J. Math. Anal. Appl. 473 (2019) 534–566

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

On explicit local solutions of Itô diffusions [☆]

Michael A. Kouritzin [a,*], Bruno Rémillard [b]

[a] Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta, T6G 2G1 Canada

[b] Service de l'enseignement des méthodes quantitatives de gestion, HEC Montréal, Montréal, Québec, H3T 2A7 Canada

ARTICLE INFO

Article history:
Received 15 May 2017
Available online 5 January 2019
Submitted by U. Stadtmueller

Keywords:
Diffeomorphism
Itô processes
Explicit solutions

ABSTRACT

Strong solutions of p-dimensional stochastic differential equations $dX_t = b(X_t, t)dt + \sigma(X_t, t)dW_t$, $X_s = x$ that can be represented locally in *explicit simulation* form $X_t = \phi^{x,s}\left(\int_s^t V_{s,u} dW_u, t\right)$ are considered. Here; W is a multidimensional Brownian motion; $u \to V_{s,u}$, $\phi^{x,s}$ are continuous functions; and $b, \sigma, \phi^{x,s}$ are locally continuously differentiable. The following three-way equivalence is established: 1) There exists such a representation from all starting points (x, s), 2) $V_{s,u}$, $\phi^{x,s}$ satisfies a set differential equations, and 3) b, σ satisfy commutation relations. (For generality, the function $V_{s,t}$ is allowed to depend upon $\phi^{x,s}$ via $V_{s,t} = U_{s,t}\phi^{x,s}$ for some operators $U_{s,t}$.) Moreover, construction theorems, based on a diffeomorphism between the solutions X and the strong solutions to a simpler Itô integral equation, with a possible deterministic component, are given. Finally, motivating examples are provided and its importance in simulation methods, including sequential Monte Carlo, financial risk assessment and path-dependent option pricing, is explained.

1. Introduction

One often confines selection of stochastic differential equation (SDE) models to those facilitating calculation and simulation. For example, the popularity of the inaccurate Black–Scholes model is only justifiable through the evaluation ease of the resulting derivative product formulae. Indeed, Kunita [18, p. 272] writes in his notes on SDEs that "It is an important problem in applications that we can compute the output from the input explicitly". We shall call such solutions *explicit solutions*.

Doss [5] and Sussmann [21] were apparently the first to solve stochastic differential equations through use of differential equations. In the multidimensional setting, Doss imposed the Abelian condition on the

☆ Partial funding in support of this work was provided by the Natural Sciences and Engineering Research Council of Canada.

* Corresponding author.
E-mail addresses: michaelk@ualberta.ca (M.A. Kouritzin), bruno.remillard@hec.ca (B. Rémillard).
URLs: http://www.math.ualberta.ca/Kouritzin_M.html (M.A. Kouritzin), http://www.hec.ca/pages/bruno.remillard (B. Rémillard).

https://doi.org/10.1016/j.jmaa.2018.12.067
0022-247X/

Lie algebra generated by the vector fields of coefficients and showed, in this case, that strong solutions, X_t^x, of Fisk–Stratonovich equations are representable as $X_t^x = \rho(\Phi(x, W_\cdot)_t, W_t)$, for some continuous ρ, Φ solving differential equations. Under the restriction of C^∞ coefficients, Yamato [22] extended the work of Doss by dispensing with the Abelian assumption in favor of less restrictive q step nilpotency, whilst also introducing a simpler form for his explicit solutions $X_t^x = u(x, t, (W_t^I)_{I\in F})$. Here, u solves a differential equation, and $(W_t^I)_{I\in F}$ are iterated Stratonovich integrals with integrands and integrators selected from $\left(t, W_t^1, ..., W_t^d\right)$. Another substantial work on explicit solutions to stochastic differential equations is due to Kunita [18, Section III.3]. He considers representing solutions to time-homogeneous Fisk–Stratonovich equations via flows generated by the coefficients of the equation under a commutative condition, and, more generally, under solvability of the underlying Lie algebra. Kunita's work therefore generalizes Yamato [22].

