n+1},s^n}(s_{n+1})\|_H^p] < \infty$. The remaining terms in (2.33) can be handled just as in Lemma 2.12 and thus (2.34) follows. Thus for $k=0$, it remains to show (2.10), that is, that

$$\psi_1(t) := \sup_{0\leq r\leq t} \mathbb{E}[\sup_{s\in[r,t]} \|D_r^l Y_s^1\|_H^p] \leq c_1 + c_2 \int_0^t \psi_0(s)\,ds, \tag{2.35}$$

holds true. First we have to show that $Y_t^1 \in \mathbb{D}^{1,\infty}(H)$ for almost all $t \in [0,T]$, which follows if, for $U \in \{V, V_1, \dots, V_d\}$, we have $u \mapsto (dU(X_u)Y_u^0) \in \mathbb{L}^{1,2}(H)$, but this evident from Lemma 2.4, Lemma 2.7 and induction hypothesis. Then from Lemma 2.14, 2.15 and Propositions 2.8 and 2.9 we have

$$\begin{aligned} D^{l_{n+1}}_{s_{n+1}} Y^1_t &= D^{l_{n+1}}_{s_{n+1}} Y^0_t + \int_{\bar{s}^{n+1}}^t S_{t-u} D^{l_{n+1}}_{s_{n+1}} dV(X_u) Y^0_u \, du \\ &+ S_{t-s_{n+1}} dV_{l_{n+1}}(X_{s_{n+1}}) Y^0_{s_{n+1}} + \sum_{i=1}^d \int_{\bar{s}^{n+1}}^t S_{t-u} D^{l_{n+1}}_{s_{n+1}} dV_i(X_u) Y^0_u \, dW^i_u. \end{aligned}$$

Accordingly (2.35) follows from (2.31) for $k = 0$, the induction hypothesis and from (cf. Lemma 2.13)

$$D^{l_{n+1}}_{s_{n+1}} dU(X_u) Y^0_u = d^2 U(X_u)(D^{l_{n+1}}_{s_{n+1}} X_u, Y^0_u) + dU(X_u) D^{l_{n+1}}_{s_{n+1}} Y^0_u, \tag{2.36}$$

where again $U \in \{V, V_1, ..., V_d\}$. Accordingly we have shown the induction hypothesis to be true for $k = 0$. Assuming (2.8)-(2.10) and (2.31) hold for some $k > 0$ it can be shown in literally the same way that the claim holds for $k+1$. Hence the claim holds for all $k \geq 0$. Next we use Lemma 2.5 to conclude that $Y_t \in \mathbb{D}^{1,p}(H)$ for all $t \in [0,T]$. We show first that Y^k converges in $\mathcal{H}^p$. Note that we know from (2.31) that $Y^k \in \mathcal{H}^p$ for all $k \geq 0$. We have by standard arguments that for $k \geq 1$

$$\|Y^{k+1} - Y^k\|_{\mathcal{H}^p} \leq c(T,p) \|Y^k - Y^{k-1}\|_{\mathcal{H}^p}, \tag{2.37}$$

which by the usual procedure of choosing T small enough leads to $c(T,p) < 1$ and accordingly to an unique fixpoint which evidently is Y and showing also that $Y \in \mathcal{H}^p$. The final step in applying Lemma 2.5 is to show that

$$\sup_{k \geq 0} \mathbb{E}\Big[\Big(\int_0^t \|D^l_r Y^k_t\|^2_H \, dr\Big)^{p/2}\Big] < \infty. \tag{2.38}$$

We have

$$\mathbb{E}\Big[\Big(\int_0^t \|D^l_r Y^k_t\|^2_H \, dr\Big)^{p/2}\Big] \leq t^{p-2} \mathbb{E}\Big[\int_0^t \|D^l_r Y^k_t\|^p_H \, dr\Big] \leq t^{p-1} \sup_{r \in [0,t]} \mathbb{E}\Big[\|D^l_r Y^k_t\|^p\Big]$$

and

$$\sup_{r \in [0,t]} \mathbb{E}\Big[\|D^l_r Y^k_t\|^p\Big] \leq \sup_{r \in [0,t]} \mathbb{E}\Big[\sup_{t \in [r,T]} \|D^l_r Y^k_t\|^p\Big] = \psi_k(t).$$

Hence from the shown induction hypothesis (2.10) we can apply Gronwalls Lemma and the claim follows, that is, from Lemma 2.5 we deduce that $Y_t \in \mathbb{D}^{1,p}(H)$ and accordingly $X_t \in \mathbb{D}^{n+1,p}(H)$ for all $t \in [0,T]$ proving the first induction hypothesis **I1**. Now let $Z_t := D^{l_{n+1}}_{s_{n+1}} Y_t = D^{l^{n+1}}_{s^{n+1}} X_t$. For the second hypothesis **I2** we need to show that Z_t satisfies (2.27) for $n+1$ and has a continuous version. By noting that

$$D^{l_{n+1}}_{s_{n+1}} (\alpha^{l^n}_{s^n}(u) + dV(X_u) Y_u) = \alpha^{l^{n+1}}_{s^{n+1}}(u) + dV(X_u) Z_u, \tag{2.39}$$

which follows from (2.20) in Lemma 2.14 and that

$$
\begin{aligned}
& D^{l_{n+1}}_{s_{n+1}}\Big(\beta(\bar{s}^n)+\sum_{i=1}^d\int_{\bar{s}^n}^t S_{t-u}\big(\alpha^{l^n}_{i,s^n}(u)+dV_i(X_u)Y_u\big)\,dW^i_s\Big)\\
= \;& D^{l_{n+1}}_{s_{n+1}}\Big(\beta(\bar{s}^n)+\sum_{i=1}^d\int_{\bar{s}^n}^t D^{l_1,\dots,l_n}_{s_1,\dots,s_n}S_{t-s}V_i(X_s)\,dW^i_s\Big)\\
= \;& \beta(\bar{s}^{n+1})+\sum_{i=1}^d\int_{\bar{s}^{n+1}}^t D^{l_1,\dots,l_{n+1}}_{s_1,\dots,s_{n+1}}S_{t-s}V_i(X_s)\,dW^i_s\\
= \;& \beta(\bar{s}^{n+1})+\sum_{i=1}^d\int_{\bar{s}^{n+1}}^t \big(\alpha^{l^{n+1}}_{i,s^{n+1}}(u)+dV_i(X_u)Z_u\big)\,dW^i_s,
\end{aligned}
$$

where the first and third equality follows from (2.20) and the second from Lemma 2.16. Thus Z_t is a mild solution of (2.27) for $n+1$. As in the proof of Lemma 2.12 we show first condition **I3** before showing that Z has a continuous version. We have to show that

$$
\sup_{s^{n+1}\in[0,T]^{n+1}}\mathbb{E}[\sup_{\bar{s}^{n+1}\le t\le T}\|D^{l_{n+1}}_{s_{n+1}}Y_t\|^p_H]<\infty. \tag{2.40}
$$

By proceeding as above, we get

$$
\begin{aligned}
& \mathbb{E}\Big[\sup_{\bar{s}^{n+1}\le t\le T}\|D^{l_{n+1}}_{s_{n+1}}Y_t\|^p_H\Big]\le M_1(T,p)\mathbb{E}\Big[\|\tilde{\beta}(\bar{s}^{n+1})\|^p_H\Big]\\
+ \;& M_2(T,p)\Big(\mathbb{E}\Big[\int_{\bar{s}^{n+1}}^T\|\alpha^{l^{n+1}}_{s^{n+1}}(u)\|^p_H du\Big]+\sum_{i=1}^d\mathbb{E}\Big[\int_{\bar{s}^{n+1}}^T\|\alpha^{l^{n+1}}_{i,s^{n+1}}(u)\|^p_H du\Big]\Big)\\
+ \;& M_3(T,p)\mathbb{E}\Big[\int_{\bar{s}^{n+1}}^T\|D^{l_{n+1}}_{s_{n+1}}Y_u\|^p_H du\Big]=:I_1+I_2+I_3,
\end{aligned}
$$

where $M_1(T,p), M_2(T,p), M_3(T,p)$ are constants depending on T and p. The terms I_1 and I_2 can be shown to be finite, uniformly in $s^n\in[0,T]^n$, in the same way as in (2.32). Thus we have

$$
\begin{aligned}
& \sup_{s^{n+1}\in[0,T]^{n+1}}\mathbb{E}\Big[\sup_{\bar{s}^{n+1}\le t\le T}\|D^{l_{n+1}}_{s_{n+1}}Y_t\|^p_H\Big]\\
& \le C(T,p)+M(T,p)\int_0^T\sup_{s^{n+1}\in[0,t]^{n+1}}\mathbb{E}\Big[\sup_{\bar{s}^{n+1}\le u\le t}\|D^{l_{n+1}}_{s_{n+1}}Y_t\|^p_H\Big]\,dt
\end{aligned}
$$

and **I3** follows from Gronwall's Lemma. Finally we show that the mild solution Z_t of (2.27) for $n+1$ has a continuous version. This follows from [26, Proposition] upon showing that for some $p>2$

$$
\mathbb{E}\Big[\int_0^T\Big\|\alpha^{i,l^{n+1}}_{s^{n+1}}(s)+dV_i(X_s)Z_s\Big\|^p_H\,ds\Big]<\infty, \tag{2.41}
$$

but this follows now immediately from (2.22) and the just shown induction hypothesis **I3** by noting that $Z_s=D^{l+1}_{n+1}Y_s$ and thus completing the proof. □

3 Asymptotic Expansions

The next result is based on Malliavin's *Integration by Parts Formula.* For $F \in \mathbb{D}^{1,2}(\mathbb{R}^N)$ the *Malliavin Covariance Matrix* $\gamma(F)$ is defined as

$$\gamma(F) := (\langle DF^i, DF^j \rangle_{L^2([0,T];\mathbb{R}^d)})_{i,j=1,\dots,N}. \tag{3.1}$$

Let G be another $\mathbb{R}^N$-valued random variable. If we now assume that $\gamma(F)$ is invertible and $(DF)^T\gamma^{-1}(F)G \in dom(\delta)$, then for any continuously differentiable function $f : \mathbb{R}^N \to \mathbb{R}$ we have (see [63, Proposition 6.2.1])

$$\mathbb{E}[f'(F)G] = \mathbb{E}[f(F)\delta((DF)^T\gamma^{-1}(F)G)] = \mathbb{E}[f(F)\pi], \tag{3.2}$$

where π is called the *Malliavin Weight* and is given by the Skohorod-Integral

$$\pi := \delta((DF)^T\gamma^{-1}(F)G). \tag{3.3}$$

By iterating this procedure and generalizing it to functions f that are only bounded and measurable in [68, Definition 2.2] the *weak Taylor approximation* is introduced and in [68, Theorem 2.3] sufficient conditions for the convergence are given. We summarize both in the following Theorem.

Theorem 3.1. *Let $\epsilon \mapsto F_\epsilon$ be a random $\mathbb{R}^N$-valued curve and $U \subset \mathbb{R}$ be a neighborhood around 0 such that both conditions are satisfied:*

WTA 1 *the curve $\epsilon \mapsto F_\epsilon$ is smooth from U into $\mathbb{D}^\infty(\mathbb{R}^N)$ and*

WTA 2 *the Malliavin Covariance Matrix $\gamma(F_\epsilon)$ is invertible with $\det(\gamma(F_\epsilon))^{-1} \in L^p(\Omega)$ for every $p \geq 1$ for every $\epsilon \in U$.*

Then for each bounded, measurable $f : \mathbb{R}^N \to \mathbb{R}$ and $n \geq 1$, there is a weak Taylor approximation *$W_\epsilon^n(f, F_\epsilon)$ of order n, that is,*

$$W_\epsilon^n(f, F_\epsilon) := \sum_{i=0}^{n} \frac{\epsilon^i}{i!}\mathbb{E}[f(F_0)\pi_i],$$

such that

$$|\mathbb{E}[f(F_\epsilon) - W_\epsilon^n(f, F_\epsilon)]| = o(\epsilon^n), \quad \textit{as } \epsilon \to 0.$$

Proof. This is [68, Theorem 2.3]. □

Recalling the discussion at the beginning of this Chapter, we see that in order to apply this Theorem for $F_\epsilon = l \circ X_T^\epsilon$, where $l \in (H, \mathbb{R}^N)$ and X_T^ϵ denotes the mild solution of (1.1), we need to check **WTA 1** and **WTA 2**. From Theorem 2.17 and Lemma 2.4 we already know sufficient conditions on the vector fields of (1.1) such that $l \circ X_T^\epsilon \in \mathbb{D}^\infty(\mathbb{R}^N)$. The condition that is most hard to check is the second part of **WTA 2**, that is, the p-integrability of $\gamma^{-1}(l \circ X_T^\epsilon)$ whereas for the invertibility a generalized Hörmander condition can be used that is given in [4] which in many applications is easy to check. We assume the following special situation and conjecture the general one below.

Theorem 3.2. *If both conditions*

SWTA 1 *the coefficients* $V, V_1, ..., V_d$ *in* (1.1) *are in* $C_b^\infty(U \times H; H)$ *and*

SWTA 2 *there are finite dimensional submanifolds* M^ϵ *being left invariant by the solution of* (1.1) *starting at* X_0 *for each* $\epsilon \in U$*, whose tangent spaces are generated by Lie-bracketing from the given vector fields of* (1.1) *for* $\epsilon \neq 0$*. At* $\epsilon = 0$ *only a lower dimensional space is generated locally around* X_0*. Furthermore assume that* l *maps the lower dimensional space of directions at* X_0*, a finite dimensional subspace of* H*, onto* $\mathbb{R}^N$*.*

hold true. Then for every $n \geq 1$*, the solutions* $l \circ X_T^\epsilon$ *admits a weak Taylor expansion up to order* n*.*

Proof. First note that under the conditions of **SWTA 1** we immediately get **WTA 1** just as in [26, Theorem 9.4]. The proof that **SWTA 2** together with **SWTA 1** gives also **WTA 2** in Theorem 3.1 follows the lines of the classical proof of p-integrability of the Maliavin covariance matrix' inverse: first notice that for any $\epsilon \neq 0$ we are locally in a situation where the classical proof can be applied on the finite dimensional submanifold M^ϵ, with respect to which the process is stochastically invariant, and on which the law is locally absolutely continuous with respect Lebesgue measure, by the generation property. Of course also the projection by l satisfies the p-integrability of the inverse. For $\epsilon = 0$ on the other hand we have a finite dimensional realization and again by the generation property for the factor process, a smooth density exists. Hence **WTA 2** is satisfied. This can be compared to a finite dimensional situation, where, of course, the p-integrability of the inverse of the covariance matrix is an open condition, i.e., if the Hörmander condition holds for vector fields $V, V_1, \ldots, V_d$ at a point in state space, then it also holds for small perturbations of those vector fields at this point, see, e.g., the Norris lemma (see [63, Lemma 2.3.3]). The twist here is somehow that for $\epsilon = 0$ the generated dimension can be considerably lower if we only look at certain projections by l. □

Conjecture 3.3. *The statement of the Theorem 3.2 remains true, if instead of condition **SWTA 2**, we have that the inverse Malliavin covariance matrix is invertible for all* $\epsilon \in U$ *and is* p*-integrable for all* $p \geq 1$ *for* $\epsilon = 0$*, given some regularity conditions.*

The proof of this conjecture is work in progress and will be similar to the proof of Theorem 3.2.

Chapter V

Consistent Recalibration Models

1 Introduction

So far we looked at generalizations of two popular models for forward variances given by the spot-variance realizations and the Bergomi model and found necessary and sufficient conditions for the existence of generic finite dimensional realizations. The findings extend those in [15] into two directions. First, as the term *generic* suggests, the resulting finite dimensional realizations can take initial curves from an open subset in an infinite dimensional Hilbert space, whereas these are restricted in [15] to some finite dimensional submanifold. Second, we included a stochastic volatility process into both forward variance models and found conditions such that the joint system admit generic finite dimensional realizations. Nevertheless, in both cases the resulting forward variances evolve (after choosing the initial curve from an open set) in a fixed finite dimensional submanifold with boundary and hence the set of curves that can be attained by such models is limited.

In some recent research it is suggested that for some applications this is too restrictive. One set of examples arise from noting that the short-end of the forward variance process determines the spot-variance process and hence determines the volatility surface. Here it is suggested in a series of papers from Jim Gatheral and coworkers, including [6], to formally extend the Bergomi model such that the resulting model admits no finite dimensional realization and hence is inherently infinite dimensional from the point of view of Markov processes (see Section III.6 for a discussion of this model). This extension is motivated by a superior reflection of some (empirically observed) stylized facts. A second example includes the so-called *consistent recalibration models* introduced for forward interest rates in discrete time in [66] and reconsidered in continuous time in [50]. This models build on the observed necessity to recalibrate certain parameters of an affine realization in order to maintain a decent fit to the observed data, which implies that the model specifications are too narrow. They suggest to extend the model by replacing (some of) the model parameters by a finite dimensional diffusion processes. By doing so, consistent recalibration becomes possible and hence justifying the name. A defining characteristic of such models is the possibility to reach any curve within an open subset of an infinite dimensional Hilbert space of curves. Or, loosely speaking, to not admit a finite dimensional realization. The common feature of both examples is that certain models that admit generic finite dimensional realizations (exponentially affine in the first case and affine in the second) are extended in a way such that the model becomes inherently infinite dimensional. However, in both cases (to some extent) the models can utilize the corresponding finite dimensional models. In the first example by using an finite dimensional representation with respect to a fractional Brownian motion and in the second case by suitably discretizing the model.

In the following we will extend the models found in the previous chapter in the spirit of consistent recalibration models and hence adopt the name. In [50] it is shown that the models can be approximated for small times by concatenations of models that admit (generic) affine realizations. This allows for consistent recalibration on a discrete grid of recalibration times such that in between this times the process can be assumed to admit an affine realization. This is very well-suited for the simulation of curves and for calibration (against time-series of curves) as the affine structure can be used. However, this discretization can not be used to utilize possibly existing closed form solutions of the finite dimensional representation for certain derivatives. This will be possible with the weak Taylor expansion for SPDEs introduced in the last chapter.

We will see that consistent recalibration models are tailor-made for this expansions in that we will show that from the perspective of the parameterized system (IV.1.1) this models can be chosen such that the conditions of Conjecture IV.3.3 are satisfied. In the next chapter we will see that the weak expansion will consist of polynomials of partial derivatives (i.e. so-called *Greeks*) of the price of the unperturbed state (i.e. $\epsilon = 0$), which in the case of consistent recalibration models are expected values of functions of the model that admits generic finite dimensional realizations.