There is related, more recent work on simulating stochastic differential equations through stochastic Taylor's theorem, exponential Lie series and sinh-log series. These methods employ iterated stochastic integrals and/or ordinary differential equation (ODE) approximation over small time. One can learn more about these methods from e.g. Ben Arous [1], Castell [3], Hu [9], Kloeden and Platen [11], Castell and Gaines [4] and Malham and Wiese [19]. These methods are general in the sense that they usually do not require commutator conditions between the coefficient vector fields. Still, significant coefficient smoothness is often required and it is usually found that the computational costs associated with numerically solving the ODEs or iteratively integrating are greater than direct use of Euler or Milstein methods on the SDE of interest. As our interest stems from computationally intensive applications, we turn our attention to less-general, computationally-efficient methods.

Our representations do not employ stochastic integrals (even non-iterated ones) nor ODEs in the manner mentioned above and consequently can facilitate efficient simulation compared to Euler and Milstein methods. A typical use is the following (Explicit Simulation Algorithm):

(1) Simulate a Gauss–Markov process, which will be denoted Y_t herein.
(2) Use ϕ to map to a desired process X_t, where ϕ is some average of the $\phi^{x,s}$ used herein.
(3) Possibly project down to a weak solution of a lower dimensional SDE.
(4) Possibly use importance sampling to create a weak solution to yet another SDE with different drift.

At each successive step the number and complexities of the SDEs that can be handled increases.

Example 1. We summarize a current use of the Explicit Simulation Algorithm with results from this paper in simulation based option pricing, financial risk assessment and sequential Monte Carlo.

Heston [8] introduced a stochastic volatility model with closed form European-call-option prices for stock, bond and foreign currency spot prices. Let B, β to be (scalar) independent standard Brownian motions. Then, the Heston model is:

$$d\begin{pmatrix} S_t \\ V_t \end{pmatrix} = \begin{pmatrix} \mu S_t \\ \nu - \varrho V_t \end{pmatrix} dt + \begin{pmatrix} \sqrt{1-\rho^2} S_t V_t^{\frac{1}{2}} & \rho S_t V_t^{\frac{1}{2}} \\ 0 & \kappa V_t^{\frac{1}{2}} \end{pmatrix} \begin{pmatrix} dB_t \\ d\beta_t \end{pmatrix}, \tag{1.1}$$

with parameters $\mu \in \mathbb{R}$, $\rho \in [-1, 1]$ and $\nu, \varrho, \kappa > 0$. S_t, the price part, is a stochastic exponential but the exponent involves $\int_0^t V_s^{\frac{1}{2}} d\beta_s$ with β and V being dependent so stochastic integral approximations would appear needed. The volatility component is just the Cox–Ingersoll–Ross (CIR) model. The diffusion vector fields *do not commute* i.e. $[\sigma_1, \sigma_2] \doteq (\nabla\sigma_1)\sigma_2 - (\nabla\sigma_2)\sigma_1 \neq 0$ so we can not obtain explicit strong solutions of the desired form but the Explicit Simulation Algorithm still works. Kouritzin [14] used Theorem 2 below in steps (1), (2) above to show that the extended Heston price model, consisting of S_t above along with a collection of Ornstein–Uhlenbeck processes, has an explicit strong solution of the form considered here under Condition (C) of Kouritzin [14]. From there, an explicit weak solution for the Heston price was obtained by projection as in step (3) above. Finally, Condition (C) was dispensed with using Girsanov's theorem as in step (4). Suppose $\varepsilon > 0$. Kouritzin [14] shows that the Heston (price and volatility) model (1.1) has explicit weak solution:

$$S_t = S_0 \exp\left(\sqrt{1-\rho^2}\int_0^t V_s^{\frac{1}{2}} dB_s + \left[\mu - \frac{\nu\rho}{\kappa}\right]t + \left[\frac{\rho\varrho}{\kappa} - \frac{1}{2}\right]\int_0^t V_s ds + \frac{\rho}{\kappa}(V_t - V_0)\right)$$

$$V_t = \sum_{i=1}^n (Y_t^i)^2, \quad \eta_\varepsilon = \inf\{t : V_t \le \varepsilon\} \quad \text{and}$$

$$L_t = \exp\left\{\frac{\nu - \nu_\kappa}{\kappa^2}\left[\ln(V_t) - \ln(V_0) + \int_0^t \frac{\kappa^2 - \nu_\kappa - \nu}{2V_s} + \varrho\, ds\right]\right\}$$

up until η_ε with respect to new probability measure

$$\widehat{P}(A) = E[1_A L_{T\wedge\eta_\varepsilon}] \ \ \forall A \in \mathcal{F}_T,$$

where $\nu_\kappa = \frac{n\kappa^2}{4}$ and $\{Y_t^i\}_{i=1}^n$ are Ornstein–Uhlenbeck processes. It is important to note: V and B are independent so $\int_0^t V_s^{\frac{1}{2}} dB_s$ is conditionally Gaussian and there is *no need to approximate stochastic integrals.* (This Condition (C) in Kouritzin [14] would make $L_t \equiv 1$ and $\widehat{P} = P$.) American and Asian options were then priced efficiently using these explicit formulas and Monte Carlo simulations. Kouritzin and MacKay [17] also use the Explicit Simulation Algorithm based upon work herein to produce explicit weak solutions to a generalized Bates model (with jumps), where the adjective *generalized* is used because there is an extra drift term in the price equation (that arises for certain insurance product prices). Further, they assess insurer's risk in Guaranteed Minimum Withdrawl Benefit insurance using Monte Carlo simulations with these explicit solutions. In current work, Kouritzin and MacKay [16] use (branching particle) sequential Monte Carlo to improve performance of path-dependent option pricing. Kouritzin [13] and the results (Theorem 2 and Example 5) herein are used to show that the Heston model yields a second weak solution with the formula for S_t unchanged but

$$dV_t = \left(\frac{\kappa^2}{4} + \kappa\chi\sqrt{V_t} - \varrho V_t\right)dt + \kappa\sqrt{V_t}d\widehat{\beta}_t$$

$$L_t = \exp\left\{\frac{\nu - \frac{\kappa^2}{4}}{\kappa^2}\left[\ln\left(\frac{\widehat{V}_t}{V_0}\right) + \int_0^t \frac{\frac{3\kappa^2}{8} - \frac{\nu}{2}}{\widehat{V}_s} ds + \varrho t\right] - \chi\widehat{\beta}_t - \frac{\chi^2}{2}t\right\}$$

with respect to the new probability measure $\widehat{P}(A) = E[1_A L_{T\wedge\eta_\varepsilon}] \ \ \forall A \in \mathcal{F}_T$ up until the time $\eta_\varepsilon = \inf\{t : V_t \le \varepsilon\}$ that the volatility dips too low. (We presented the time-homogeneous-coefficient case for V here for simplicity. The more general case is given in Example 5 below.) If one simulates multiple independent copies $\{(S^i, V^i, L^i)\}_{i=1}^N$ with either Heston representation in this example, then one finds that the weighted empirical measures of the path processes converge a.s. to the process distribution of the Heston model

$$\frac{1}{N}\sum_{i=1}^N L_T^i \delta_{S^i_{[0,T]}, V^i_{[0,T]}}(s_{[0,T]}, v_{[0,T]}) \to P^{Heston}(s_{[0,T]}, v_{[0,T]}),$$

where $(s_{[0,T]}, v_{[0,T]})$ solve (1.1) with respect to P^{Heston}, and path-dependent option pricing can be done (with the celebrated LS algorithm). However, if the option is over any significant time period, then the weights L_T^i will diverge without some type of (unbiased) resampling, branching or interaction to level them.

Kouritzin and MacKay [16] use branching particle sequential Monte Carlo with these explicit solutions to keep the weights relatively equal and all particles effective.

The first two steps of the Explicit Simulation Algorithm involve classification of which Itô processes $X_t^{x,s}$, starting at (x,s), are representable as a time-dependent function of a simple stochastic integral $\phi^{x,s}\left(\int_s^t V_{s,u}dW_u, t\right)$, which was initially motivated by filtering applications (see Kouritzin [12]). Our determination of $\phi^{x,s}$, $V_{s,u}$ also facilitates an effective means of calculation and simulation in other applications. To simulate, one merely needs to compute the Gauss–Markov process $\int_s^t V_{s,u}dW_u$ at discrete times and substitute these samples into $\phi^{x,s}$, which is often known in closed form and otherwise is the solution of differential equations that can be solved numerically a priori. $\int_s^t V_{s,u}dW_u = \int_s^t V_{s,u}(X_u)dW_u$ can depend upon X but not in a way that will destroy its Gaussian distribution nor make simulation difficult.