Before introducing the type of models that we will consider, we give in the following a short (formal) summary of the ideas and models used in [50] from the perspective of forward variance models. In general they consider the case where the short-rate is given by an affine function of an affine processes but consider as prime examples (i.e. the Hull-White extended Vasiček and CIR model) short-rates which are itself given by real-valued affine processes. Thus we restrict this summary for notational convenience to this case. Thus, we assume that the spot-variance process ξ is given by a real-valued (time-inhomogeneous) affine process

$$d\xi_t = (\theta(t) + b_y \xi_t)\, dt + \sqrt{a_y + \alpha_y \xi_t}\, d\beta_t, \quad \xi_0 = \xi,$$

where θ is understood as the *Hull-White extension* (cf. Section II.3.1) and $-b_y, a_y, \alpha_y$ are non-negative real-numbers depending on the parameter $y \in W$ with W being some subset of $\mathbb{R}^m$. For a sufficiently regular curve u we can choose (cf. (II.3.6))

$$\theta(t) := \frac{d}{dt} u(t) - b_y u(t)$$

and the forward variance process then satisfies

$$du_t = \frac{d}{dx} u_t\, dt + \phi(u_t, y) e^{b_y \cdot}\, d\beta_t, \quad u_0 = u, \tag{1.1}$$

where $\phi(u, y) := \sqrt{a_y + \alpha_y u(0)}$. As the solution can be represented by (cf. (II.3.7))

$$u_t = S_t u + (\xi_t - u(t)) e^{b_y \cdot}$$

we see that for each fixed parameter value $y \in W$, the forward variance curve u_t can only attain values in

$$\{S_t u + \langle e^{b_y \cdot} \rangle \mid t \in \mathbb{R}_+\}. \tag{1.2}$$

In particular, as is by now well-known, besides the initial curve u, this set of attainable curve depends only on the *mean-reversion speed* b_y and is independent of the diffusion parameters a_y and α_y. Thus, if we want the forward variance process to be able to reach any curve within some open subset with positive probability we see that, at least formally, a necessary condition is that the parameter y has to be chosen as a state process such that the mean-reversion speed b_y can attain infinitely many values. To simplify notation we assume that $a_y = a$ and $\alpha_y = \alpha$ are constants and choose $m = 1$ and let $b_y = y$. Then, if we replace the parameter y by a stochastic process Y_t it is evident by construction (we will also rigorously show this later on, compare this also to [50, Lemma 4.5]) that in general the joint process

$$\begin{cases} du_t &= \frac{d}{dx} u_t \, dt + \phi(u_t) e^{Y_t \cdot} \, d\beta_t, \quad u_0 = u, \\ dY_t &= c(Y_t) \, dt + c_1(Y_t) \, d\tilde{\beta}_t, \quad Y_0 = y, \end{cases} \tag{1.3}$$

for $\tilde{\beta}$ being a real-valued Brownian motion independent of β, will not admit a generic finite dimensional realization, which in the current context is a desirable feature. Thus by replacing the parameter y with a stochastic process we arrived at a forward variance process (as we will show later) that is able to reach any curve in some open subset with positive probability. However, in this generality the analytic tractability is diminished. Inspired from real-world applications where such models are recalibrated at some discrete points in time (e.g. daily) in [50] it is suggested to look at the case where the process Y_t is piece-wise constant. That is, given a fixed grid of recalibration times $\{0 = t_0 < t_1 < ... < t_n < ...\}$ and setting $\tilde{Y}_t := \sum_{n \in \mathbb{N}_0} Y_{t_n} \mathbf{1}_{[t_n, t_{n+1})}(t)$ we see that formally (1.3) admits on each of this intervals $[t_n, t_{n+1})$ on the set $\{(u_{t_n}, Y_{t_n}) = (u_n, y_n)\}$ for $(u_n, y_n) \in U \times W$, where U is some open subset of $D(d/dx)$, a finite dimensional realization, as

$$du_t = \frac{d}{dx} u_t \, dt + \phi(u_t) e^{y_n \cdot} \, d\beta_t, \quad t > t_n, \qquad u_{t_n} = u_n, \tag{1.4}$$

is of the form (1.1) and accordingly the solution can again be represented (cf. (II.3.7)) by

$$u_t = S_{t-t_n} u_n + (\xi_t^n - u_n(t - t_n)) e^{y_n \cdot}, \quad t \in [t_n, t_{n+1}), \tag{1.5}$$

where the spot-variance process for this interval is given as the solution of

$$d\xi_t^n = (\theta^n(t) + y_n \xi_t^n) \, dt + \phi(\xi_t^n) \, d\beta_t, \quad \xi_{t_n}^n = u_n(0) \tag{1.6}$$

and the n-th Hull-White extension is

$$\theta^n(t) := \frac{d}{dt} u_n(t - t_n) - y_n u_n(t - t_n). \tag{1.7}$$

By repeating this steps we arrive at the suggested simulation algorithm [50, Definition 3.1], that is, starting with an initial forward variance curve u_0 and initial parameter value y_0 we can construct the initial Hull-White extension θ^0 according to (1.7). Then we can determine the solution of the initial spot-variance process ξ_t^0 given in (1.6) and simulate the next forward variance $u_1 := u_{t_1}$ by (1.5). By simulating $Y_{t_1} =: y_1$ we get the simulated output (u_1, y_1). By regarding this tupel as the input for the next simulation interval $[t_1, t_2)$ we arrive in the same fashion by using (1.7), (1.6) and (1.5) at (u_2, y_2) and more generally at (u_n, y_n). Thus we have constructed a discrete forward variance process that can change direction at every recalibration date and in particular is not stuck to the set (1.2). This becomes more evident by reparameterizing the forward variance realization (1.5). Indeed, notice that the algorithm is based on iteratively updating the Hull-White extension operator θ^n such that the next forward variance curve starts where the previous one ended. Thereby the effective algorithm consists of constructing a sequence of spot-variance processes (1.6) and to retrieve the corresponding forward variance curve from the short-end condition as in Proposition II.3.5. An equivalent formulation consists in constructing a sequence of the SPDEs (1.4) and using the Frobenius Theory of Chapter III to retrieve the finite dimensional realizations (1.5). In this case we get a reparamaterized representation of (1.5) that is given by (cf. Theorem III.3.9)

$$u_t = S_{t-t_n} u_n + Z_t^n e^{y_n \cdot}, \quad t \in [t_n, t_{n+1}),$$

where the coordinate process Z^n is given as the solution of

$$dZ_t^n = y_n Z_t^n \, dt + \phi(S_{t-t_n} u_n + Z_t^n e^{y_n \cdot}) \, d\beta_t, \quad Z_{t_n} = 0.$$

The resulting forward variance process in discrete time has the desired feature that it can change the direction (given by $e^{y_n \cdot}$, $n \in \mathbb{N}_0$) at every recalibration step. Indeed, looking at the Representation (1.5) we see that the initial curve u can be chosen freely within U and for each $n \geq 1$ the forward curve u_{t_n} satisfies

$$u_{t_n} \in \{S_{t_n - t_{n-1}} u_{t_{n-1}} + \langle e^{y_{n-1} \cdot} \rangle\} = \{S_{t_n} u + \sum_{i=1}^{n} \langle S_{t_n - t_i} e^{y_{i-1} \cdot} \rangle\}.$$

We end this summary by noting that this simulation algorithm, appropriately formulated, corresponds to an exponential Euler splitting scheme for the continuous-time consistent recalibration model given in (1.3) as discussed in [50, Section 3.1.2]. Given sufficient regularity of the vector fields in (1.3), it can be shown using [29] and [48] that this splitting scheme converges weakly as the mesh-size tends to zero. In [50, Theorem 4.4] this is shown for the consistent recalibration version of the Vasiček model. As the logarithm of the Bergomi forward variance model corresponds to a slightly friendlier version of the Vasiček model essentially the same proof can be used to show weak convergence. Also when aiming for affine realizations for forward interest rates the short-rate process (or more generally the state processes) necessarily (see [72]) belong to the class of affine processes which gives for the short-rate basically the choice between the processes of Vasiček or CIR type. Hence the possible choices of the diffusion coefficients is quite limited. On the other hand, when considering affine realizations for forward variance processes the corresponding class of processes is given by the affine drift models (cf. Definition II.3.4) and hence the choice of the diffusion coefficients is arbitrary (within the obvious regularity requirements). In particular these can be chosen arbitrarily regular to meet the conditions of [48] to show weak convergence. However, we do not pursue this further but introduce in the following the continuous-time forward variance consistent recalibration (henceforth CRC) models and show that these meet the conditions of Conjecture IV.3.3 which will allow us to look at the weak Taylor approximations.

2 CRC Forward Variance Models

In the following we will look at the continuous time CRC models corresponding to the affine and exponentially affine models from Sections III.3 and III.4 respectively.

When looking in the latter case at the logarithm of the forward variance process then we have in both cases models that lead to affine realizations under the conditions stated in Chapter III. We extend this models by turning certain parameters into state variables by replacing them with finite dimensional diffusion processes. For the small-parameter expansion we have in mind, we look for extensions that satisfy an infinite dimensional version of Hörmander's theorem on a very regular space of forward curves, i.e. on a space on which the operator $\frac{d}{dx}$ generates a strongly continuous group (and not only a strongly continuous semigroup). Examples of such spaces that also satisfy the conditions **(H1)** and **(H2)** of Section II.5 are given by [4, Example 3] and [4, Example 4], where (basically) the latter corresponds to the Hilbert space given in Remark II.5.1 (which was introduced in [32, Chapter 5]) extended to the whole real line which accordingly is too narrow to accommodate the function $x \mapsto e^{bx}$ where b is a strictly negative real number. As this vector will play a crucial role in the following, we will consider the Hilbert space of [4, Example 3] which was introduced by Tomas Björk and Lars Svensson in [13, Definition 4.1]. A major drawback of this space in the context of forward interest models is that it does not accommodate the invariant submanifold generated by a short-rate model that is given by the (Hull-White) extended CIR process (for more details on this, see [38, Page 3]). This lack ultimately lead to a series of papers by Damir Filipović and Josef Teichmann including [38] and [40] that utilized a Frobenius Theorem on Frèchet spaces using convenient calculus (cf. also to [73]). However, when it comes to forward variance models this problem is much less severe. In particular, all finite dimensional realizations including the short-variance realisations corresponding to the (Hull-White extended) Heston model (i.e. the short-variance follows a CIR process) are accommodated by this space and thus, for our applications, the utilization of this space means no (major) loss of generality. For fixed real numbers $\beta > 1$ and $\gamma > 0$ this Hilbert space $H_{\beta,\gamma}$ is given by the space of infinitely differentiable functions h on $\mathbb{R}_+$ satisfying the norm condition $\|h\|_{\beta,\gamma} < \infty$, where the norm is defined as

$$\|h\|_{\beta,\gamma}^2 := \sum_{n=0}^{\infty} \beta^{-n} \int_0^{\infty} \left(\frac{d^n}{dx^n} h(x) \right)^2 e^{-\gamma x}\, dx. \tag{2.1}$$

Remark 2.1. Notice that with the choice of the Hilbert space $H_{\beta,\gamma}$ the condition **(A3)** in Section III.3 is not satisfied anymore, however, this condition was needed only to prove necessity in certain representations, such as in the proof of Theorem III.3.3. For the sufficient conditions on the existence of generic finite dimensional realizations as given for example in Theorem III.3.9 and III.4.9 that condition is not required.

According to [13, Proposition 4.2] this space consists of entire analytic functions and satisfies the conditions **(H1)** and **(H2)** of Section II.5 with the additional property that the operator $\frac{d}{dx}$ is bounded on $H_{\beta,\gamma}$. As already mentioned, it will be crucial in the following that the function $x \mapsto e^{bx}$ belongs to this space.

Lemma 2.2. *We have $e^{b\cdot} \in H_{\beta,\gamma}$ if and only if $b \in (-\sqrt{\beta}, \frac{\gamma}{2})$.*

Proof. This follows immediately from the definition of the norm (2.1), as

$$\|e^{b\cdot}\|_{\beta,\gamma}^2 = \sum_{n=0}^{\infty} \left(\frac{b^2}{\beta} \right)^n \int_0^{\infty} e^{(-\gamma+2b)x}\, dx,$$

and hence the claim. □

It holds also true that for every polynomial p, the pointwise product $m(p, e^{b\cdot})$ belongs to $H_{\beta,\gamma}$ (see [13, Page 218]), for a suitable parameter combination β, γ. In the next Lemma we show that the conditions of Lemma 2.2 are sufficient for this, that is, a sufficient set of parameters β, γ can be found such that $m(p, e^{b\cdot}) \in H_{\beta,\gamma}$ for all polynomials p (i.e. the parameters depend only on b).

Lemma 2.3. *Let $b \in (-\sqrt{\beta}, \frac{\gamma}{2})$. Then $m(p, e^{b\cdot}) \in H_{\beta,\gamma}$ for every polynomial p.*

Proof. Let p be a polynomial of degree m. Then we have for $n > m$ by the Leibniz rule

$$\frac{d^n}{dx^n}p(x)e^{bx} = \sum_{k=0}^{n} \binom{n}{k} p^{(n-k)}(x)b^k e^{bx} = e^{bx} \sum_{k=n-m}^{n} \binom{n}{k} p^{(n-k)}(x)b^k$$

and by denoting $\tilde{p}(x) := \max(p^{(0)}(x), ..., p^{(m)}(x))$, $C_1 := \int_0^\infty e^{(2b-\gamma)x}\tilde{p}^2(x)\,dx$ and $C_2 := \sum_{n=0}^{m} \beta^{-n} \int_0^\infty \left(\frac{d^n}{dx^n}p(x)e^{bx}\right)^2 e^{-\gamma x}\,dx$ we get

$$\begin{aligned} \|m(p, e^{b\cdot})\|^2_{\beta,\gamma} &\leq C_2 + \sum_{n=m+1}^{\infty} \beta^{-n} \int_0^\infty e^{(2b-\gamma)x} m^2 \tilde{p}^2(x) \Big(\sum_{k=n-m}^{n} \binom{n}{k} b^k \Big)^2 dx \\ &= C_2 + C_1 m^2 \sum_{n=m+1}^{\infty} \Big(\frac{b^2}{\beta}\Big)^n \Big(\sum_{k=0}^{m} \binom{n}{k} b^{-k} \Big)^2 \end{aligned}$$

and for $n > 2m$ we get

$$\begin{aligned} \|m(p, e^{b\cdot})\|^2_{\beta,\gamma} &\leq C_3 + C_1 m^2 (m+1)^2 b^{-2m} \sum_{n=2m+1}^{\infty} \Big(\frac{b^2}{\beta}\Big)^n \binom{n}{m}^2 \\ &\leq C_3 + C_1 m^2 (m+1)^2 b^{-2m} (m!)^2 \sum_{n=2m+1}^{\infty} \Big(\frac{b^2 (n^{2m})^{\frac{1}{n}}}{\beta}\Big)^n \end{aligned}$$

and as $(n^{2m})^{\frac{1}{n}} \to 1$ for $n \to \infty$ the result follows from dominated convergence as soon as $\frac{b^2}{\beta} < 1$ which gives the claim. □

In the following we will consider the product space $\mathcal{H}_{\beta,\gamma} := H_{\beta,\gamma} \times \mathbb{R}^m$ associated with the norm $\|(h, y)\|_{\mathcal{H}_{\beta,\gamma}} := \sqrt{\|h\|^2_{\beta,\gamma} + \|y\|^2_{\mathbb{R}^m}}$. For normed spaces X and Y we denote by $C_b^\infty(X; Y)$ the space of mappings from X into Y such that each higher derivative exists and is bounded.

Lemma 2.4. *Let $\phi : H_{\beta,\gamma} \to \mathbb{R}$ be a bounded function that is in $C_b^\infty(H_{\beta,\gamma}; \mathbb{R})$. Further let b be a bounded function in $C_b^\infty(\mathbb{R}^d; \mathbb{R})$ such that there are positive real numbers b_1, b_2 and $b(\mathbb{R}^d) = [-b_1, b_2] \subset (-\sqrt{\beta}, \frac{\gamma}{2})$. Then the vector field $V : \mathcal{H}_{\beta,\gamma} \to \mathcal{H}_{\beta,\gamma}$ given by*

$$V(h, y) := \begin{pmatrix} \phi(h)e^{b(y)\cdot} \\ 0 \end{pmatrix} \tag{2.2}$$

is bounded and belongs to $C_b^\infty(\mathcal{H}_{\beta,\gamma})$.

Proof. First it follows from Lemma 2.3 that $V(\mathcal{H}_{\beta,\gamma}) \subset \mathcal{H}_{\beta,\gamma}$. For the second claim, we need to show that each higher derivative exists and is bounded, that is,

$$\sup_{(h,y)\in\mathcal{H}_{\beta,\gamma}} \|D^l V(h, y)\|^2_{L^l(\mathcal{H}_{\beta,\gamma})} \leq C(l) < \infty,$$

for $l \geq 0$ where $C(l)$ is some positive real number depending only on l. It is evident that it is sufficient to show that $V^1 \in C_b^\infty(\mathcal{H}_{\beta,\gamma}; H_{\beta,\gamma})$, where as usual V^1 denotes the first coordinate. For $l = 0$ we have

$$\|\phi(h)e^{b(y)\cdot}\|^2_{H_{\beta,\gamma}} = |\phi(h)|^2\|e^{b(y)\cdot}\|^2_{H_{\beta,\gamma}} \leq C_{\phi,0}C_{b,0} =: C(0) < \infty,$$

where $C_{\phi,0} := \sup_{h\in H_{\beta,\gamma}} |\phi(h)|^2$ and $C_{b,0} := \sup_{y\in\mathbb{R}^m} \|e^{b(y)\cdot}\|^2_{H_{\beta,\gamma}} = \|e^{b_2\cdot}\|^2_{H_{\beta,\gamma}}$ (recall that it follows from Lemma 2.2 that $b \mapsto \|e^{b\cdot}\|^2_{H_{\beta,\gamma}}$ is increasing). For $l \geq 1$ it is straight forward to show that

$$D^l V(h, y) = \sum_{j=0}^{l} p_{l,j}(h, y) Id(\cdot)^j e^{b(y)\cdot}, \tag{2.3}$$

where $Id(\cdot)^j \in H_{\beta,\gamma}$ is the monomial $x \mapsto x^j$ and $p_{l,j} : \mathcal{H}_{\beta,\gamma} \to L(\mathcal{H}^l_{\beta,\gamma}, \mathbb{R})$ are bounded mappings, that are continuously differentiable, e.g.

$$p_{l,l}(h, y)(h_1, y_1)...(h_l, y_l) := \phi(h)D_y b(y)(y_1)\cdots D_y b(y)(y_l). \tag{2.4}$$

As it follows from Lemma 2.3 that under the conditions on the map b, we have $Id^j(\cdot)e^{b(y)\cdot} \in \mathcal{H}_{\beta,\gamma}$ for all $j \geq 0$ we see that $V \in C^\infty(\mathcal{H}_{\beta,\gamma})$. Now, letting $C_{l,i} := \sup_{(h,y)\in\mathcal{H}_{\beta,\gamma}} \|p_{l,i}(h, y)\|^2_{L^l(\mathcal{H}_{\beta,\gamma})}$ the boundedness follows immediately as

$$\sup_{(h,y)\in\mathcal{H}_{\beta,\gamma}} \|D^l V(h, y)\|^2_{L^l(\mathcal{H}_{\beta,\gamma})} \leq (l+1)\sum_{j=0}^{l} C_{l,j}\|Id^j(\cdot)e^{b_2\cdot}\|^2_{\mathcal{H}_{\beta,\gamma}} =: C(l) < \infty,$$

which is the claim. □

We can now introduce the CRC versions of the models that admit affine and exponentially affine realizations and show that these satisfy the conditions of Conjecture IV.3.3.