We require commutator conditions for (step (2) and) our explicit strong solutions herein, which is a significant restriction. However, (i) the drift vector field need not strictly commute with the diffusion vector fields, (ii) non-commuting diffusion vector fields can sometimes be handled (in a weak sense) by considering a higher dimensional SDE (see the Explicit Simulation Algorithm), (iii) importance sampling methods can be used to handle more non-commuting drift vector fields. Actually, it is already known that one can have explicit solutions under commutator conditions and our work is quite related to the earlier works of Yamato [22] and Kunita [18]. However, compared to these early works, our work features time-dependent coefficients and a different representation that is very useful in simulation. We compare our results to Yamato [22] and Kunita [18] in Section 4.

Our explicit simulation is without (Euler or Milstein) bias and is often orders of magnitude faster than Euler or Milstein methods when applicable and high accuracy is desired (see Kouritzin [14]). Our representations also make properties of certain stochastic differential equations readily discernible. Finally, as demonstrated in Karatzas and Shreve [10, Proposition 5.2.24], explicit solutions can be useful in establishing convergence for solutions of stochastic differential equations.

In order to describe our method, we recall the state-space diffeomorphism mapping method has been used to construct solutions to interesting stochastic differential equations from solutions to simpler ones. The idea of this method is to change the infinitesimal generator L of a simple Itô process to the generator $\mathcal{L}$ corresponding to a more complicated Itô process via $\mathcal{L}f(x) = \{L(f \circ \Lambda^{-1})\} \circ \Lambda(x)$. This corresponds to using Itô's formula on $X_t = \Lambda^{-1}(\xi_t)$, where ξ is a diffusion process with infinitesimal generator L. For related examples, we refer the reader to the problems in Friedman [7, page 126] or Ethier and Kurtz [6, page 303].

Motivated by applications in filtering, Kouritzin and Li [15] and Kouritzin [13] used differential equation methods to study: "When can global, time-dependent diffeomorphisms be used to construct solutions to Itô equations?", "What scalar Itô equations can be solved via diffeomorphisms?", and "How can one construct these diffeomorphisms?". They considered scalar solutions in an open interval D to the time-homogeneous stochastic differential equation

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t,\ X_0 = x, \tag{1.2}$$

which are of the form $\phi^x\left(\int_0^t V_u dW_u, t\right)$, and showed that all nonsingular solutions of this form were actually (time-dependent) diffeomorphisms $\Lambda_t^{-1}(\xi_t)$ with ξ satisfying

$$d\xi_t = (\chi - \kappa\xi_t)dt + dW_t,\ \xi_0 = \Lambda_0(x).$$

A nonsingular solution in this scalar case was interpreted as finiteness of $\int_\lambda^y \sigma^{-1}(x)dx$ for some fixed point λ and all $y \in D$. (Their methods involve non-stochastic differential equations that can continue to hold in the singular situations when global diffeomorphisms fail.)

For our current work, we suppose henceforth that $D \subset \mathbb{R}^p$ is a bounded convex domain, $T > 0$, and define

$$D_T = \begin{cases} D & \text{if } \sigma, b \text{ do not depend on } t \\ D \times [0,T) & \text{if either do} \end{cases}$$

so $(x,s) \in D_T$ means $x \in D$ when $D_T = D$. We also let $\sigma(X_t)dW_t$ imply Itô integration and $\sigma(X_t) \bullet dW_t$ Stratonovich. Then, we resolve the question: "When can we explicitly solve vector-valued Itô equations

$$dX_t = b(X_t,t)dt + \sigma(X_t,t)dW_t,\ X_s = x, \tag{1.3}$$

with the dimensions of X_t, W_t being p, d respectively, through representations of the form $X_t^{x,s} = \phi^{x,s}\left(\int_s^t V_{s,u}dW_u, t\right)$?". This question is more precisely broken into two separate important questions: "For which σ and b does such a strong-local-solution representation exists?" and "What conditions are required on ϕ and V for such representations with $\int_s^t V_{s,u}dW_u = \int_s^t V_{s,u}(X_u)dW_u$ still being Gauss–Markov?" Equivalently, we consider "When can the solutions to the Fisk–Stratonovich equation