2.1 Affine CRC Models

The affine CRC models that we will consider in the following are generalized versions of the processes considered in the introduction (cf. (1.3)) and given by

$$\begin{cases} du = \frac{d}{dx}u_t\, dt + \sum_{j=1}^{d} \phi_j(u_t)e^{b_j(Y_t)\cdot}\, d\beta_t^j \\ dY_t = c(Y_t)\, dt + \sum_{j=d+1}^{d+m} c_j(Y_t)\, d\beta_t^j \\ (u_0, Y_0) \in \mathcal{H}_{\beta,\gamma}, \end{cases} \tag{2.5}$$

where $b_1, ..., b_d$ and $\phi_1, ..., \phi_d$ are as in Lemma 2.4 and $c, c_{d+1}, ..., c_{d+m}$ are in $C_b^\infty(\mathbb{R}^m)$. Then it follows from Lemma 2.2 that the joint process $h = (u, Y)$ taking values in $\mathcal{H}_{\beta,\gamma}$ is given as the strong solution of

$$\begin{cases} dh_t = (Ah_t + V(h_t))\, dt + \sum_{j=1}^{d+m} V_j(h_t)\, d\beta_t^j \\ h_0 \in \mathcal{H}_{\beta,\gamma}, \end{cases} \tag{2.6}$$

as the vector fields $V, V_1, ..., V_d$ on $\mathcal{H}_{\beta,\gamma}$ are each in $C_b^\infty(\mathcal{H}_{\beta,\gamma})$. As we will assume in the weak Taylor expansion that the vector fields depend linearly on the expansion parameter ϵ it follows that the solution h of (2.6) satisfy Condition **SWTA 1** of Theorem IV.3.2.

Notice that in (2.6) we have not chosen the most general configuration possible in that we could have chosen ϕ in (2.5) to depend on the stochastic parameter process Y as well. However, as discussed in the introduction, this would not have changed the set of attainable curves of the corresponding model where $b_1, ..., b_d$ are constants. For this reason and also for notational convenience we are choosing this setting. It follows from Theorem III.3.9 (recall also Remark 2.1) that for a constant parameter process $Y_t \equiv y$ (which will correspond to the case $\epsilon = 0$ in the expansion) and each $b_1(y), ..., b_d(y)$ taking distinct values, the corresponding *parameterized equation*

$$\begin{cases} du_t = \frac{d}{dx}u_t\, dt + \sum_{j=1}^d \phi_j(u_t)e^{b_j(y)\cdot}\, d\beta_t^j \\ u_0 \in H_{\beta,\gamma} \end{cases} \tag{2.7}$$

admits a generic finite dimensional realization around each $u_0 = u$ in $H_{\beta,\gamma}$ given by (cf. (III.3.22) with $m = 0$)

$$u_t = S_t u + \sum_{i=1}^d Z_t^i e^{b_i(y)\cdot}, \tag{2.8}$$

with

$$dZ_t = B(y)Z_t\, dt + \sum_{i=1}^d \phi_i\Big(S_t u + \sum_{k=1}^d Z_t^k e^{b_k(y)\cdot}\Big)e_i\, d\beta_t^i, \quad Z_0 = 0, \tag{2.9}$$

where $B(y)$ is the diagonal matrix with entries $b_1(y), ..., b_d(y)$. Thus we arrived at a very regular extension of forward variance processes that admit affine realizations. In particular, as the vector-fields in (2.6) are in C_b^∞, it can easily be shown that the corresponding simulation algorithm indicated in the introductorey Section 1 converges weakly with formal order one using the setting of [29] and [48].

2.2 Exponentially Affine CRC Models

The CRC versions of forward variance models that admit exponentially affine realizations, as investigated in III.4, are formally given by

$$\begin{cases} du_t = \frac{d}{dx}u_t\, dt + \sum_{j=1}^d \phi_j m(e^{b_j(Y_t)\cdot}, u_t)\, d\beta_t^j \\ dY_t = c(Y_t)\, dt + \sum_{j=d+1}^{d+m} c_j(Y_t)\, d\beta_t^j \\ (u_0, Y_0) \in \mathcal{H}_{\beta,\gamma}). \end{cases} \tag{2.10}$$

However, it immediately follows from Lemma 2.2 that the multiplication operator can not be defined on the full space $H_{\beta,\gamma}$. Indeed, letting $b \in (\frac{\gamma}{4}, \frac{\gamma}{2})$ we would have $e^{b\cdot} \in H_{\beta,\gamma}$ but $m(e^{b\cdot}, e^{b\cdot}) = e^{2b\cdot} \notin H_{\beta,\gamma}$. Also, even if the multiplication operator were defined on the full space $H_{\beta,\gamma}$ (or on some open subset that is left invariant by the solutions of (2.10)) for the conditions of Theorem IV.3.2 to be satisfied, we need to have forward variance equations that have vector fields in C_b^∞(as we have in (2.6)) and it is evident due to product structure in (2.10) that this will not be the case. Therefore we restrict our analysis to the logarithm of the forward variances (cf.(III.4.3)) as in this case we arrive at the sufficiently regular representation, which is given by

$$\begin{cases} dJ_t = (\frac{d}{dx}J_t - \frac{1}{2}\sum_{j=1}^d \phi_j^2 e^{2b_j(Y_t)\cdot})\, dt + \sum_{j=1}^d \phi_j e^{b_j(Y_t)\cdot}\, d\beta_t^j \\ dY_t = c(Y_t)\, dt + \sum_{j=d+1}^{d+m} c_j(Y_t)\, d\beta_t^j \\ (J_0, Y_0) \in \mathcal{H}_{\beta,\gamma}, \end{cases} \tag{2.11}$$

where the coefficients are as in (2.5) with the difference that now $2b_j(\mathbb{R}^m) \subset (-\sqrt{\beta}, \frac{\gamma}{2})$ for $j = 1, ..., d$ (cf. Lemma 2.2) and $\phi_1, ..., \phi_d$ are real numbers. It is obvious that, with the same reasoning as in (2.6), the joint process $h = (J, Y)$ is given by the strong solution of

$$\begin{cases} dh_t = (Ah_t + V(h_t))\,dt + \sum_{j=1}^{d+m} V_j(h_t)\,d\beta_t^j \\ h_0 \in \mathcal{H}_{\beta,\gamma} \end{cases}, \tag{2.12}$$

and hence as above satisfies the Condition **SWTA 1** of Theorem IV.3.2. The *parameterized version* given by

$$\begin{cases} dJ_t = (\frac{d}{dx}J_t - \frac{1}{2}\sum_{j=1}^{d}\phi_j^2 e^{2b_j(y)\cdot})\,dt + \sum_{j=1}^{d}\phi_j e^{b_j(y)\cdot}\,d\beta_t^j \\ J_0 \in H_{\beta,\gamma}, \end{cases} \tag{2.13}$$

admits a generic finite dimensional realization around any $J_0 \in H_{\beta,\gamma}$ and is given by (cf. (III.4.24) with $m = 0$ and $p = 0$)

$$J_t := Fl_t^{\pi}(J_0) + \sum_{j=1}^{d} Z_t^j e^{b_j(y)\cdot} \tag{2.14}$$

where Z is the $\mathbb{R}^d$-valued time-inhomogeneous diffusion process given as the solution of

$$dZ_t = B(y)Z_t\,dt + \sum_{j=1}^{d}\phi_j e_j\,d\beta_t^j, \quad Z_0 = 0, \tag{2.15}$$

where $B(y)$ is the diagonal matrix with entries $b_1(y), ..., b_d(y)$. Thus we arrive at a similar situation as in Section 2.1 but now only for the logarithm of the forward variance process. We will see that this will restrict the range (and usefulness) of possible applications.

3 Hypoellipticity

We have shown that both models (2.6) and (2.12) satisfy the condition **SWTA 1** of Theorem IV.3.2. In this section we show that both models satisfy conditions of the Conjecture IV.3.3 as well. We will divide this problem into two parts. In the first part we will show that for a linear map $l \in L(\mathcal{H}_{\beta,\gamma}; \mathbb{R}^N)$ the projected model $l \circ h_t$ admits a density with respect to the Lebesgue measure on $\mathbb{R}^N$ and in particular has an Malliavin-Covariance matrix (see (IV.3.1)) that is invertible. In [4, Theorem 1] a sufficient condition for this is given, which corresponds to an infinite dimensional version of the Hörmander's Theorem (cf. [63, Theorem 2.3.3] for the finite dimensional case). Note that due to the boundedness of the operator A (given in (2.6) and (2.12)) we do not have to perform the analysis on the Frechét space $dom(A^\infty)$ but can work on $\mathcal{H}_{\beta,\gamma}$. The same is true for [4, Proposition 2], where again we can restrict the analysis to $\mathcal{H}_{\beta,\gamma}$ as we have strong solutions (again for (2.6) and (2.12)) for every initial curve in $\mathcal{H}_{\beta,\gamma}$. According to [4, Theorem 1], a sufficient condition for the invertibility of the *Malliavin-Covariance Matrix* (and than also for the absolute-continuity of the probability distribution of $l \circ h_t$ with respect to the Lebesgue measure on $\mathbb{R}^N$) is given by the following version of Hörmander's condition:

(H) D_{h_0} is dense is $\mathcal{H}_{\beta,\gamma}$, where for $h_0 \in \mathcal{H}_{\beta.\gamma}$, D_{h_0} denotes the distribution generated by V_1,..., V_{d+m} and all multiple Lie brackets of $\Xi, V_1, ..., V_{d+m}$ at h_0, (where Ξ denotes the *Strato- novich drift* (see (III.2.2) of h_t), that is,

$$D_{h_0} := \langle V_1(h_0), ..., V_{d+m}(h_0), [V_i, V_j](h_0), ..., [\Xi, V_i](h_0), ...\rangle. \tag{3.1}$$

Before investigating this condition separately for (2.6) and (2.12), we make the following condition on the parameter process Y appearing as the autonomous process in the second coordinate of h.

(HY) The parameter process Y satisfies the finite dimensional Hörmander condition for every $y \in \mathbb{R}^m$.

3.1 Affine CRC Models

The vector-fields in (2.6) are given by

$$V(u,Y) = \begin{pmatrix} 0 \\ c(Y) \end{pmatrix}, \quad V_i(u,Y) = \begin{pmatrix} \phi_i(u)e^{b_i(Y)\cdot} \\ 0 \end{pmatrix}, \quad V_j(u,Y) = \begin{pmatrix} 0 \\ c_j(Y) \end{pmatrix},$$

for $i = 1, ..., d$ and $j = d+1, ..., d+m$. Accordingly the Stratonovich drift Ξ is

$$\begin{aligned} \Xi(u,Y) &= A(u,Y) + V(u,Y) - \frac{1}{2}\sum_{j=1}^{d+m} DV_j(u,Y)V_j(u,Y) \\ &= \begin{pmatrix} \frac{d}{dx}u \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ c(Y) \end{pmatrix} - \frac{1}{2}\sum_{j=1}^{d} \begin{pmatrix} \zeta_j(u,Y)e^{b_j(Y)\cdot} \\ 0 \end{pmatrix} - \frac{1}{2}\sum_{j=d+1}^{d+m} \begin{pmatrix} 0 \\ D_Y c_j(Y)\cdot c_j(Y) \end{pmatrix} \end{aligned}$$

where the bounded scalar-fields ζ_j on $\mathcal{H}_{\beta,\gamma}$ for $j = 1, ..., d$ are given by

$$\zeta_j(u,Y) := (D_u\phi_j(u)\cdot\phi_j(u)e^{b_j(Y)\cdot}). \tag{3.2}$$

We have for $i = 1, ..., d$ that

$$\begin{aligned} &[\Xi, V_i](u,Y) \\ = &\begin{pmatrix} p_i^1(u,Y)e^{b_i(Y)\cdot} - \frac{1}{2}\sum_{j=1}^d p_{ij}^2(u,Y)e^{b_j(Y)\cdot} + \phi_i(u)p_i^3(u,Y)Id(\cdot)e^{b_i(Y)\cdot} \\ 0 \end{pmatrix}, \end{aligned}$$

with smooth bounded scalar-fields p_i^1, p_{ij}^2 and p_i^3 given by

$$\begin{aligned} p_i^1(u,Y) &:= \phi_i(u)b_i(Y) - D_u\phi_i(u)\cdot\Xi^1(u,Y), \\ p_{ij}^2(u,Y) &:= D_u\zeta_j(u,Y)\cdot V_i^1(u,Y), \\ p_i^3(u,Y) &:= D_Y b_i(Y)\cdot\Xi^2(u,Y). \end{aligned}$$

Thus if $\phi_i(u) \neq 0$ and $p_i^3(u,Y) \neq 0$ for $i = 1, ..., d$ we have

$$[\Xi, V_i](u,Y) \notin \langle V_1(u,Y), ..., V_d(u,Y)\rangle$$

for $i = 1, ..., d$. By iterating this procedure we see that the distribution generated by $V_1, ..., V_d$ and iterated Lie brackets of $V_1, ..., V_d$ with Ξ at (u,Y) contains the distribution

$$\begin{aligned} D^1(u,Y) &:= \left\langle\left\{V_1(u,Y), ..., V_d(u,Y), \begin{pmatrix} Id(\cdot)^m e^{b_j(Y)\cdot} \\ 0 \end{pmatrix} \;\middle|\; j = 1, ..., d, \quad m \in \mathbb{N}\right\}\right\rangle \\ &= \left\langle\left\{\begin{pmatrix} Id(\cdot)^m e^{b_j(Y)\cdot} \\ 0 \end{pmatrix} \;\middle|\; j = 1, ..., d, \quad m \in \mathbb{N}_0\right\}\right\rangle. \end{aligned} \tag{3.3}$$

Also, due to Condition **(HY)** we have that

$$D^2(u,Y) := \langle V(u,Y), V_{d+1}(u,Y), ..., V_{d+m}(u,Y), [V_j, V_k](u,Y), ...\rangle = \mathbb{R}^m. \tag{3.4}$$

As $\langle D^1(u,Y) \times D^2(u,Y)\rangle$ is dense in $\mathcal{H}_{\beta,\gamma}$ it follows from $\langle D^1(u,Y) \times D^2(u,Y)\rangle \subset D(u,Y)$ that the same is true for $D(u,Y)$. In particular we have in this case that $D^1(u,Y)$ is dense in $H_{\beta,\gamma}$.

Proposition 3.1. *For every $l \in L(\mathcal{H}_{\beta,\gamma}; \mathbb{R}^k)$, the (probability) distribution of $l \circ h_t$, where h_t is the solution of (2.6), is absolutely continuous with respect to the Lebesgue measure. In particular, the Malliavin-Covariance Matrix of $l \circ h_t$ is invertible.*

Proof. This follows from [4, Theorem 1] as Condition **(H)** is satisfied. □

Thus we have shown the first part of the Conjecture IV.3.3. For the second part, namely the integrability of the inverse Malliavin matrix at $\epsilon = 0$ we can utilize the parameterized model (2.7) and the finite dimensional representation of the solution given by (2.8) with coordinate process (2.9). Thus by choosing a linear map $L(H_{\beta,\gamma}; \mathbb{R}^N)$ (which we understand as a map on $\mathcal{H}_{\beta,\gamma}$) we can represent $l \circ h_t = l \circ u_t$ as

$$l \circ u_t = l(S_t u) + \sum_{i=1}^{d} Z_t^i l(e^{b_i(y)\cdot}). \tag{3.5}$$

We show first that the inverse of the Malliavin matrix of the coordinate process Z exists and its determinant is in $L^p(\Omega)$ for all $p \geq 1$. As Z is a time-inhomogeneous diffusion we can use a slightly modified version of the classic Hörmander's Theorem given in [23, (1.10)] (see also [41]). For $u \in H_{\alpha,\gamma}$ we set

$$U_j(t,h) := \phi_j\Big(S_t u + \sum_{i=1}^{d} h_i e^{b_i(y)\cdot}\Big) e_j, \quad j = 1, ..., d,$$

and

$$U(t,h) := B(y)h - \frac{1}{2}\sum_{j=1}^{d} \phi_j\Big(S_t u + \sum_{i=1}^{d} h_i e^{b_i(y)\cdot}\Big) \frac{\partial}{\partial_{h^j}} \phi_j\Big(S_t u + \sum_{i=1}^{d} h_i e^{b_i(y)\cdot}\Big) e_j,$$

giving the diffusion vector fields on $\mathbb{R}^d$ and Stratonovich drift of (2.9). In the following we understand the mappings as vector fields on $\mathbb{R}^{d+1}$ by adding a 0-th coordinate consisting of zeros. Then it is apparent (recall that $u \in H_{\beta,\gamma}$ and hence $t \mapsto S_t u$ is in C^∞) that for every $T > 0$ the extended vector fields $U, U_1, ..., U_d$ (we use the same notation) are in $C_b^\infty([0,T] \times \mathbb{R}^d)$ and hence we can apply the (extended) Hörmander's condition from [23, (1.10)] given by

$$\langle U_0(0,h), U_1(0,h), ..., U_d(0,h), [U_i, U_j](0,h), ...\rangle = \mathbb{R}^{d+1}, \tag{3.6}$$

where $U_0 := e_0 + U$ and e_0 denotes the zeroth basis vector. If we assume that $\phi_1, ..., \phi_d > 0$ then $U_1(0,h), ..., U_d(0,h)$ evidently are linearly independent and hence span $0 \times \mathbb{R}^d$. Thus the claim follows as $U_0(0,h)$ has a non zero zeroth component (given by 1). We summarize this in the following Lemma.