$$dX_t^x = h(X_t^x,t)dt + \sigma(X_t^x,t) \bullet dW_t, \tag{1.4}$$

with

$$h = b - \frac{1}{2}\sum_{j=1}^{d}\{\nabla_\varphi \sigma_j\}\sigma_j \text{ on } D_T \tag{1.5}$$

and σ_j denoting the jth column of the matrix σ, be locally represented in this manner?" It follows from, for example, Kunita [18, p. 239] that the unique local solutions to (1.3) and (1.4) are equal if (1.5) holds and σ is twice continuously differentiable or satisfies the Fisk–Stratonovich acceptable condition in D, the latter being discussed in Protter [20, Chapter 5]. We work with Itô equations to avoid these stronger assumptions on σ but still relate b and h through (1.5). Also, to obtain simple, concrete necessary and sufficient conditions for such a representation, we consider all solutions starting from each $(x,s) \in D_T$. Under natural regularity conditions, we answer these question by showing the equivalence of the following three conditions: 1) The SDEs (1.3) have our local-solution-representations for all starting points $(x,s) \in D_T$. 2) The representation pair $\phi^{x,s}, V_{s,t}$ satisfy a system of differential equations. 3) The SDE coefficients σ and h satisfy simple commutator conditions. In the process of establishing this three-way equivalence, we also answer the question "When is (1.3) locally diffeomorphic to an SDE with a simple diffusion coefficient?" i.e. "When will it have a representation as in (1.6), (1.7) to follow?".

Given precise conditions of when an Itô equation has such a representation, the next natural questions we answer are: "What form do the solutions have?" and "How do you construct such solutions?" In order to include as many interesting examples as possible we will only require *local* representation $X_t^{x,s} = \phi^{x,s}\left(\int_s^t V_{s,u}dW_u, t\right)$ and allow σ to have rank less than $\min(p,d)$. By allowing the rank of $\sigma(x)$ to be less than p one can handle time-dependent coefficients, treating time as an extra state. The second advantage from allowing lesser rank than $\min(p,d)$ is the extra richness afforded by appending a deterministic equation into the diffeomorphism solution. A third, important benefit of this general rank condition is the possibility of producing explicit *weak* solutions to SDEs where no explicit strong solution exists (see Kouritzin [14]). In our construction results, we show that ϕ is constructed via a time-dependent diffeomorphism Λ_t, which in turn is defined in terms of σ. The diffeomorphism separates a representable SDEs into deterministic and stochastic differential equations: $\Lambda_t(X_t) = (\overline{X}_t, \widetilde{X}_t)$, where $\widetilde{X}_t \in \mathbb{R}^{p-r}$ is deterministic and satisfies the differential equation

$$\frac{d}{dt}\widetilde{X}_t = \widetilde{h}(\widetilde{X}_t, t), \tag{1.6}$$

while $\overline{X}_t$ is a Gauss–Markov process satisfying

$$d\overline{X}_t = (\overline{\theta}(\widetilde{X}_t, t) + \overline{\beta}(\widetilde{X}_t, t)\overline{X}_t)dt + \left(I_r \middle| \, \overline{\kappa}(\widetilde{X}_t, t)\right) dW_t. \tag{1.7}$$

$\overline{\kappa}$ is determined (within an equivalence class) by σ while $\overline{\theta}$, $\widetilde{h}$ and $\overline{\beta}$ can be anything (subject to dimensional and differentiability regularity conditions). These parameters allow us to handle a whole class of *nonlinear* drift coefficients b for a given σ in the SDE (1.3) for $X_t = \Lambda_t^{-1}(\overline{X}_t, \widetilde{X}_t)$.

In the next section, we introduce notation and state the main existence results. In Section 3, we build off of these existence results to give our construction results, illustrated with simple applications. We compare our work to prior work of Yamato and Kunita in Section 4. The proofs of all main results are postponed to Section 5.

2. Notation and existence results

Let $(W_t)_{t\geq 0}$ be a standard d-dimensional Brownian motion with respect to filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual hypotheses on a complete probability space $(\Omega, \mathcal{F}, P)$. We will use $\phi^{x,s}$ to denote a representation function and x to denote a starting point as in the introduction. On the other hand, φ will denote a variable with the same dimension p as $\phi^{x,s}$ and x.