Lemma 3.2. *For every $T > 0$, the coordinate process Z_t (see (2.9)), for $t \in [0,T]$, has a smooth density that is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$. In particular, the Malliavin-covariance matrix $\gamma_t(Z)$, for $t \in [0,T]$ is invertible and $(\det \gamma(Z_t))^{-1} \in L^p(\Omega)$ for all $p \geq 1$.*

Proof. This follows from condition [23, (1.10)]. □

Now we look at the Malliavin-Covariance matrix of (3.5). Recalling (IV.3.1), we see that for $l \in L(H_{\beta,\gamma}; \mathbb{R}^N)$ with coordinate maps $l^1, ..., l^N$ we have

$$\gamma(l \circ u_t) = (\langle D(l^i \circ u_t), D(l^j \circ u_t)\rangle_{L^2([0,T];\mathbb{R}^d)})_{i,j=1,...,N}. \tag{3.7}$$

Thus from the chain-rule given in Lemma IV.2.4 and the Representation (3.5) we get (notice that the Malliavin derivative of $l \circ S_t u$ is zero, see Section IV.2)

$$\gamma(l \circ u_t) = \gamma\Big(\sum_{i=1}^{d} Z_t^i l(e^{b_i(y)\cdot})\Big).$$

If we set $L(y) := (l^{ij}(y))_{i=1,...,N, j=1,...,d}$ with $l^{ij}(y) := l^i(e^{b_j(y)\cdot})$ then we can represent

$$\gamma(l \circ u_t) = L(y)\gamma(Z_t)L^T(y) \tag{3.8}$$

and accordingly we arrive at the following proposition.

Proposition 3.3. *Let B denote the dimension of $V(y) := \langle e^{b_1(y)\cdot}, ..., e^{b_d(y)\cdot}\rangle$. If Z_t has a Malliavin Covariance matrix that is invertible with p-integrable inverse for all $p \geq 1$ then the same is true for $\gamma(l \circ u_t)$ if and only if $N \leq B \leq d$ and the map $l : V(y) \to \mathbb{R}^N$ satisfies $\langle l(e^{b_1(y)\cdot}), ..., l(e^{b_d(y)\cdot})\rangle = \mathbb{R}^N$.*

Proof. From (3.8) it is evident the claim follows if $L(y)$ has rank N but this is equivalent to the stated condition on l. □

Thus for affine CRC models we have shown that, subject to the conditions of the last proposition, the conditions of Conjecture IV.3.3 are satisfied.

3.2 Exponentially Affine CRC Models

The situation is very similar when looking at the model (2.12). The vector fields are given by

$$V(u, Y) = \begin{pmatrix} -\frac{1}{2}\sum_{j=1}^{d} \phi_j^2 e^{2b_j(Y)\cdot} \\ c(Y) \end{pmatrix},$$

$$V_i(u, Y) = \begin{pmatrix} \phi_i e^{b_i(Y)\cdot} \\ 0 \end{pmatrix}, \quad V_j(u, Y) = \begin{pmatrix} 0 \\ c_j(Y) \end{pmatrix},$$

for $i = 1, ..., d$ and $j = d+1, ..., d+m$. Accordingly the Stratonovich drift Ξ is given by

$$\Xi(u, Y) = A(u, Y) + V(u, Y) - \frac{1}{2}\sum_{j=1}^{d+m} DV_j(u, Y)V_j(u, Y)$$

$$= \begin{pmatrix} \frac{d}{dx}u \\ 0 \end{pmatrix} + \begin{pmatrix} -\frac{1}{2}\sum_{j=1}^{d} \phi_j^2 e^{2b_j(Y)\cdot} \\ c(Y) \end{pmatrix} - \frac{1}{2}\sum_{j=d+1}^{d+m} \begin{pmatrix} 0 \\ D_Y c_j(Y) \cdot c_j(Y) \end{pmatrix}.$$

Similar to above we find that for $i = 1, ..., d$

$$[\Xi, V_i](u, y) = \begin{pmatrix} p_i^1(Y)e^{b_i(Y)\cdot} - \phi_i p_i^3(u, Y)Id(\cdot)e^{b_i(y)\cdot} \\ 0 \end{pmatrix},$$

where

$$p_i^1(Y) := \phi_i b_i(Y) \quad p_i^3(u, Y) := D_y b_i(y)\Xi^2(u, Y).$$

Thus with the same reasoning as above we can state the following corollary to Proposition 3.1. Also when looking at (2.14) and (2.15) it is evident that the claims of Lemma 3.2 and Proposition 3.3 also hold.

Corollary 3.4. *The claims of Proposition 3.1, Lemma 3.2 and Proposition 3.3 remain true for h_t given as the solution of* (2.12) *and Z_t given as the solution of* (2.15).

Thus the Conjecture IV.3.3 are satisfied for the process given in (2.12) as well. However, for the exponentially affine CRC models the conditions are only true for the logarithm of the forward variances.

Chapter VI

Weak Taylor PDE Expansions

1 Introduction

So far we have introduced the Weak Taylor Expansions for SPDEs in Chapter IV and found with the CRC models in Chapter V a class of processes satisfying the conditions of Conjecture IV.3.3. From this we get the existence and the construction of the Weak Taylor approximation. The construction consists in computing the iterative Malliavin weights $\pi_0, \pi_1, ...$ in terms of the Skohorod integrals, cf. (IV.3.3) for π_1. However, from (IV.3.2) it is apparent that the Representation of the Malliavin weights is not unique. Indeed, we see in the notation of (IV.3.2) that $\tilde{\pi} := \mathbb{E}[\pi \,|\, F]$ satisfies the same Representation (IV.3.2). Moreover, $\tilde{\pi}$ corresponds in this situation to the *variance-minimal* Malliavin weight (see [42]) which is most suited for Monte-Carlo Simulations. More importantly, we will see in the following that the *variance-minimal* Malliavin weight lead in some cases to particularly nice representations that are most suited in situations where for the case $\epsilon = 0$ closed from expressions are available. Unfortunately, the construction via Skohorod integrals leads in general not to the variance-minimal weights. In the finite dimensional setting, in [69] the concept of *push-down* weights is introduced which provides a mean of explicit calculation of the variance minimal weights. We do not pursue this approach here, but look at a different but related expansion based on the Kolmogorov PDE. If we denote by X_t^ϵ the (mild) solution of (IV.1.1) and let $f \in C_b^\infty(H;\mathbb{R})$, then under the condition **SWTA 1** of Theorem IV.3.2 it follows that for every $n \geq 0$ we have the Taylor approximation of n-th order

$$\mathbb{E}[f(X_T^\epsilon) \,|\, X_t^\epsilon = x] = \sum_{i=0}^{n} \frac{\epsilon^i}{i!} \frac{\partial^i}{\partial \epsilon^i}\Big|_{\epsilon=0} \mathbb{E}[f(X_T^\epsilon) \,|\, X_t^\epsilon = x] + o(\epsilon^n) \quad \text{ as } \epsilon \to 0. \tag{1.1}$$

By denoting $v^\epsilon(t,x) := \mathbb{E}[f(X_T^\epsilon) \,|\, X_t^\epsilon = x]$ and $v_i(t,x) := \frac{1}{i!}\frac{\partial^i}{\partial \epsilon^i}\Big|_{\epsilon=0} \mathbb{E}[f(X_T^\epsilon) \,|\, X_t^\epsilon = x]$ for $i \geq 0$ we can state this expansion as

$$v^\epsilon(t,x) = \sum_{i=0}^{n} \epsilon^i v_i(t,x) + o(\epsilon^n), \quad \text{ as } \epsilon \to 0. \tag{1.2}$$

This expansion formally agrees with the weak Taylor expansion from Theorem IV.3.1 but requires only the smoothness of $\epsilon \mapsto v^\epsilon(t,x)$, which follows for example from condition **SWTA 1** of Theorem IV.3.2. The hypoellipticity condition **WTA 2** that is required for Malliavin's integration by parts formula is in this generality not necessary as the terms $v_i(t,x)$ are not expressed in terms of Malliavin weights. However, if the condition **WTA 2** is satisfied as well, then both expansions agree and we can use this PDE expansion to derive representations of the Malliavin weights. We will utilize here a similar expansion as in [70], where a PDE expansion for finite dimensional diffusions is considered. However, as opposed to [70] where again the concept of push-down weights are used to determine this coefficients, we employ a technique (formally) suggested by [8] for a direct calculation of the coefficients.

In the following we will look at two applications of this approach. The first case is motivated by [8] and correspond to a generalization of well-known small noise expansion for diffusive stochastic volatility models (cf. [69]) where the case $\epsilon = 0$ corresponds to the situation of an Black & Scholes model with time dependent but deterministic volatility. Here the expansion occurs around the partial derivatives (i.e. the so-called *Greeks*) of the Black & Scholes price. In the second application we will look at representations of the forward variance processes that depend on a finite-dimensional stochastic parameter process such that the unperturbed state $\epsilon = 0$ corresponds to the case where the parameter process becomes a constant. This is most suited for the CRC models introduced in Chapter V as the unperturbed state admits in this case a generic finite dimensional realization.

2 Implied Volatility Expansion

We consider the following generalized stochastic volatility model, where X denotes the log-price, u the forward variance and Y a stochastic parameter.

$$\begin{cases} dX_t^\epsilon = -\frac{1}{2}u_t^\epsilon(0)\,dt + \sqrt{u_t^\epsilon(0)}\,d\beta_t^1 \\ du_t^\epsilon = \frac{d}{dx}u_t^\epsilon\,dt + \epsilon\sum_{i=1}^d \sigma_i(u_t^\epsilon, Y_t^\epsilon)\,d\beta_t^i \\ dY_t^\epsilon = \epsilon^2\,c_0(Y_t^\epsilon)\,dt + \epsilon\sum_{i=1}^d c_i(Y_t^\epsilon)\,d\beta_t^i, \\ (X_0^\epsilon, u_0^\epsilon, Y_0^\epsilon) = (x,u,y) \in \mathbb{R}\times H\times\mathbb{R}^m, \end{cases} \tag{2.1}$$

where we understand that the square-root in the log-price process X_t^ϵ is approximated by a smooth function (see [69] for a discussion of this) such that the joint system $Z^\epsilon = (X^\epsilon, u^\epsilon, Y)$ on $\mathcal{H} := \mathbb{R}\times H\times\mathbb{R}^m$ satisfies the condition **SWTA 1** of Theorem IV.3.2 for the linear map $\ell \in L(\mathcal{H},\mathbb{R})$ that is given by the projection onto the first coordinate, that is, $\ell(x,u,y) = x$ for every $z = (x,u,y) \in \mathcal{H}$. In this case we have for every function $\tilde{f} : \mathbb{R}\to\mathbb{R}$ in C_b^∞ that the expansion (1.1) holds true with $f \in C_b^\infty(\mathcal{H};\mathbb{R})$ given by $f := \tilde{f}\circ\ell$. The expansion that we consider is conceptually very similar to the expansion suggested in [8] where the forward variances v^ϵ are parametrized in time-of-maturity, that is, $v_t^\epsilon(t+x) = u_t^\epsilon(x)$ for all $t, x \ge 0$. They consider the curve-valued process $T \mapsto v_t^\epsilon(T)$, where $T \ge t$, in which case the state-space moves with time. Indeed, let for each $t \ge 0$, B_t be some space of functions from $\mathbb{R}_{\ge t} \to \mathbb{R}$. Then it follows that $v_t^\epsilon \in B_t$ for each $t \ge 0$. Thus, they consider the following system

$$\begin{cases} dX_t^\epsilon &= -\frac{1}{2}v^\epsilon(t,t)\,dt + \sqrt{v^\epsilon(t,t)}\,d\beta_t^1, \quad X_0^\epsilon = x \in \mathbb{R}_+, \\ dv_t^\epsilon &= \epsilon\sum_{i=1}^d \sigma^i(t,\cdot,v_t^\epsilon)\,d\beta_t^i, \quad v^\epsilon(0,\cdot) = v \in B_0, \end{cases} \tag{2.2}$$

where

$$\sigma^i : \mathbb{R}_{\ge t}\times B_t \to B_t, \quad (t, v_t^\epsilon) \mapsto (\mathbb{R}_{\ge t} \ni T \mapsto \sigma^i(t,T,v_t^\epsilon)) \in B_t.$$

The state-space can be fixed to B_0, when we consider the process in Musiela's parametrization by setting $t \mapsto u_t^\epsilon(T) := v_t^\epsilon(T-t)$ which then leads formally to the SPDE formulation (1.1) (under some mild conditions, see [32]).

2.1 Expansion of Price

Under the condition **SWTA 1** of Theorem IV.3.2 the joint process $Z^\epsilon = (X^\epsilon, u^\epsilon, Y^\epsilon)$ on $\mathcal{H}$ satisfies the conditions of [26, Theorem 9.17] and accordingly for every $\tilde{f} \in C_b^\infty(\mathbb{R})$ and for all $\epsilon \in U$ the Kolmogorov equation with boundary condition $f := \tilde{f} \circ \ell$

$$\begin{cases} \partial_t v^\epsilon(t,(x,u,y)) - \frac{1}{2} D_x v^\epsilon(t,(x,u,y))(u(0)) + D_u v^\epsilon(t,(x,u,y))(\frac{d}{dx}u) \\ +\frac{1}{2} D_x^2 v^\epsilon(t,(x,u,y))(u(0)) + \epsilon D_{ux}^2 v^\epsilon(t,(x,u,y))(\sigma_1(u,y), \sqrt{u(0)}) \\ +\epsilon^2 D_y v^\epsilon(t,(x,u,y)) c_0(y) + \epsilon^2 \frac{1}{2} \sum_{j=1}^d D_y^2 v^\epsilon(t,(x,u,y)) c_j^2(y) \\ +\epsilon^2 \frac{1}{2} \sum_{j=1}^d D_u^2 v^\epsilon(t,(x,u,y))(\sigma_j(u,y), \sigma_j(u,y)) \\ +\epsilon^2 \sum_{i=1}^d D_{u,y}^2 v^\epsilon(t,(x,u,y))(\sigma_i(u,y), c_i(y)) = 0 \\ v^\epsilon(T,(x,u,y)) = f(x), \end{cases} \tag{2.3}$$

has a unique *strict solution* (see [26, (9.43)]) given by

$$v^\epsilon(t,(x,u,y)) = \mathbb{E}[f(X_T^\epsilon) | (X_t^\epsilon, u_t^\epsilon, Y_t^\epsilon) = (x,u,y)]. \tag{2.4}$$

That is $v^\epsilon : \mathbb{R}_+ \times \mathcal{H} \to \mathbb{R}$ satisfies

$$\begin{cases} v^\epsilon \in C_b^{0,0}([0,T] \times \mathcal{H}) \\ v^\epsilon(t,\cdot) \in C_b^\infty(\mathcal{H}) \quad \text{for all } t \geq 0, \\ v^\epsilon \in C^1([0,T] \times \mathbb{R} \times D(d/dx) \times \mathbb{R}^m) \\ \text{Equation (2.3) holds for any } u \in D(d/dx),\ t \in [0,T]. \end{cases} \tag{2.5}$$

In fact, by noting that it follows from the condition **SWTA 1** of Theorem IV.3.2 that $\epsilon \mapsto v^\epsilon(t,(x,u,y))$ is smooth it is easy to see that the following holds true.

Lemma 2.1. *We have*

$$v^{\cdot} \in C^{1,\infty}([0,T] \times \mathbb{R} \times D(d/dx) \times \mathbb{R}^m \times U). \tag{2.6}$$

Proof. This can be shown as in the proof of [26, Theorem 9.17]. □

By aggregating the operators in (2.3) we can write

$$\begin{cases} (\partial_t + L^\epsilon) v^\epsilon(t,(x,u,y)) = 0 \\ v^\epsilon(T,(x,u,y)) = f(x), \end{cases} \tag{2.7}$$

where

$$L^\epsilon := L^0 + \epsilon L^1 + \epsilon^2 L^2$$

such that for any $v \in C_b^2(\mathcal{H})$ and letting $z = (x, u, y)$

$$L^0 v(z) := \frac{1}{2}(D_x^2 v(z) - D_x v(z))(u(0)) + D_u v(z)(\frac{d}{dx}u) \tag{2.8}$$

$$L^1 v(z) := D_{ux}^2 v(z)(\sigma_1(u, y), \sqrt{u(0)}) \tag{2.9}$$

$$\begin{aligned} L^2 v(z) := & \frac{1}{2}\sum_{j=1}^{d} D_u^2 v(z)(\sigma_j(u, y), \sigma_j(u, y)) + D_y v(z) c_0(y) \\ & + \sum_{j=1}^{d} D_{uy}^2 v(z)(\sigma_j(u, y), c_j(y)) + \frac{1}{2}\sum_{j=1}^{d} D_y^2 v(z) c_j^2(y). \end{aligned} \tag{2.10}$$

Lemma 2.2. *Let* $\tilde{v}_n := \frac{\partial^n}{\partial \epsilon^n}\Big|_{\epsilon=0} v^\epsilon$*, then for each* $n \geq 0$*,* $\tilde{v}_n$ *is the solution of*

$$\begin{cases} (\partial_t + L^0)\tilde{v}_n + nL^1\tilde{v}_{n-1} + n(n-1)L^2\tilde{v}_{n-2} = 0 \\ \tilde{v}_n(T, (x, u, y)) = \frac{\partial^n}{\partial \epsilon^n}\Big|_{\epsilon=0} f(x). \end{cases} \tag{2.11}$$

Proof. It follows from Lemma 2.1 that for each $n \geq 0$ we can apply the operator $\frac{\partial^n}{\partial \epsilon^n}\Big|_{\epsilon=0}$ to (2.3). Hence we have for $t < T$ (setting $v^\epsilon := v^\epsilon(t, (x, u, y))$

$$\frac{\partial^n}{\partial \epsilon^n}\Big|_{\epsilon=0} (\partial_t + L^\epsilon) v^\epsilon = 0 \tag{2.12}$$

and as the mixed partial derivatives commute

$$\begin{aligned} & \frac{\partial^n}{\partial \epsilon^n}(\partial_t + L^0 + \epsilon L^1 + \epsilon^2 L^2) v^\epsilon \\ = & (\partial_t + L^0)\frac{\partial^n}{\partial \epsilon^n} v^\epsilon + \frac{\partial^n}{\partial \epsilon^n}\epsilon L^1 v^\epsilon + \frac{\partial^n}{\partial \epsilon^n}\epsilon^2 L^2, \end{aligned}$$

where after performing the necessary calculations we get

$$\frac{\partial^n}{\partial \epsilon^n}\epsilon L^1 v^\epsilon = nL^1\frac{\partial^{n-1}}{\partial \epsilon^{n-1}} v^\epsilon + \epsilon L^1 \frac{\partial^n}{\partial \epsilon^n} v^\epsilon$$

and

$$\frac{\partial^n}{\partial \epsilon^n}\epsilon^2 L^2 v^\epsilon = n(n-1)\frac{\partial^{n-2}}{\partial \epsilon^{n-2}} L^2 v^\epsilon + 2n\epsilon \frac{\partial^{n-1}}{\partial \epsilon^{n-1}} L^2 v^\epsilon + \epsilon^2 L^2 \frac{\partial^n}{\partial \epsilon^n} v^\epsilon.$$