For functions of time or paths of a stochastic process, we use Z_t and $Z(t)$ interchangeably. For a matrix V, V_j will denote its jth column vector and $V_{i,j}$ the ith element of this jth column. 0_j (0^j) means a row (column) of j 0's.

$B_z(\delta)$ denotes an open Euclidean ball centered at z with radius $\delta > 0$. Suppose $m, r \in \mathbb{N}$, $O \subset \mathbb{R}^m$ is open and $I \subset [0, T)$ is an interval. Then, $C(I)$ is the continuous functions on I and $C^r(O)$ denotes the continuous functions whose partial derivatives up to order r exist and are continuous on O. Moreover, $C^{r,1}(O \times I)$ denotes the continuous functions $g(\varphi, t)$ whose mixed partial derivatives in $\varphi \in O$ up to order r and in $t \in I$ up to order 1 all exist and are continuous functions on $O \times I$. $C^1(O \times I) = C^{1,0}(O \times I) \cap C^{0,1}(O \times I)$. (We only require one-sided derivatives in time to exist at interval endpoints.) For a vector function g of both $\varphi \in \mathbb{R}^p$ and t, $\nabla_\varphi g$ is the Jacobian matrix of g, that is $(\nabla_\varphi g)_{i,j} = \partial_{\varphi_j} g_i$, while ∇g will include the time derivative as the last column. (A similar notation will be used for vector functions of $y \in \mathbb{R}^d$ and t.) $f \circ g$ will denote the composition of functions $f \circ g(x) = f(g(x))$.

The purpose of our representations is to simulate a class of processes in an efficient manner, which leads to a dilemma. We would like to allow $V_{s,t}$ to depend upon $X^{x,s}$ for generality but not in a way that would destroy the ease of simulation. Our approach to this dilemma is to allow $V_{s,t} = U_{s,t}\phi^{x,s}(y_\cdot, \cdot)$ to be defined by operators $U_{s,t}$ on the functions $\phi^{x,s}(y_u, u)\big|_{u\in[s,t]}$ but then impose the condition that the result $U_{s,t}\phi^{x,s}(y_\cdot, \cdot)$ can not depend upon y. As we will expose below, this basically allows $V_{s,t}$ to depend upon some hidden deterministic part of X but not the purely stochastic part, saving the Gaussian nature of

$$Y_t^s = \int_s^t U_{s,u}\phi^{x,s}(Y_\cdot^s, \cdot)dW(u) = \int_s^t U_{s,u}\phi^{x,s}(0, \cdot)dW(u) \tag{2.1}$$

so it can be computed off-line, which is the point of this work.

$\phi^{x,s}$ must be differentiable enough to apply Itô's formula and allow room for random process Y_t^s to move. For fixed s, t and path $y \in C([s, t]; \mathbb{R}^d)$, $U_{s,t}$ is a mapping $C([s, t]; \mathbb{R}^p) \to C([s, t]; \mathbb{R}^{d\times d})$. ($U_{s,t}$ will be forced to be constant in y. Hence, when we apply $U_{s,t}$ to $\phi^{x,s}(\cdot, \cdot)$ below we are effectively applying it to $\phi^{x,s}(0, \cdot)$.) Further constraints on $t \to U_{s,t}\phi^{x,s}$, in particular the imposition of a group structure, will be set in Conditions $\mathcal{C}_2$, $\mathcal{C}_3$ below while the role of $U_{s,t}$ in preserving the Guassian character of Y_t^s will become clearer in Example 4. The precise regularity conditions for potential representations $X_t^{x,s} = \phi^{x,s}(Y_t^s, t)$, $Y_t^s = \int_s^t U_{s,u}\phi^{x,s}(0,\cdot)dW(u)$ follow:

$\mathcal{C}_1$: For each $(x,s) \in D_T$, there is a $t_0 = t_0^{x,s} > s$ and a convex neighborhood $\mathcal{N}^{x,s} \subset \mathbb{R}^d$ of 0 such that $\phi^{x,s} \in C^{2,1}(\mathcal{N}^{x,s} \times [s,t_0); \mathbb{R}^p)$ and $t \to U_{s,t}\phi^{x,s}(y.,\cdot) \in C^1([s,t_0); \mathbb{R}^{d\times d})$.

$\mathcal{C}_2$: $\phi^{x,s}, U_{s,t}$ start correctly

$$\phi^{x,s}(0,s) = x, \quad U_{s,s}\phi^{x,s}(0,s) = I_d, \ \forall (x,s) \in D_T. \tag{2.2}$$

$\mathcal{C}_3$: $U_{s,t}\phi^{x,s}$ is non-singular on $\mathcal{N}^{x,s} \times [s,t_0)$ (with matrix inverse denoted by $U_{s,t}^{-1}\phi^{x,s}$) and satisfies

$$U_{s,t}\phi^{x,s}(y,u) = U_{s,t}\phi^{x,s}(0,u) \tag{2.3}$$

as well as

$$U_{s,t}^{-1}\phi^{x,s}(y_t,t)\frac{d}{dt}U_{s,t}\phi^{x,s}(y_u,u)\big|_{u=t} = \frac{d}{dt}U_{u,t}\phi^{\phi^{x,s}(y_u,u),u}(y_u,u)\big|_{u=t}. \tag{2.4}$$

The purpose of the first part of $\mathcal{C}_3$, (2.3), is to preserve the Gaussian nature of X (while still allowing $V_{s,t}$ to depend on X in some way) as discussed above. The role of the second part of $\mathcal{C}_3$, (2.4), is to force a type of (semi-)group structure on $U_{s,t}\phi^{x,s}$. Combined, $\mathcal{C}_3$ will allow our representation function $t \to \phi^{x,s}(\cdot,t)$ to contain a deterministic, dynamic portion of X.

(2.2), (2.4) imply

$$U_{s,t}^{-1}\phi^{x,s}(y_t,t)\frac{d}{dt}U_{s,t}\phi^{x,s}(y_u,u)\big|_{u=t} = U_{t,t}^{-1}\phi^{\phi^{x,s}(y_t,t),t}\frac{d}{dt}U_{u,t}\phi^{\phi^{x,s}(y_u,u),u}\big|_{u=t} \tag{2.5}$$

and therefore that $U_{s,t}\phi^{x,s}$ is a (two parameter) group. We use (2.3) to economize the notation $U_{s,t}\phi^{x,s}(y.,\cdot)$ to $U_{s,t}\phi^{x,s}$.

Now, define the $\mathcal{F}_t$-stopping time

$$\tau^{x,s} = \min\left(t_0^{x,s}, \inf\{t > s : \ Y_t^s \notin \mathcal{N}^{x,s} \text{ or } (\phi^{x,s}(Y_t^s,t),t) \notin D_T\}\right)$$

and let

$$\mathcal{R}^{x,s} = \bigcup_{t\geq 0}\left\{(y,t) : P((Y_t^s,t) \in B_{(y,t)}(\delta), t \leq \tau^{x,s}) > 0, \ \forall \delta > 0\right\}. \tag{2.6}$$

There is structure that can be imposed upon $\phi^{x,s}, U_{s,t}$ that will turn out to be equivalent to the existence of our explicit strong local solutions.

Definition 1. An (x,s,σ,h)-representation is a pair $\phi^{x,s}, U_{s,t}$ satisfying $(\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3)$ such that the following system of differential equations:

$$\nabla_y\phi^{x,s}(y,t) = \sigma(\phi^{x,s}(y,t),t)U_{s,t}^{-1}\phi^{x,s}, \tag{2.7}$$

$$\partial_t\phi^{x,s}(y,t) = h(\phi^{x,s}(y,t),t) \tag{2.8}$$

hold for all $(y,t) \in \mathcal{R}^{x,s}$ and $\partial_s\nabla_y\phi^{x,s}(0,s)$, $\partial_s\partial_t\phi^{x,s}(0,s)$, $\partial_{x_i}\nabla_y\phi^{x,s}(0,s)$ and $\partial_{x_i}\partial_t\phi^{x,s}(0,s)$ exist as continuous functions of $(x,s) \in D_T$. Here and below, $\partial_t\phi^{x,s}(0,s)$ means $\partial_t\phi^{x,s}(0,t)\big|_{t=s}$.