Evaluating this terms at $\epsilon = 0$ and plugging them into (2.12) gives the claim. □

The functions in (1.2) satisfy $v_n = \frac{1}{n!}\tilde{v}_n$ and accordingly by dividing in (2.11) by $n!$ we see that the functions v_n satisfy the PDEs given by

$$(\partial_t + L^0)v_0 = 0, \quad v_0(T, (x, u, y)) = f(x) \tag{2.13}$$

$$(\partial_t + L^0)v_1 + L^1 v_0 = 0, \quad v_1(T, (x, u, y)) = 0 \tag{2.14}$$

$$(\partial_t + L^0)v_n + L^1 v_{n-1} + L^2 v_{n-2} = 0, \quad v_n(T, (x, u, y)) = 0, \quad n \geq 2. \tag{2.15}$$

From (2.1) we see that v_0 is (functionally) independent of y and satisfies

$$\begin{aligned} v_0(t, (x, u)) &= \mathbb{E}[f(X_T^0) \,|\, (X_t^0, u_t^0) = (x, u)] \\ &= \mathbb{E}\Big[f\Big(x - \frac{1}{2}\int_t^T u_s^0(0)\, ds + \int_0^t \sqrt{u_s^0(0)}\, d\beta_s^1\Big)\Big] \\ &= \mathbb{E}\Big[f\Big(x - \frac{1}{2}\int_t^T u(s-t)\, ds + \int_0^t \sqrt{u(s-t)}\, d\beta_s^1\Big)\Big], \end{aligned} \tag{2.16}$$

where we used in the last equation that $u_s^0 = S_{s-t}u$ for $s \geq t$ is the solution to

$$du_s^0 = \frac{d}{dx}u_s^0, \quad u_t^0 = u, \tag{2.17}$$

for every $u \in \mathcal{D}(\frac{d}{dx})$. For the functions v_n with $n \geq 1$ we can conclude from the representations (2.14)-(2.15) and the Feynman-Kac formula that

$$v_n(t,(x,u,y)) = \int_t^T \mathbb{E}[H_n(s,(X_s^0,u_s^0,y)) \,|\, (X_t^0,u_t^0) = (x,u)]\,ds, \tag{2.18}$$

holds true, where

$$\begin{cases} H_1(t,(x,u,y)) := L^1 v_0(t,(x,u,y)) \\ H_n(t,(x,u,y)) := L^1 v_{n-1}(t,(x,u,y)) + L^2 v_{n-2}(t,(x,u,y)), \ n \geq 2. \end{cases} \tag{2.19}$$

The following lemma will be crucial in the following as it will allow us to find explicit solutions to (2.18) and under the conditions of the next subsection to find the variance-minimal Malliavin weights. It can be found in [8] in a very similar form. Again, we notice that in [8] the system (2.2) is considered and that accordingly the differential operators (2.8)-(2.10) differ. Therefore we provide a proof.

Lemma 2.3. *For all $n \geq 1$ and $s \geq t$ it holds true that*

$$\mathbb{E}[D_x^n v_0(s,(X_s^0,u_s^0)) \,|\, (X_t^0,u_t^0) = (x,u)] = D_x^n v_0(t,(x,u)).$$

Proof. We see that $D_x^n v_0(t,(x,u))$ satisfies the PDE (2.13) with boundary condition $D_x^n f(x)$ (i.e. due to Lemma 2.6 we can apply D_x^n to (2.13)) and hence satisfies

$$D_x^n v_0(t,(x,u)) = \mathbb{E}[D_x^n f(X_T^0) \,|\, (X_t^0,u_t^0) = (x,u)].$$

Accordingly

$$D_x^n v_0(s,(X_s^0,u_s^0)) = \mathbb{E}[D_x^n f(X_T^0) \,|\, (X_s^0,u_s^0)]$$

and

$$\begin{aligned} & \mathbb{E}[\mathbb{E}[D_x^n f(X_T^0) \,|\, (X_s^0,u_s^0)] \,|\, (X_t^0,u_t^0) = (x,u)] \\ = \ & \mathbb{E}[D_x^n f(X_T^0) \,|\, (X_t^0,u_s^0) = (x,S_{s-t}u)] = \mathbb{E}[D_x^n f(X_T^0) \,|\, (X_t^0,u_t^0) = (x,u)] \end{aligned}$$

which gives the claim. □

The following lemma corresponds to the *vega-gamma relationship* for the log-price of the Black & Scholes model. It will allow us in the following to replace partial derivatives with respect to u by partial derivatives with respect to x.

Lemma 2.4. *$v_0(t,(x,u))$ satisfies the following* vega-gamma relationship

$$\begin{cases} D_u v_0(t,(x,u))h = \frac{1}{2}(D_x^2 - D_x)v_0(t,(x,u)) \int_t^T h(s-t)\,ds \\ v_0(t,(x,0)) = f(x). \end{cases} \tag{2.20}$$

for every $h \in H$.

Proof. This is straight forward upon noting that on $(X_t^0,u_t^0) = (x,u)$, X_T^0 is normally distributed with mean and variance given by $x - \frac{1}{2}\int_t^T u(s-t)\,ds$ and $\int_t^T u(s-t)\,ds$ respectively, from which the claim follows upon differentiating the normal density. □

In the following we will write $\mathbb{E}^{t,x,u}[\cdot] := \mathbb{E}[\cdot \mid (X_t^0, u_t^0) = (x,u)]$. For the computation of the prices $v_i(t,(x,u,y))$ the following Lemma will be useful.

Lemma 2.5. *Let $C \in C_b^\infty(\mathbb{R}_+ \times H_\alpha \times \mathbb{R}^m)$ and $n \geq 0$. Then*

$$\begin{cases} \int_t^T \mathbb{E}^{t,x,u}[L^1 D_x^n v_0(s, X_s^0, u_s^0, y) C(s, u_s^0, y)]\, ds \\ = D_x^n (D_x^3 - D_x^2) v_0(t,x,u) C_1(t,u,y) + D_x^n D_x v_0(t,x,u) C_2(t,u,y) \end{cases} \tag{2.21}$$

and

$$\begin{cases} \int_t^T \mathbb{E}^{t,x,u}[L^2 D_x^n v_0(s, X_s^0, u_s^0, y) C(s, u_s^0, y)]\, ds \\ = D_x^n (D_x^2 - D_x)^2 v_0(t,x,u) C_3(t,u,y) \\ + D_x^n (D_x^2 - D_x) v_0(t,x,u) C_4(t,u,y) + D_x^n v_0(t,x,u) C_5(t,u,y) \end{cases} \tag{2.22}$$

with

$$\begin{aligned} C_1(t,u,y) &:= \frac{1}{2} \int_t^T \int_t^r \mathbb{E}^{t,u}[S_{r-s}\sigma_1(u_s^0, y)(0)\, C(s, u_s^0, y) \sqrt{u_s^0(0)}]\, ds\, dr \\ C_2(t,u,y) &:= \int_t^T \mathbb{E}^{t,u}[\sqrt{u_s^0(0)} D_u C(s, u_s^0, y)(\sigma_1(u_s^0, y))]\, ds \\ C_3(t,u,y) &:= \frac{1}{8} \sum_{j=1}^d \int_t^T \mathbb{E}^{t,u}[\Big(\int_s^T S_{r-s}\sigma_j(u_s^0, y)(0)\, dr\Big)^2 C(s, u_s^0, y)]\, ds \\ C_4(t,u,y) &:= \frac{1}{2} \sum_{j=1}^d \int_t^T \mathbb{E}^{t,u}[\Big(\int_s^T S_{r-s}\sigma_j(u_s^0, y)(0)\, dr\Big) D_u C(s, u_s^0, y)(\sigma_j(u_s^0, y))]\, ds \\ &+ \frac{1}{2} \sum_{j=1}^d \int_t^T \mathbb{E}^{t,u}\Big(\int_s^T S_{r-s}\sigma_j(u_s^0, y)(0))\, dr\Big) D_y C(s, u_s^0, y)(c_j(y))]\, ds \\ C_5(t,u,y) &:= \frac{1}{2} \sum_{j=1}^d \int_t^T \mathbb{E}^{t,u}[D_u^2 C(s, u_s^0, y)(\sigma_j(u_s^0, y), \sigma_j(u_s^0, y))]\, ds \\ &+ \sum_{j=1}^d \int_t^T \mathbb{E}^{t,u}[D_y D_u C(s, u_s^0, y)(\sigma_j(u_s^0, y), c_j(y))]\, ds \\ &+ \int_t^T \mathbb{E}^{t,u}[(\frac{1}{2} D_y^2 C(s, u_s^0, y) c_j^2(y) + D_y C(s, u_s^0, y) c_0(y))]\, ds \end{aligned}$$

Proof. By recalling the Definition of L^1 in (2.9), we see that

$$\begin{aligned} & L^1 D_x^n v_0(s, X_s^0, u_s^0) C(s, u_s^0, y) \\ = \; & D_x^n D_x [D_u v_0(s, X_s^0, u_s^0)(\sigma_1(u_s, y))] C(s, u_s^0, y) \sqrt{u_s^0(0)} \\ + \; & D_x^n D_x v_0(s, X_s^0, u_s^0) \sqrt{u_s^0(0)} D_u C(s, u_s^0, y)(\sigma_1(u_s, y)) \end{aligned}$$

and recalling (2.20)

$$\int_t^T D_u v_0(s, X_s^0, u_s^0)(\sigma_1(u_s, y)) = \frac{1}{2}(D_x^2 - D_x) v_0(s, X_s^0, u_s^0) \int_s^T S_{r-s}\sigma_1(u_s^0, y)(0)\, dr$$

and finally with Lemma 2.3 and Fubini's Theorem

$$
\begin{aligned}
&\int_t^T \mathbb{E}^{t,x,u}[L^1 D_x^n v_0(s, X_s^0, u_s^0) C(s, u_s^0, y)]\, ds \\
= \quad & D_x^n (D_x^3 - D_x^2) v_0(t, x, u) \frac{1}{2} \int_t^T \int_t^r \mathbb{E}^{t,u}[S_{r-s} \sigma_1(u_s^0, y)(0)\, C(s, u_s^0, y) \sqrt{u_s^0(0)}]\, ds\, dr \\
+ \quad & D_x^n D_x v_0(t, x, u) \int_t^T \mathbb{E}^{t,u}[\sqrt{u_s^0(0)} D_u C(s, u_s^0, y)(\sigma_1(u_s, y))]\, ds,
\end{aligned}
$$

which gives the first claim. Similarly

$$
\begin{aligned}
& L^2 D_x^n v_0(s, X_s^0, u_s^0) C(s, u_s^0, y) \\
= \quad & \frac{1}{2} \sum_{j=1}^d D_x^n D_u^2 v_0(s, X_s^0, u_s^0) C(s, u_s^0, y)(\sigma_j(u_s^0, y), \sigma_j(u_s^0, y)) && (2.23) \\
+ \quad & \sum_{j=1}^d D_x^n D_y D_u v_0(s, X_s^0, u_s^0) C(s, u_s^0, y)(\sigma_j(u_s^0, y), c_j(y)) && (2.24) \\
+ \quad & \frac{1}{2} D_x^n v_0(s, X_s^0, u_s^0) D_y^2 C(s, u_s^0, y) c_j^2(y) + v_0(s, X_s^0, u_s^0) D_y C(s, u_s^0, y) c_0(y)
\end{aligned}
$$

and (2.23) satisfies

$$
\begin{aligned}
& \frac{1}{2} \sum_{j=1}^d D_x^n D_u^2 v_0(s, X_s^0, u_s^0) C(s, u_s^0, y)(\sigma_j(u_s^0, y), \sigma_j(u_s^0, y)) \\
= \quad & \frac{1}{2} \sum_{j=1}^d D_x^n D_u^2 v_0(s, X_s^0, u_s^0)(\sigma_j(u_s^0, y), \sigma_j(u_s^0, y)) C(s, u_s^0, y) \\
+ \quad & \sum_{j=1}^d D_x^n D_u v_0(s, X_s^0, u_s^0)(\sigma_j(u_s^0, y)) D_u C(s, u_s^0, y)(\sigma_j(u_s^0, y)) \\
+ \quad & \frac{1}{2} \sum_{j=1}^d D_x^n v_0(s, X_s^0, u_s^0) D_u^2 C(s, u_s^0, y)(\sigma_j(u_s^0, y), \sigma_j(u_s^0, y)) \\
= \quad & \frac{1}{8} \sum_{j=1}^d D_x^n (D_x^2 - D_x)^2 v_0(s, X_s^0, u_s^0) \\
& \Big(\int_s^T S_{r-s} \sigma_j(u_s^0, y)(0)\, dr \Big)^2 C(s, u_s^0, y) \\
+ \quad & \frac{1}{2} \sum_{j=1}^d D_x^n (D_x^2 - D_x) v_0(s, X_s^0, u_s^0) \\
& \Big(\int_s^T S_{r-s} \sigma_j(u_s^0, y)(0)\, dr \Big) D_u C(s, u_s^0, y)(\sigma_j(u_s^0, y)) \\
+ \quad & \frac{1}{2} \sum_{j=1}^d D_x^n v_0(s, X_s^0, u_s^0) D_u^2 C(s, u_s^0, y)(\sigma_j(u_s^0, y), \sigma_j(u_s^0, y))
\end{aligned}
$$

and for (2.24)

$$\begin{aligned}
&\sum_{j=1}^{d} D_y D_u D_x^n v_0(s, X_s^0, u_s^0) C(s, u_s^0, y)(\sigma_j(u_s^0, y), c_j(y)) \\
= \;& \frac{1}{2} \sum_{j=1}^{d} D_x^n (D_x^2 - D_x) v_0(s, X_s^0, u_s^0) \Big(\int_s^T S_{r-s} \sigma_j(u_s^0, y)(0))\, dr \Big) D_y C(s, u_s^0, y)(c_j(y)) \\
+ \;& \sum_{j=1}^{d} D_x^n v_0(s, X_s^0, u_s^0) D_y D_u C(s, u_s^0, y)(\sigma_j(u_s^0, y), c_j(y))
\end{aligned}$$

and finally putting all together and using Lemma 2.3

$$\begin{aligned}
&\int_t^T \mathbb{E}^{t,x,u}[L^2 v_0(s, X_s^0, u_s^0) C(s, u_s^0, y)]\, ds \\
= \;& D_x^n (D_x^2 - D_x)^2 v_0(t, x, u) \frac{1}{8} \sum_{j=1}^{d} \int_t^T \mathbb{E}^{t,u}\Big[\Big(\int_s^T S_{r-s} \sigma_j(u_s^0, y)(0)\, dr \Big)^2 C(s, u_s^0, y)\Big]\, ds \\
+ \;& D_x^n (D_x^2 - D_x) v_0(t, x, u) \\
& \frac{1}{2} \sum_{j=1}^{d} \int_t^T \mathbb{E}^{t,u}\Big[\Big(\int_s^T S_{r-s} \sigma_j(u_s^0, y)(0)\, dr \Big) D_u C(s, u_s^0, y)(\sigma_j(u_s^0, y))\Big]\, ds \\
+ \;& D_x^n v_0(t, x, u) \frac{1}{2} \sum_{j=1}^{d} \int_t^T \mathbb{E}^{t,u}[D_u^2 C(s, u_s^0, y)(\sigma_j(u_s^0, y), \sigma_j(u_s^0, y))]\, ds \\
+ \;& D_x^n (D_x^2 - D_x) v_0(t, x, u) \\
& \frac{1}{2} \sum_{j=1}^{d} \int_t^T \mathbb{E}^{t,u} \Big(\int_s^T S_{r-s} \sigma_j(u_s^0, y)(0))\, dr \Big) D_y C(s, u_s^0, y)(c_j(y))]\, ds \\
+ \;& D_x^n v_0(t, x, u) \sum_{j=1}^{d} \int_t^T \mathbb{E}^{t,u}[D_y D_u C(s, u_s^0, y)(\sigma_j(u_s^0, y), c_j(y))]\, ds \\
+ \;& D_x^n v_0(t, x, u) \int_t^T \mathbb{E}^{t,u}[(\frac{1}{2} D_y^2 C(s, u_s^0, y) c_j^2(y) + D_y C(s, u_s^0, y) c_0(y))]\, ds
\end{aligned}$$

gives the second claim. □

From (2.18) the price at first order is given by

$$v_1(t, (x, u, y)) = \int_t^T \mathbb{E}^{t,x,u}[L^1 v_0(s, X_s^0, u_s^0)]\, ds,$$

and hence from Lemma 2.5 with $n = 0$ and $C \equiv 1$ we see that

$$v_1(t, (x, u, y)) = (D_x^3 - D_x^2) v_0(t, x, u) C_{1,1}(t, (u, y)), \tag{2.25}$$

with $C_{1,1}(t, u, y) := C_1(t, u, y)$ and $C_1(t, u, y)$ as in the Lemma 2.5 for $C \equiv 1$. We notice that

$$\begin{aligned}
C_{1,1}(t, (u, y)) &= \frac{1}{2} \int_t^T \int_t^r \mathbb{E}^{t,u}[S_{r-s} \sigma^1(u_s^0, y)(0) \sqrt{u_s^0(0)})]\, ds\, dr \\
&= \frac{1}{2} \int_t^T \mathbb{E}^{t,u}[\langle \tilde{u}^0(0), X^0 \rangle_r - \langle \tilde{u}^0(0), X^0 \rangle_t]\, dr,
\end{aligned} \tag{2.26}$$

with $\tilde{u}_t^0 = \frac{\partial}{\partial \epsilon}\Big|_{\epsilon=0} u_t^\epsilon$. For the price at second order we have according to (2.18)

$$v_2(t,x,u,y) = \int_t^T \mathbb{E}^{t,x,u}[L^1 v_1(s, X_s^0, u_s^0, y) + L^2 v_0(s, X_s^0, u_s^0)]\, ds$$

and thus from (2.25) we can again apply Lemma (2.5) on each term, i.e. for the first term it is $n = 3$ and $n = 2$ and $C = C_{1,1}$ while for the second it is $n = 0$ and $C = 1$. This gives

$$\begin{cases} v_2(t,(x,u,y)) = (D_x^3 - D_x^2)^2 v_0(t,x,u) C_{2,1}(t,(u,y)) \\ +(D_x^4 - D_x^3) v_0(t,x,u) C_{2,2}(t,u,y) + (D_x^2 - D_x)^2 v_0(t,x,u) C_{2,3}(t,u,y) \end{cases} \tag{2.27}$$

where $C_{2,1}$ and $C_{2,2}$ corresponds to C_1 and C_2 given in Lemma 2.5 for $C = C_{1,1}$ and $C_{2,3}$ to C_3 with $C = 1$. More generally, we can show the following.

Proposition 2.6. *The function $v_n(t,(x,u,y))$, for $n \geq 0$, satisfies*

$$v_n(t,(x,u,y)) = \sum_{i=0}^{m(n)} D_x^i v_0(t,x,u) \tilde{C}_i(t,u,y),$$

where $n \mapsto m(n)$ is an increasing sequence of natural numbers and $\tilde{C}_i$ are smooth, deterministic functions.

Proof. This can be shown by induction in n. As shown in (2.25) and (2.27) the claim holds for $n = 0, 1, 2$. Assuming it holds for some $n \geq 3$, it follows from (2.19), (2.18), 2.3 and the fact that L^1 and L^2 commute with D_x that it is sufficient to show that

$$L^k v_0(t,x,u) C(t,(u,y)) = \sum_{i=0}^{m} \partial_x^i v_0(t,x,u) \bar{C}_{k,i}(t,u,y)$$

for $k = 1, 2$, some natural number m and smooth, deterministic functions $\bar{C}_{k,i}$, $i = 0, ..., m$. But this follows immediately from (2.20) and the definitions of L^1 and L^2 given in (2.9) and (2.10). □

Remark 2.7. We comment here on the functions $C_{1,1}, C_{2,2}$ and $C_{2,3}$. The function $C_{1,1}(0,(u,y))$ given in (2.26) satisfies after an application of the Fubini Theorem

$$C_{1,1}(0,u,y) = \int_0^T \int_0^r \sigma_1(S_s u, y)(r-s)\sqrt{u(s)}\, ds\, dr.$$

By noting that the process $\tilde{u}_t^0 := \frac{\partial}{\partial \epsilon}\Big|_{\epsilon=0} u_t^\epsilon$ satisfies

$$\tilde{u}_t^0 = \sum_{j=1}^{d} \int_0^t S_{t-s} \sigma_j(S_s u, y)\, d\beta_s^j,$$

we find that

$$C_{1,1}(0,u,y) = \int_0^T \langle X^0, \tilde{u}^0(0) \rangle_r \, dr \tag{2.28}$$

and hence is the integrated quadratic covariation between the log-price process X^ϵ and $\tilde{u}^0(0)$. Similarly

$$\begin{aligned} & C_{2,3}(0,u,y) \\ = \; & \frac{1}{8}\sum_{j=1}^{d}\int_0^T\int_0^T\int_0^{\min(r_1,r_2)} \sigma_j(S_s u,y)(r_1-s)\sigma_j(S_s u,y)(r_2-s)\,ds\,dr_1\,dr_2 \\ = \; & \frac{1}{4}\sum_{j=1}^{d}\int_0^T\int_0^{r_2}\int_0^{r_1} S_{r_1-s}\sigma_j(S_s u,y)(0)S_{r_1-s}\sigma_j(S_s u,y)(r_2-r_1)\,ds\,dr_1\,dr_2 \\ = \; & \frac{1}{4}\int_0^T\int_0^{r_2}\langle \tilde{u}^0(0),\tilde{u}^0(r_2-r_1)\rangle_{r_1}\,dr_1\,dr_2 \end{aligned}$$

corresponds to the integrated auto-covariation of $\tilde{u}^0$. This quantities agree with the corresponding quantities in [8].

2.2 Malliavin Weights

We assume now that the Conjecture IV.3.3 is satisfied as well, and accordingly in this case (2.1) admits a weak Taylor expansion. Hence each term of the expansion (1.2) satisfies

$$v_n(0,x,u,y) = \mathbb{E}[f(X_T^0)\pi_n] = \mathbb{E}[f(X_T^0)\mathbb{E}[\pi_n \,|\, X_T^0]], \tag{2.29}$$

where π_n and $\mathbb{E}[\pi_n \,|\, X_T^0]$ denote the n-th Malliavin weight and n-th variance-minimal Malliavin weight, respectively. Thus using Lemma 2.2 and Proposition 2.6 for $f \in C_b^\infty$ as above

$$\mathbb{E}^{x,u}[f(X_T^0)\mathbb{E}^{x,u}[\pi_n \,|\, X_T^0]] = n! \sum_{i=0}^{m(n)} \partial_x^i v_0(0,x,u)\tilde{C}_i(0,u,y).$$

If we now denote the density of X_T^0 for $(X_0^0,u_0^0)=(x,u)$ by $p(\cdot;x,u)$ we see that

$$\begin{aligned} & \int_{-\infty}^{\infty} f(z)\mathbb{E}^{x,u}[\pi_n \,|\, X_T^0 = z]\,p(z;x,u)\,dz \\ = \; & \int_{-\infty}^{\infty} f(z)\Big(\sum_{i=0}^{m(n)} \frac{\partial_x^i p(z;x,u)}{p(z;x,u)}\tilde{C}_i(0,u,y)\Big)p(z;x,u)\,dz, \end{aligned}$$

and hence have shown the following proposition.

Proposition 2.8. *For every bounded and measurable function f and $n \geq 1$, the push-down Malliavin weight satisfies*

$$\mathbb{E}^{x,u}[\pi_n \,|\, X_T^0 = z] = \sum_{i=0}^{m(n)} \frac{\partial_x^i p(z;x,u)}{p(z;x,u)}\tilde{C}_i(0,u,y).$$

Proof. For $f \in C_b^\infty$ the claim follows from the above and for general f bounded and measurable from a monotone class argument. □

2.3 Expansion of Implied Volatilities

As noticed above (cf. Lemma 2.20), the function at first order $v_0(t,x,u)$ corresponds to the Black & Scholes price in terms of the log-price. In fact, recalling that on $(X_t^0, u_t^0) = (x,u)$, X_T^0 is normally distributed with mean and variance given by $x - \frac{1}{2}\int_t^T u(s-t)\,ds$ and $\int_t^T u(s-t)\,ds$ respectively, we can switch to the usual Black & Scholes parameterization $v_{BS}(x, \sqrt{V(t,u)})$ in terms of the initial log-price x and volatility $\sqrt{V(t,u)}$ where

$$V(t,u) := \frac{1}{T-t}\int_t^T u(s-t)\,ds.$$

We will now consider an implied volatility expansion for (2.1) by proceeding similarly to [69] where this was done for finite dimensional diffusions. We take f to be the payoff of a call option with strike K (the call-option payoff is not bounded but the put payoff is so we can proceed with the put-call parity) and recall that the implied volatility $\sigma^\epsilon := \sigma^\epsilon(T,K)$ is defined by the condition

$$v_{BS}(x,\sigma^\epsilon) = v^\epsilon(0,(x,u,y)). \tag{2.30}$$

Lemma 2.9. *If $V(t,u) > 0$, then the mapping $\epsilon \mapsto \sigma^\epsilon$ is smooth.*

Proof. Let $\psi(\sigma) := v_{BS}(x,\sigma)$. It is known that $\sigma \mapsto \psi(\sigma)$ is smooth and that $\psi'(\sigma) > 0$ (i.e. the vega of an call-option is strictly positive). Hence ψ has a continuously differentiable inverse. Moreover, it follows from (2.30) that $\epsilon \mapsto \psi(\sigma^\epsilon)$ is smooth. Hence $\epsilon \mapsto \sigma^\epsilon = \psi^{-1} \circ \psi(\sigma^\epsilon)$ is continuously differentiable with representation $\frac{\partial}{\partial \epsilon}\sigma^\epsilon = \frac{\frac{\partial}{\partial \epsilon}\psi(\sigma^\epsilon)}{\psi'(\sigma^\epsilon)}$ from which the higher order differentiability follows. □

Hence for any $n \geq 0$ we have

$$\sigma^\epsilon = \sum_{i=0}^{n} \epsilon^i \sigma_i + o(\epsilon^n), \qquad \text{as } \epsilon \to 0,$$

where $\sigma_i := \frac{1}{i!}\frac{\partial^i}{\partial \epsilon^i}\Big|_{\epsilon=0}\sigma^\epsilon$. On the other hand, it follows from (2.30) and Proposition 2.6 that

$$\frac{\partial^i}{\partial \epsilon^i}\Big|_{\epsilon=0} v_{BS}(x,\sigma^\epsilon) = v_i(0,(x,u,y))$$

for $i \geq 0$. Thus by using the Faà di Bruno formula and matching terms explicit representations for σ_i, $i \geq 0$, can be found. We illustrate this for first three terms. It is evident that $\sigma_0 = \sqrt{V(0,u)}$ and from $\frac{\partial}{\partial \epsilon}\Big|_{\epsilon=0} v_{BS}(x,\sigma^\epsilon) = \partial_\sigma v_{BS}(x,\sigma_0)\sigma_1$ and

$$\frac{\partial^2}{\partial \epsilon^2}\Big|_{\epsilon=0} v_{BS}(x,\sigma^\epsilon) = \partial_\sigma^2 v_{BS}(x,\sigma_0)\sigma_1 + \partial_\sigma v_{BS}(x,\sigma_0)\sigma_2$$

we find that

$$\sigma_1 = \frac{v_1(0,x,u,y)}{\partial_\sigma v_{BS}(x,\sigma_0)} \qquad \text{and} \qquad \sigma_2 = \frac{v_2(0,x,u,y) - \partial_\sigma^2 v_{BS}(x,\sigma_0)\sigma_1}{\partial_\sigma v_{BS}(x,\sigma_0)}, \tag{2.31}$$

where v_1 and v_2 are given in (2.25) and (2.27) respectively. We give now a more explicit representation for σ^1. Recalling the Black & Scholes formula in terms of the log-price

$$v_{BS}(x,\sigma_0) = e^x N(d_1(x,\sigma)) - KN(d_2(x,\sigma)),$$

where $N(z) = \int_{-\infty}^{z} n(r)\,dr$, $n(r) := \frac{1}{\sqrt{2\pi}} e^{-\frac{r^2}{2}}$ and

$$d_1(x,\sigma) = \frac{\log(e^x/K)}{\sigma\sqrt{T}} + \frac{1}{2}\sigma\sqrt{T}, \quad d_1(x,\sigma) = \frac{\log(e^x/K)}{\sigma\sqrt{T}} - \frac{1}{2}\sigma\sqrt{T},$$

we see that $D_\sigma v_{BS}(x,\sigma) = e^x\sqrt{T} n(d_1(x,\sigma))$ and

$$D_x D_\sigma v_{BS}(x,\sigma) = D_\sigma v_{BS}(x,\sigma) - \frac{D_\sigma v_{BS}(x,\sigma) d_1(x,\sigma)}{\sigma\sqrt{T}}.$$

Letting $C'_{1,1}(0,(u,y)) \in H$ such that $C_{1,1}(0,(u,y)) = \int_0^T C'_{1,1}(0,(u,y))(s)\,ds$ (see (2.26)) we can apply (2.20) to (2.25) to arrive at

$$\begin{aligned} v_1(0,x,u,y) &= D_x D_u v_{BS}(x, \sqrt{V(0,u)}) C'_{1,1}(0,(u,y)) \\ &= D_x D_\sigma v_{BS}(x, \sqrt{V(0,u)}) \frac{V(0, C'_{1,1}(0,(u,y)))}{2\sigma_0}. \end{aligned}$$

Putting all together and plugging into (2.31) gives

$$\begin{aligned} \sigma_1 &= \frac{1}{2} \frac{(D_x^3 - D_x^2) v_{BS}(x,\sigma_0) C_{1,1}(0,(u,y))}{D_\sigma v_{BS}(x,\sigma_0)} \\ &= \Big(\frac{1}{2} - \frac{\log(e^x/K)}{\sigma_0^2 T}\Big) \frac{V(0, C'_{1,1}(0,(u,y)))}{2\sigma_0}, \end{aligned} \tag{2.32}$$

where as above $\sigma_0 = \sqrt{V(0,u)}$.

2.4 Termstructure of ATM volatility skew

According to [5] the termstructure of at-the-money volatility skew defined as

$$\psi(T) := \Big| \frac{\partial}{\partial_k}\Big|_{k=0} \sigma_{T,k} \Big|,$$

where $k = log(e^x/K)$, *is a feature of the volatility surface that really does distinguish between (stochastic volatility) models.* They found empirically (cf. [5, Figure 1.2]) that $T \mapsto \psi(T)$ should be proportional to $1/T^\gamma$ with $\gamma \in (0, 1/2)$ (see Section III.6 for more on this topic). Recalling the representation of σ_1 in (2.32) we see that at first order

$$\psi(T) = \Big| \frac{V(0, C'_{1,1}(0,(u,y)))}{2\sigma_0^3 T} \Big| = \Big| \frac{\int_0^T \int_s^T \sigma_1(u(s-t), y)(r-s)\sqrt{u(s-t)}\,ds\,dr}{2\sqrt{T}(\int_0^T u(s)\,ds)^{3/2}} \Big|.$$

For a flat initial curve $u(x) = c$ for all $x \geq 0$ we accordingly have

$$\psi(T) = \Big| \frac{\int_0^T \int_s^T \sigma_1(c,y)(r-s)\,dr\,ds}{2T^2 c} \Big|.$$

Notice that in the rBergomi model (see Section III.6) we have $\sigma_1(u,Y)(r) = u\frac{1}{r^\gamma}$ and hence (formally, as we do not know whether this expansion holds for the rBergomi model) for $u = c$

$$\psi_{\text{rBergomi}}(T) = \Big| \frac{\int_0^T \int_0^{T-s} \frac{1}{r^\gamma}\,dr\,ds}{2T^2} \Big| = \Big| \frac{1}{2(2-\gamma)(1-\gamma)} \Big| T^{-\gamma}$$

agreeing with the empirically observed form.

2.5 Example I: CRC Affine Realizations

In the case of the affine CRC models given in (V.2.5) we have $\sigma_1(u, y) = \phi_1(u)e^{b_1(y)\cdot}$ and accordingly

$$\psi_{\text{affCRC}}(T) = \left| \frac{\int_0^T \phi(c) \frac{e^{b_1(y)(T-s)}-1}{b_1(y)}\, ds}{2T^2 c} \right|.$$

2.6 Example II: CRC Exponentially Affine Realizations

For the exponentially affine CRC models given in (V.2.10) we have

$$\sigma_1(u, Y) = m(u, e^{b_1(Y)\cdot}),$$

which gives for the constant initial curve $u = c$ essentially the same ATM volatility skew as for the affine CRC models, the only difference is that here $\phi(c) = c$.

3 Term Structure Expansion

We saw in the last section that with the combination of (condition **SWTA 1** of) Theorem IV.3.2 and the Kolmogorov equation we could find an asymptotic expansion around the unperturbed (i.e. $\epsilon = 0$) state. This is very convenient since in the unperturbed state the first two coordinates of the joint system $Z = (X, u, Y)$ admitted a particular simple kind of a finite dimensional realization in the sense that the forward variance reduced to the solution of the deterministic transport equation. Accordingly the log-price process X reduced to a 1-dimensional (time-inhomogeneous) Markov process that in particular corresponded to the time-dependent Black & Scholes model for which closed-form solutions for many derivatives on the log-pice exist. In this section we will continue with this example by looking at derivatives of the forward variance curve u (instead of the log-price X) by looking at the second and third coordinate of (2.1). By doing so, we can use the same system for pricing and calibration (using this asymptotic expansion) of derivatives on the log-price X (like plain vanilla call options) and on the forward variance u (like call options on the VIX).

3.1 Expansion of Price

We look at the parameterized forward variance model with stochastic volatility (which includes the CRC processes) given by

$$\begin{cases} du_t^\epsilon &= \frac{d}{dx} u_t^\epsilon\, dt + \epsilon \sum_{i=1}^d \sigma^i(u_t^\epsilon, Y_t^\epsilon)\, d\beta_t^i, \quad u_0^\epsilon = u \in H, \\ dY_t^\epsilon &= \epsilon^2\, c_0(Y_t^\epsilon)\, dt + \epsilon \sum_{i=1}^d c_i(Y_t^\epsilon)\, d\beta_t^i, \quad Y_0^\epsilon = y \in \mathbb{R}^m, \end{cases} \tag{3.1}$$

and assume that the joint process (u, Y) in $\mathcal{H} := H \times \mathbb{R}^m$ satisfies the conditions of Theorem IV.3.2 for a certain linear map $l \in L(\mathcal{H}, \mathbb{R})$ satisfying $l(u, Y) = \tilde{l}(u)$ for some linear map in $L(H, \mathbb{R})$. Then as above for a smooth function $\tilde{f} \in C_b^\infty(\mathbb{R}; \mathbb{R})$ we consider the composite map $f := \tilde{f} \circ l$. Finally we assume that for $\epsilon = 0$ and each $y \in \mathbb{R}^m$ the first coordinate of (3.1) admits a strong solution for every $u_0 \in D(d/dx)$ which is for example satisfied for the CRC processes. Then under the condition **SWTA 1** of Theorem IV.3.2 the joint process $Z^\epsilon = (u^\epsilon, Y^\epsilon)$ on $\mathcal{H}$ satisfies for each $\epsilon \in U$ the conditions of [26, Theorem 9.17] and accordingly the Kolmogorov equation

$$\begin{cases} \partial_t v^\epsilon(t,(u,y)) + D_u v^\epsilon(t,(u,y))(\frac{d}{dx}u) \\ +\epsilon^2 D_y v^\epsilon(t,(u,y))c_0(y) + \epsilon^2 \frac{1}{2}\sum_{j=1}^d D_y^2 v^\epsilon(t,(u,y))c_j^2(y) \\ +\epsilon^2\,\frac{1}{2}\sum_{j=1}^d D_u^2 v^\epsilon(t,(u,y))(\sigma_j(u,y),\sigma_j(u,y)) \\ +\epsilon^2\sum_{i=1}^d D_{u,y}^2 v^\epsilon(t,(u,y))(\sigma_i(u,y),c_i(y)) = 0 \\ v^\epsilon(T,(u,y)) = f(u), \end{cases} \tag{3.2}$$

has a unique strict solution for each $\epsilon \in U$ given by

$$v^\epsilon(t,(u,y)) = \mathbb{E}[f(u_T^\epsilon)\,|\,(u_t^\epsilon, Y_t^\epsilon) = (u,y)],$$

satisfying the corresponding properties of (2.5) and Lemma 2.1. By aggregating as above the operators in (3.2) we can write

$$\begin{cases} (\partial_t + L^\epsilon)v^\epsilon(t,(u,y)) = 0 \\ v^\epsilon(T,(u,y)) = f(u), \end{cases}$$

where $L^\epsilon := L^0 + \epsilon^2 L^2$ are such that for any $v \in C_b^2(\mathcal{H})$ and letting $z = (u,y)$

$$L^0 v(z) \;:=\; D_u v(z)(\frac{d}{dx}u) \tag{3.3}$$

$$\begin{aligned} L^2 v(z) \;:=\;& D_y v(z)c_0(y) + \frac{1}{2}\sum_{j=1}^d D_y^2 v(z)c_j^2(y) \\ +\;& \frac{1}{2}\sum_{j=1}^d D_u^2 v(z)(\sigma_j(u,y),\sigma_j(u,y)) + \sum_{i=1}^d D_{u,y}^2 v(z)(\sigma_i(u,y),c_i(y)). \end{aligned} \tag{3.4}$$

And from the corresponding version of Lemma 2.2 we see that the functions in (1.2) satisfy the PDEs given by

$$(\partial_t + L^0)v_0 \;=\; 0, \quad v_0(T,(u,y)) = f(u) \tag{3.5}$$

$$(\partial_t + L^0)v_n + L^2 v_{n-2} \;=\; 0, \quad v_n(T,(u,y)) = 0, \quad n \geq 2. \tag{3.6}$$

The price at order 0 is given by

$$v_0(t,u) = f(S_{T-t}u), \tag{3.7}$$

as for $\epsilon = 0$ and $u_t^0 = u$, u_T^0 corresponds to the solution at time T of

$$\begin{cases} du_s^0 = \frac{d}{dx}u_s^0, \quad s \geq t \\ u_t^0 = u. \end{cases}$$

The price at 1-st order is defined as

$$\begin{aligned} v_1(t,u,y) \;:=\;& \frac{\partial}{\partial\epsilon}\Big|_{\epsilon=0}\mathbb{E}[f(u_T^\epsilon)\,|\,(u_t^\epsilon, Y_t^\epsilon) = (u,y)] \\ =\;& \frac{\partial}{\partial\epsilon}\Big|_{\epsilon=0}\mathbb{E}\Big[f(S_{T-t}u + \epsilon\sum_{i=1}^d \int_t^T S_{T-s}\sigma_i(u_s^\epsilon, Y_s^\epsilon)\,d\beta_s^i)\Big] \end{aligned}$$

and be recalling that $f := \tilde{f} \circ l$ we see that

$$\frac{\partial}{\partial \epsilon}\Big|_{\epsilon=0} \mathbb{E}\Big[f(S_{T-t}u + \epsilon \sum_{i=1}^{d} \int_t^T S_{T-s}\sigma_i(u_s^\epsilon, Y_s^\epsilon)\, d\beta_s^i)\Big]$$

$$= \tilde{f}'(l(S_{T-t}u)) \sum_{i=1}^{d} \mathbb{E}\Big[\int_t^T l\Big(S_{T-s}\sigma_i(S_{s-t}u, y)\Big)\, d\beta_s^i\Big] = 0,$$

due to the properties of the stochastic integral. Hence we have shown that

$$v_1(t, u, y) = 0. \tag{3.8}$$

For the prices at higher order, we have as in (2.18) that

$$v_n(t, u, y) = \int_t^T L^2 v_{n-2}(s, S_{s-t}u, y)\, ds, \tag{3.9}$$

which follows now from the fundamental theorem of calculus and (3.6) upon noting that for a smooth v we have $\partial_t v(t, u_t^0) = \partial_t v(t, u_t^0) + D_u v(t, u_t^0) \frac{d}{dx} u_t^0$ where on the left hand side of the equation ∂_t acts on $t \to v(t, u_t^0)$ and on the right-hand side on the first argument, i.e. $t \mapsto v(t, u)$. We also recall from the above that $u_s^0 = S_{s-t}u$ and that accordingly

$$v_0(s, u_s^0, y) = f(S_{T-s}u_s^0) = f(S_{T-s}S_{s-t}u) = f(S_{T-t}u), \tag{3.10}$$

which is a property reminiscent to the martingality property stated in Lemma 2.3. Now using (3.9) and (3.10) we can compute the higher order prices. We see from (3.8) that the price at every odd order is zero. In the following we compute the prices at second and fourth order. At second order we have

$$v_2(t, u, y) = \int_t^T L^2 v_0(s, u_s^0)\, ds = \frac{1}{2} \sum_{j=1}^{d} \int_t^T D_u^2 f(S_{T-s}u_s^0)(\sigma_j(u_s^0, y), \sigma_j(u_s^0, y))\, ds$$

$$= \tilde{f}''(l(S_{T-t}u)) \frac{1}{2} \sum_{j=1}^{d} \int_t^T l(S_{T-s}\sigma_j(S_{s-t}u, y))^2\, ds =: \tilde{f}''(l(S_{T-t}u)) C_1(t, u, y)$$

and similarly at fourth order

$$\begin{aligned} v_4(t, u, y) &= \int_t^T L^2 v_2(s, u_s^0, y)\, ds \\ &= \tilde{f}''(l(S_{T-t}u)) \int_t^T \Big(D_y C_1(s, u_s^0, y) c_0(y) + \frac{1}{2} \sum_{j=1}^{d} D_y^2 C_1(s, u_s^0, y) c_j^2(y)\Big)\, ds \\ &+ \frac{1}{2} \sum_{j=1}^{d} \int_t^T \Big(D_u^2 \tilde{f}''(l(S_{T-s}u_s^0)) C_1(s, u_s^0, y)\Big)(\sigma_j(u_s^0, y), \sigma_j(u_s^0, y))\, ds \\ &+ \sum_{i=1}^{d} \int_t^T D_{u,y} \tilde{f}''(l(S_{T-s}u_s^0)) C_1(s, u_s^0, y)(\sigma_i(u_s^0, y), c_i(y))\, ds \\ &=: \tilde{f}''(l(S_{T-t}u)) C_2(t, u, y) + I + II, \end{aligned}$$

where I satisfies

$$\begin{aligned}
I &= \frac{1}{2}\sum_{j=1}^{d}\int_t^T D_u^2\Big(\tilde{f}''(l(S_{T-s}u_s^0))C_1(s,u_s^0,y)\Big)(\sigma_j(u_s^0,y),\sigma_j(u_s^0,y))\,ds \\
&= \frac{1}{2}\sum_{j=1}^{d}\int_t^T \Big(D_u^2\tilde{f}''(l(S_{T-s}u_s^0))(\sigma_j(u_s^0,y),\sigma_j(u_s^0,y))\Big)C_1(s,u_s^0,y)\,ds \\
&+ \sum_{j=1}^{d}\int_t^T \Big(D_u\tilde{f}''(l(S_{T-s}u_s^0))\sigma_j(u_s^0,y)\Big)\Big(D_uC_1(s,u_s^0,y)\sigma_j(u_s^0,y)\Big)\,ds \\
&+ \frac{1}{2}\sum_{j=1}^{d}\int_t^T \tilde{f}''(l(S_{T-s}u_s^0))\Big(D_u^2C_1(s,u_s^0,y)(\sigma_j(u_s^0,y),\sigma_j(u_s^0,y))\Big)\,ds \\
&= \tilde{f}^{(4)}(l(S_{T-t}u))\frac{1}{2}\sum_{j=1}^{d}\int_t^T l(S_{T-s}\sigma_j(S_{s-t}u,y))^2C_1(s,u_s^0,y)\,ds \\
&+ \tilde{f}^{(3)}(l(S_{T-t}u))\sum_{j=1}^{d}\int_t^T l(S_{T-s}(\sigma_j(u_s^0,y))\Big(D_uC_1(s,u_s^0,y)\sigma_j(u_s^0,y)\Big)\,ds \\
&+ \tilde{f}''(l(S_{T-t}u))\frac{1}{2}\sum_{j=1}^{d}\int_t^T \Big(D_u^2C_1(s,u_s^0,y)(\sigma_j(u_s^0,y),\sigma_j(u_s^0,y))\Big)\,ds \\
&=: \tilde{f}^{(4)}(l(S_{T-t}u))C_3(t,u,y)+\tilde{f}'''(l(S_{T-t}u))C_4(t,u,y)+\tilde{f}''(l(S_{T-t}u))C_5(t,u,y)
\end{aligned}$$

and for II we have

$$\begin{aligned}
II &= \sum_{i=1}^{d}\int_t^T D_{u,y}\Big(\tilde{f}''(l(S_{T-s}u_s^0))C_1(s,u_s^0,y)\Big)(\sigma_i(u_s^0,y),c_i(y))\,ds \\
&= \sum_{i=1}^{d}\int_t^T \Big(D_u\tilde{f}''(l(S_{T-s}u_s^0))\sigma_i(u_s^0,y)\Big)\Big(D_yC_1(s,u_s^0,y)c_i(y)\Big)\,ds \\
&+ \sum_{i=1}^{d}\int_t^T \tilde{f}''(l(S_{T-s}u_s^0))D_{u,y}C_1(s,u_s^0,y)(\sigma_i(u_s^0,y),c_i(y))\,ds \\
&= \tilde{f}'''(l(S_{T-t}u))\sum_{i=1}^{d}\int_t^T l(S_{T-s}\sigma_i(u_s^0,y))D_yC_1(s,u_s^0,y)c_i(y)\,ds \\
&+ \tilde{f}''(l(S_{T-t}u))\sum_{i=1}^{d}\int_t^T D_{u,y}C_1(s,u_s^0,y)(\sigma_i(u_s^0,y),c_i(y))\,ds \\
&=: \tilde{f}'''(l(S_{T-t}u))C_6(t,u,y)+\tilde{f}''(l(S_{T-t}u))C_7(t,u,y)
\end{aligned}$$

and hence putting all together

$$\begin{aligned}
v_4(t,u,y) &= \tilde{f}''(l(S_{T-t}u))(C_2(t,u,y)+C_5(t,u,y)+C_7(t,u,y)) \\
&= \tilde{f}'''(l(S_{T-t}u))(C_4(t,u,y)+C_6(t,u,y))+\tilde{f}^{(4)}(l(S_{T-t}u))C_3(t,u,y).
\end{aligned}$$

More generally we can show as in Proposition 2.6 the following.

Proposition 3.1. *The function* $v_n(t,(u,y))$, *for every odd number* $n \geq 1$, $v_n(t,(u,y)) = 0$ *and for every even number* $n \geq 2$

$$v_n(t,(u,y)) = \sum_{i=0}^{m(n)} \tilde{f}^{(i)}(l(S_{T-t}u))\tilde{C}_i(t,u,y),$$

where $n \mapsto m(n)$ *is an increasing sequence of natural numbers and* $\tilde{C}_i$ *are smooth, deterministic functions.*

Proof. The claim holds as shown above for $n = 1, 2, 3, 4$ and the general case follows from induction. □

3.2 VIX Options

As an example of derivatives on the forward variance we can consider options on the VIX, which is given for time T by

$$\mathrm{VIX}_T := \sqrt{\frac{1}{x}\int_0^x u_T^\epsilon(z)\,dz}, \tag{3.11}$$

where x denotes time-to-maturity and is about 20 business days. Hence for a smooth payoff functions $\hat{f}$ we consider the pricing function

$$\mathbb{E}[\hat{f}(\mathrm{VIX}_T)] = \mathbb{E}[f(u_T^\epsilon)] = \mathbb{E}[\tilde{f}(l(u_T^\epsilon))], \tag{3.12}$$

where f, $\tilde{f}$ and l correspond to the notations used above, i.e. l is the linear mapping in $L(H;(R))$ given by $l(h) = \frac{1}{x}\int_0^x h(z)\,dz$ and $\tilde{f}$ is the real-valued function on $\mathbb{R}_+$ given by $\tilde{f}(x) := \hat{f}(\sqrt{x})$ and $f := \tilde{f} \circ l$. Now under the conditions of the previous subsection the expansion given in Proposition 3.1 can be used.

Appendix A

Stochastic Processes in Hilbert Spaces

We give a short overview on stochastic partial differential equations. The main reference here is [26], but also results from [27], [25], [32] and [45] are used. Throughout we let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P})$ denote a complete filtered probability space satisfying the usual conditions, that is, $\mathcal{F}$ is $\mathbb{P}$-complete, $\mathcal{F}_0$ contains all $\mathbb{P}$-nullsets, and the filtration $(\mathcal{F})_{t\in\mathbb{R}_+}$ is right continuous. We also assume that we are given a d-dimensional Brownian motion $\beta = (\beta^1, ..., \beta^d)$ relative to the probability space. Also we denote by H a separable Hilbert space and by $\mathcal{B}(H)$ the *Borel* σ-field on H. Further we denote by $\mathcal{P}$ the *predictable* σ-field on $\mathbb{R}_+ \times \Omega$ and by $\mathcal{P}_T$ its restriction to $[0, T] \times \Omega$ (cf. [26, page 76]).

1 The Itô Integral

We follow the construction given in [26, Chapter 4] and [32, Chapter 2] for the stochastic integral wit values in H, however as we consider finite dimensional driving noise (given by the d-dimensional Brownian motion β) the procedure is very similar to the well-known case of the stochastic integral taking values in $\mathbb{R}^n$, see for example [57]. Let $T \in \mathbb{R}_+$ be arbitrary but fixed. A H-valued (local) martingale is defined exactly as in the real-valued case (see [26]) such that also the following property holds.

Proposition 1.1. *The space $\mathcal{M}^2_T(H)$ of H-valued continuous martingales M on $[0, T]$ with the norm*

$$\|M\|^2_{\mathcal{M}^2_T(H)} := \mathbb{E}\Big[\sup_{t\in[0,T]} \|M(t)\|^2_H\Big]$$

is a Hilbert space. The closed subspace $\mathcal{M}^{0,2}_T(H)$ consists of those martingales $M \in \mathcal{M}^2_T(H)$ with $M_0 = 0$.

Proof. This is [26, Proposition 3.9]. □

Definition 1.2. *We call $\mathcal{L}^2_T(H)$ the Hilbert space of equivalence classes of H-valued predictable processes Φ with norm*

$$\|\Phi\|^2_{\mathcal{L}^2_T(H)} := \mathbb{E}\Big[\int_0^T \|\Phi_t\|^2_H \, dt\Big].$$

Now denote by $\mathcal{E}_T(H)$ the subset of $\mathcal{L}_T^2(H)$ consisting of *elementary processes* Φ, i.e. there exists a sequence $0 = t_0 < t_1 < ... < t_k = T$ and a sequence of random variables $\Phi_0, \Phi_1, ..., \Phi_{k-1}$ such that Φ_m is $\mathcal{F}_{t_m}$ measurable and

$$\Phi(t) = \Phi_m, \quad \text{for } t \in (t_m, t_{m+1}], \quad m = 0, 1, ..., k-1.$$

For $\Phi \in \mathcal{E}_T(H)$ the *stochastic integral* with respect to a real-valued Brownian motion β^1 is defined by

$$(\Phi \cdot \beta^1)_t := \sum_{m=0}^{k-1} \Phi_m(\beta^1_{t_{m+1}\wedge t} - \beta^1_{t_m \wedge t}), \quad t \in [0, T].$$

Proposition 1.3. *The subset of elementary processes $\mathcal{E}_T(H)$ is dense in $\mathcal{L}_T^2(H)$, and the map from $\mathcal{E}_T(H)$ into $\mathcal{M}_T^{0,2}(H)$ given by $\Phi \mapsto \Phi \cdot \beta^1$ is an isometry.*

Proof. This [26, Proposition 4.5 and 4.7] □

Definition 1.4. *The unique extension of the isometry to the map from $\mathcal{L}_T^2(H)$ into $\mathcal{M}_T^{0,2}(H)$ will be called the stochastic integral of Φ with respect to β^1. It will also be denoted by*

$$(\Phi \cdot \beta^1)_t = \int_0^t \Phi_s \, d\beta^1_s.$$

Definition 1.5. *We call $\mathcal{L}_T^{loc}(H)$ the space of equivalence classes of H-valued predictable processes Φ such that*

$$\mathbb{P}\Big[\int_0^T \|\Phi_t\|_H^2 \, dt < \infty\Big] = 1.$$

Proposition 1.6. *Let $\Phi \in \mathcal{L}_T^{loc}(H)$. Then there exists a unique H-valued continuous local martingale M on $[0, T]$ that is characterized by*

$$M_{t\wedge\tau} = ((\Phi 1_{[0,\tau]}) \cdot \beta^1)_t$$

whenever $\Phi 1_{[0,\tau]} \in \mathcal{L}_T^2(H)$. Again, M is called the stochastic integral of Φ with respect to β^1 and it is written

$$M_t =: (\Phi \cdot \beta^1)_t = \int_0^t \Phi_s \, d\beta^1_s.$$

Proof. This is [32, Proposition 2.2.3]. □

Finally we can consider the spaces $\mathcal{L}^2(H) := \mathcal{L}^2_\infty(H)$ and $\mathcal{L}^{loc} := \bigcap_{T\in\mathbb{R}_+} \mathcal{L}_T^{loc}$ in the obvious way, see [32, page 19].

2 Itô's Formula

An H-valued continuous adapted process X is called an Itô process if it is of the form

$$X_t = X_0 + \int_0^t b_s \, ds + \sum_{i=1}^{d} \int_0^t \Phi^i_s \, d\beta^i_s, \tag{2.1}$$

where $\Phi^i \in \mathcal{L}^{loc}(H)$, for $i = 1, ..., d$ and b is an H-valued predictable process such that

$$\mathbb{P}\Big[\int_0^T \|\Phi_t\|_H \, dt < \infty\Big] = 1$$

and X_0 is $\mathcal{F}_0$ measurable.

Proposition 2.1. *Let X be an Itô process as in* (2.1) *and let $F \in C_b^{1,2}([0,T] \times H; E)$ where E is another Hilbert space. Then $t \mapsto F(t, X_t)$ is an E-valued Itô process with representation*

$$\begin{aligned} F(t, X_t) &= F(0, X_0) + \sum_{i=1}^d \int_0^t D_x F(s, X_s)(\Phi_s^i) \, d\beta_s^i \\ &+ \int_0^t \Big(D_s F(s, X_s) + D_x F(s, X_s)(b_s) + \frac{1}{2}\sum_{i=1}^d D_{xx} F(s, X_s)(\Phi_s^i, \Phi_s^i)\Big) \, ds \end{aligned}$$

Proof. This follows from [26, Theorem 4.17] and [32, Theorem 2.3.1]. □

3 Stochastic Equations

Here we will look to the stochastic equation in H given by

$$\begin{cases} dX_t &= (AX_t + F(t, X_t)) \, dt + \sum_{j=1}^d B^j(t, X_t) \, d\beta_t^j \\ X_0 &= h_0, \end{cases} \tag{3.1}$$

where A is the infinitesimal generator of a strongly continuous semigroup $\{S_t \,|\, t \in \mathbb{R}_+\}$ and F, B^j, $j = 1, ..., d$ are mappings from $(\mathbb{R}_+ \times \Omega \times H, \mathcal{P} \otimes \mathcal{B}(H))$ into $(H, \mathcal{B}(H))$. In the following we will follow the exposition given in [32, Section 2.4] for the different solution concepts.

Definition 3.1. *Suppose that X is an H-valued predictable process and $\tau > 0$ a stopping time satisfying*

$$\mathbb{P}\Big[\int_0^{t\wedge\tau} \big(\|X_s\|_H + \|F(s, X_s)\|_H + \sum_{j=1}^d \|B^j(s, X_s)\|_H^2\big) \, ds < \infty\Big] = 1,$$

for all $t \in \mathbb{R}_+$. We call X

- *a local mild solution to* (3.1)*, if the following holds*

$$\begin{aligned} X_t &= S_{t\wedge\tau} h_0 + \int_0^{t\wedge\tau} S_{(t\wedge\tau)-s} F(s, X_s) \, ds \\ &+ \sum_{j=1}^d \int_0^{t\wedge\tau} S_{(t\wedge\tau)-s} B^j(s, X_s) \, d\beta_s^j, \quad \mathbb{P} - a.s., \forall t \in \mathbb{R}_+. \end{aligned}$$

- *a local weak solution to* (3.1)*, if for arbitrary $\zeta \in D(A^*)$*

$$\begin{aligned} \langle \zeta, X_t \rangle_H &= \langle \zeta, h_0 \rangle_H + \int_0^{t\wedge\tau} \big(\langle A^*\zeta, X_s \rangle_H + \langle \zeta, F(s, X_s) \rangle_H\big) \, ds \\ &= \sum_{j=1}^d \int_0^{t\wedge\tau} \langle \zeta, B^j(s, X_s) \rangle_H \, d\beta_s^j, \quad \mathbb{P} - a.s., \forall t \in \mathbb{R}_+. \end{aligned}$$

- *a local strong solution to* (3.1), *if* $X \in D(A)$, $dt \otimes d\mathbb{P}$*-a.s.*

$$\mathbb{P}\Big[\int_0^{t\wedge\tau} \|AX_s\|_H \, ds < \infty\Big] = 1, \quad \forall t \in \mathbb{R}_+$$

and the integral version of (3.1)

$$X_t = X_0 + \int_0^{t\wedge\tau} (AX_s + F(s, X_s))\, ds + \sum_{j=1}^{d} \int_0^{t\wedge\tau} B^j(s, X_s)\, d\beta_s^j$$

holds $\mathbb{P}$*-a.s.* $\forall t \in \mathbb{R}_+$.

We call τ *the lifetime of* X, *if* $\tau = \infty$ *the solutions are just called mild, weak and strong respectively.*

As the names suggest, a (local) strong solution is also a (local) weak solution which in turn is also a (local) mild solution. On the other hand, it follows from [26, Theorem 6.5] that if $B^i \in \mathcal{L}^2(H)$, for $i = 1, ..., d$, then a (local) mild solution is also a (local) weak solution.

Theorem 3.2. *Let* $r \in [0, T]$ *and assume that the mappings* $F, B_1..., B_d : [r, T] \times \Omega \times H \to H$ *satisfy the conditions*

SI1 *The mappings* $F, B_1, ..., B_d : [r, T] \times \Omega \times H \to H$ *are measurable from* $((\Omega \times [r, T]) \times H, \mathcal{P}_T \times \mathcal{B}(H))$ *into* $(H, \mathcal{B}(H))$

SI2 *There exists a constant* $C > 0$ *such that*

$$\begin{cases} \|F(t,\omega;x) - F(t,\omega;y)\|_H + \sum_{i=1}^d \|B_i(t,\omega;x) - B_i(t,\omega;y)\|_H \\ \le C\|x-y\|_H, \quad \text{for all } x, y \in H, t \in [r, T], \omega \in \Omega, \end{cases}$$

$$\begin{cases} \|F(t,\omega;x)\|_H^2 + \sum_{i=1}^d \|B_i(t,\omega;x)\|_H^2 \le C^2(1 + \|x\|_H^2), \\ \text{for all } x \in H, t \in [r, T], \omega \in \Omega. \end{cases}$$

Then the equation for $t \ge r$

$$\begin{cases} dX_t &= AX_t + F(t, X_t)\, dt + \sum_{i=1}^d B_i(t, X_t)\, d\beta_t^i \\ X_r &= \xi. \end{cases} \tag{3.2}$$

with ξ *being* $\mathcal{F}_r$ *measurable has a unique continuous mild solution*

$$X_t = S_{t-r}\xi + \int_r^t S_{t-s}F(s, X_s)\, ds + \sum_{i=1}^d \int_r^t S_{t-s}B_i(s, X_s)\, d\beta_s^i,$$

that satisfies for $p \ge 2$,

$$\sup_{t\in[r,T]} E[\|X_t\|_H^p] \le C_{p,T}(1 + E[\|\xi\|_h^p]).$$

and for $p > 2$

$$E[\sup_{t\in[r,T]} \|X_t\|_H^p] \le \tilde{C}_{p,T}(1 + E[\|\xi\|_H^p]).$$

The solution is the limit of the Picard Iteration scheme, defined recursively by

$$\begin{cases} X_t^0 & := S_{t-r}\xi \\ X_t^{k+1} & := S_{t-r}\xi + \int_r^t S_{t-s}F(s, X_s^k)\, ds + \sum_{i=1}^d \int_r^t S_{t-s}B_i(s, X_s^k)\, d\beta_s^i, \quad k \geq 1, \end{cases} \tag{3.3}$$

in the Banach space $\mathcal{H}_p$, $p \geq 2$, *of* H*-valued predictable processes* Y *defined on the time interval* $[r, T]$ *such that*

$$\|Y\|_{\mathcal{H}_p} = \Big(\sup_{t \in [r,T]} E[\|Y_t\|_H^p] \Big)^{1/p} < \infty.$$

Proof. This is [26, Theorem 7.4]. □

In [32, Corollary 2.4.1] there is also a corresponding version of this Theorem with local conditions yielding a local weak solution.

Bibliography

[1] Ralph H Abraham, Tudor S Ratiu, and Jerrold E Marsden, *Manifolds, tensor analysis and applications*, (1983).

[2] Yacine Ait-Sahalia, Mustafa Karaman, and Loriano Mancini, *The term structure of variance swaps and risk premia*, Available at SSRN 2136820 (2014).

[3] Peter Allen, Stephen Einchcomb, and Nicolas Granger, *Variance swaps*, London: JP Morgan (November) (2006).

[4] Fabrice Baudoin, Josef Teichmann, et al., *Hypoellipticity in infinite dimensions and an application in interest rate theory*, The Annals of Applied Probability **15** (2005), no. 3, 1765–1777.

[5] Christian Bayer, Peter K. Friz, and Jim Gatheral, *Pricing under rough volatility*, Available at SSRN: http://ssrn.com/abstract=2554754 (2015).

[6] Christian Bayer, Jim Gatheral, and Morten Karlsmark, *Fast ninomiya–victoir calibration of the double-mean-reverting model*, Quantitative Finance **13** (2013), no. 11, 1813–1829.

[7] Lorenzo Bergomi, *Smile dynamics ii*, Available at SSRN 1493302 (2005).

[8] Lorenzo Bergomi and Julien Guyon, *The smile in stochastic volatility models*, Available at SSRN 1967470 (2011).

[9] ______, *Derivatives pricing-stochastic volatility's orderly smiles*, Risk-London **25** (2012), no. 5, 60.

[10] Tomas Björk, *On the geometry of interest rate models*, Springer, 2004.

[11] Tomas Björk and Bent Jesper Christensen, *Interest rate dynamics and consistent forward rate curves*, Mathematical Finance **9** (1999), no. 4, 323–348.

[12] Tomas Björk, Camilla Landén, and Lars Svensson, *Finite-dimensional markovian realizations for stochastic volatility forward–rate models*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, vol. 460, The Royal Society, 2004, pp. 53–83.

[13] Tomas Björk and Lars Svensson, *On the existence of finite-dimensional realizations for nonlinear forward rate models.*, Mathematical Finance (2001), 11: 205243.

[14] Fischer Black and Myron Scholes, *The pricing of options and corporate liabilities*, The journal of political economy (1973), 637–654.

[15] Hans Bühler, *Volatility markets: Consistent modeling, hedging, and practical implementation of variance swap market models*, VDM Verlag Dr. Müller, 2009.

[16] Philippe Carmona and Laure Coutin, *Fractional brownian motion and the markov property*, Université de Toulouse. Laboratoire de Statistique et Probabilités [LSP], 1998.

[17] Philippe Carmona, Laure Coutin, and Gérard Montseny, *Approximation of some gaussian processes*, Statistical inference for stochastic processes **3** (2000), no. 1-2, 161–171.

[18] René Carmona and Michael Tehranchi, *A characterization of hedging portfolios for interest rate contingent claims*, Annals of Applied Probability (2004), 1267–1294.

[19] ______, *Interest rate models: an infinite dimensional stochastic analysis perspective*, Springer Science & Business Media, 2007.

[20] Peter Carr and Roger Lee, *Volatility derivatives*, Annu. Rev. Financ. Econ. **1** (2009), no. 1, 319–339.

[21] ______, *Hedging variance options on continuous semimartingales*, Finance and Stochastics **14** (2010), no. 2, 179–207.

[22] Peter Carr and Dilip Madan, *Towards a theory of volatility trading*, Option Pricing, Interest Rates and Risk Management, Handbooks in Mathematical Finance (2001), 458–476.

[23] Patrick Cattiaux and Laurent Mesnager, *Hypoelliptic non-homogeneous diffusions*, Probability Theory and Related Fields **123** (2002), no. 4, 453–483.

[24] Fabio Comte and Eric Renault, *Long memory continuous time models*, Journal of Econometrics **73** (1996), no. 1, 101–149.

[25] Giuseppe Da Prato, *Kolmogorov equations for stochastic pdes*, Springer Science & Business Media, 2004.

[26] Giuseppe Da Prato and Jerzy Zabczyk, *Stochastic equations in infinite dimensions*, Cambridge University Press, 1992.

[27] ______, *Second order partial differential equations in hilbert spaces.*, Cambridge University Press, Vol.293, 2002.

[28] Freddy Delbaen and Walter Schachermayer, *A general version of the fundamental theorem of asset pricing*, Mathematische Annalen **300** (1994), no. 1, 463–520.

[29] Philipp Dörsek and Josef Teichmann, *Efficient simulation and calibration of general hjm models by splitting schemes.*, SIAM Journal on Financial Mathematics 4.1 (2013).

[30] Darrel Duffie, Damir Filipović, and Walter Schachermayer, *Affine processes and applications in finance*, The Annals of Applied Probability, Vol. 13, No. 3, 984-1053 (2003).

[31] Daniel Egloff, Markus Leippold, and Liuren Wu, *The term structure of variance swap rates and optimal variance swap investments*, (2010).

[32] Damir Filipović, *Consistency problems for heath-jarrow-morton interest rate models*, Springer, Lecture Notes in Mathematics, 1760, 2001.

[33] ______, *Time-inhomogeneous affine processes*, Stochastic Processes and their Applications **115** (2005), no. 4, 639–659.

[34] ______, *Term-structure models, a graduate course*, Springer Finance, 2009.

[35] ______, *Affine variance swap curve models*, Seminar on Stochastic Analysis, Random Fields and Applications VII, Progress in Probability, Springer Basel (2012).

[36] Damir Filipović, Elise Gourier, and Loriano Mancinia, *Quadratic variance swap models*, Journal of Financial Economics, Volume 119, Issue 1, January 2016, Pages 44 – 68.

[37] Damir Filipović, Stefan Tappe, and Josef Teichmann, *Invariant manifolds with boundary for jump-diffusions*, Electron. J. Probab **19** (2014), no. 51, 1–28.

[38] Damir Filipović and Josef Teichmann, *Existence of invariant manifolds for stochastic equations in infinite dimension*, Journal of Functional Analysis 197.2, 398-432. (2003).

[39] ______, *Regularity of finite-dimensional realizations for evolution equations*, Journal of Functional Analysis **197** (2003), no. 2, 433–446.

[40] ______, *On the geometry of the term structure of interest rates*, Proceedings of the Royal Society London A 460, 129-167 (2004).

[41] Patrick Florchinger, *Malliavin calculus with time dependent coefficients and application to nonlinear filtering*, Probability theory and related fields **86** (1990), no. 2, 203–223.

[42] Eric Fournié, Jean-Michel Lasry, Jérôme Lebuchoux, and Pierre-Louis Lions, *Applications of malliavin calculus to monte-carlo methods in finance. ii*, Finance and Stochastics **5** (2001), no. 2, 201–236.

[43] Jim Gatheral, *Consistent modeling of spx and vix options*, Bachelier Congress, 2008.

[44] Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum, *Volatility is rough*, Available at SSRN 2509457 (2014).

[45] Leszek Gawarecki and Vidyadhar Mandrekar, *Stochastic differential equations in infinite dimensions: with applications to stochastic partial differential equations*, Springer, 2010.

[46] Axel Grorud and E Pardoux, *Intégrales hilbertiennes anticipantes par rapport à un processus de wiener cylindrique et calcul stochastique associé*, Applied Mathematics and Optimization **25** (1992), no. 1, 31–49.

[47] Martin Hairer, *Hypoellipticity in infinite dimensions*, Proceedings of 7th ISAAC Congress. Imperial College London, World Scientific, 2009.

[48] Eskil Hansen and Alexander Ostermann, *Exponential splitting for unbounded operators*, Mathematics of computation **78** (2009), no. 267, 1485–1496.

[49] Philipp Harms, David Stefanovits, et al., *Affine representations of fractional processes with applications in mathematical finance*, arXiv preprint arXiv:1510.04061 (2015).

[50] Philipp Harms, David Stefanovits, Josef Teichmann, and Mario Wüthrich, *Consistent recalibration of yield curve models*, arXiv preprint arXiv:1502.02926 (2015).

[51] David Heath, Robert Jarrow, and Andrew Morton, *Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation*, Econometrica, Vol. 60, No. 1 (Jan., 1992), pp. 77-105 (1992).

[52] Steven L Heston, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, Review of financial studies **6** (1993), no. 2, 327–343.

[53] Jean Jacod and Albert N. Shiryaev, *Limit theorems for stochastic processes, 2nd edtion*, Springer, 2002.

[54] Robert Jarrow, Younes Kchia, Martin Larsson, and Philip Protter, *Discretely sampled variance and volatility swaps versus their continuous approximations.*, Finance and Stochastics 17.2 (2013), 305–324.

[55] Jan Kallsen, *A didactic note on affine stochastic volatility models*, in Yu. Kabanov, R. Liptser and J. Stoyanov (eds.), From Stochastic Calculus to Mathematical Finance, (Springer, Berlin, 2006), pp. 343-368 (2006).

[56] Jan Kallsen and Paul Krühner, *On a heath-jarrow-morton approach for stock options.*, arXiv preprint arXiv:1305.5621 (2013).

[57] Ioannis Karatzas and Steven Shreve, *Brownian motion and stochastic calculus, second edition*, Springer, Graduate Texts in Mathematics, 1991.

[58] Raphael Kruse, *Strong and weak approximation of semilinear stochastic evolution equations*, Springer, 2014.

[59] Benoit B Mandelbrot and John W Van Ness, *Fractional brownian motions, fractional noises and applications*, SIAM review **10** (1968), no. 4, 422–437.

[60] Domenico Marinucci and Peter M Robinson, *Alternative forms of fractional brownian motion*, Journal of statistical planning and inference **80** (1999), no. 1, 111–122.

[61] Scott Mixon and Esen Onur, *Volatility derivatives in practice: activity and impact*, Available at SSRN (2015).

[62] Marek Musiela, *Stochastic pdes and term structure models*, Journées Internationales de France, IGR- AFFI, La Baule, 1993 (1993).

[63] David Nualart, *The malliavin calculus and related topics*, vol. 1995, Springer, 2006.

[64] Matthijs Pronk and Mark Veraar, *Tools for malliavin calculus in umd banach spaces*, Potential Analysis **40** (2014), no. 4, 307–344.

[65] Daniel Revuz and Marc Yor, *Continuous martingales and brownian motion*, Springer, A Series of Comprehensive Studies in Mathematics, Volume 293, 2005.

[66] Anja Richter and Josef Teichmann, *Discrete time term structure theory*, Working Paper, 2013.

[67] Evelina Shamarova, *A version of the hörmander–malliavin theorem in 2-smooth banach spaces*, Infinite Dimensional Analysis, Quantum Probability and Related Topics **17** (2014), no. 01, 1450004.

[68] Maria Siopacha and Josef Teichmann, *Weak and strong taylor methods for numerical solutions of stochastic differential equations*, Quantitative Finance **11** (2011), no. 4, 517–528.

[69] Akihiko Takahashi and Toshihiro Yamada, *An asymptotic expansion with push-down of malliavin weights*, SIAM Journal on Financial Mathematics **3** (2012), no. 1, 95–136.

[70] ———, *A remark on approximation of the solutions to partial differential equations in finance*, Available at SSRN 2007584 (2012).

[71] Stefan Tappe, *An alternative approach on the existence of affine realizations for hjm term structure models*, Proc. R. Soc. A, 2010.

[72] ———, *Affine realizations with affine state processes for stochastic partial differential equations*, http://www.stochastik.uni-hannover.de/fileadmin/institut/pdf/Affine-state-proc.pdf (2015).

[73] Josef Teichmann, *A frobenius theorem on convenient manifolds*, Monatshefte für Mathematik **134** (2001), no. 2, 159–167.

[74] ———, *A note on nonaffine solutions of term structure equations with applications to power exchanges*, Mathematical Finance **15** (2005), no. 1, 191–201.