Notice that $U_{s,t}$ only appears as (the matrix inverse of) $V_{s,t} = U_{s,t}\phi^{x,s}$ so we will only be concerned with solving for $U_{s,t}\phi^{x,s}$ for the $\phi^{x,s}$ of interest.

Now, our explicit solutions are:

$$X_t^{x,s} = \phi^{x,s}(Y_t^s, t) \text{ on } [s, \tau^{x,s}). \tag{2.9}$$

Our first main result establishes two necessary and sufficient conditions for all $X^{x,s}$, defined in (2.9), to be strong local solutions to

$$dX_t = b(X_t, t)dt + \sigma(X_t, t)dW_t, \qquad X_s = x \tag{2.10}$$

on $[s, \tau^{x,s})$. The function h is always related to b through (1.5) and $U_{s,t}\phi^{x,s}$ comes into the necessary and sufficient commutator conditions through generator

$$A(x,s) = \frac{d}{dt} U_{s,t}\phi^{x,s}\big|_{t=s}. \tag{2.11}$$

It follows from (2.3) that A does not depend upon y.

Theorem 1. *The following are equivalent:*

a) *$\sigma \in C^1(D_T; \mathbb{R}^{p\times d})$, $h \in C^1(D_T; \mathbb{R}^p)$, there is a unique strong solution to (2.10) on $[s, \tau^{x,s})$ for each $(x,s) \in D_T$, and this solution has explicit form $\phi^{x,s}(Y_t^s, t)$ with Y_t^s defined in (2.1) and $\phi^{x,s}, U_{s,t}$ satisfying $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$.*
b) *There is a (x, s, σ, h)-representation $\phi^{x,s}, U_{s,t}$ for each $(x,s) \in D_T$.*
c) *$\sigma \in C^1(D_T; \mathbb{R}^{p\times d})$, $h \in C^1(D_T; \mathbb{R}^p)$ and the commutator conditions:*

$$(\nabla_\varphi \sigma_k)\sigma_j = (\nabla_\varphi \sigma_j)\sigma_k, \text{ for all } j,k \in \{1, ..., d\}, \tag{2.12}$$

$$(\nabla_\varphi h)\sigma_j = (\nabla_\varphi \sigma_j)h + \partial_t \sigma_j - \sigma A_j, \text{ for all } 1 \le j \le d, \tag{2.13}$$

hold on D_T for some $A \in C(D_T; \mathbb{R}^{d\times d})$.

Remark 1. The case where (2.12) holds but (2.13) may not hold will be handled in Theorem 2 to follow.

Remark 2. When (a) or (c) are known, then there could be multiple $\phi^{x,s}$, $U_{s,t}$ pairs satisfying (2.7), (2.8). However, the extra solutions to these differential equations will generally not satisfy (2.3) so they will not correspond to a (x, s, σ, h)-representation nor necessarily be useful in simulation. For example, in (5.26) of the proof of (c) implies (b) below we use $A(\phi^{s,x}(u, 0), u)$, which does not depend upon y, instead of $A(\phi^{s,x}(u, y), u)$, which would generally cause $U_{s,t}$ to violate (2.3).

Remark 3. Theorem 1 simplifies in the time-invariant h, σ coefficient case. Clearly, one only needs to check the commutator conditions on D versus D_T. However, the second commutator condition actually changes in form to:

$$(\nabla_\varphi h)\sigma_j - (\nabla_\varphi \sigma_j)h = \sigma B_j, \text{ for all } 1 \le j \le d, \tag{2.14}$$

where $B(\varphi) = -A(\varphi, 0)$. Indeed, the left hand side of (2.14) does not depend on time so the right side can not either.

Remark 4. Theorem 1 also simplifies when $d = 1$, which corresponds to appending a deterministic equation and allowing time dependence to the case considered in Kouritzin [13]. In this $d = 1$ case, (2.12) is automatically true and (2.13) becomes

$$(\nabla_\varphi h)\sigma = (\nabla_\varphi \sigma)h + \partial_t \sigma - \sigma A. \tag{2.15}$$

2.1. Simple examples solving the commutator conditions

Often, we are interested in determining which SDEs (within a class) have the representation. In this case, the commutator conditions often can be solved quickly. The easiest way to ensure (2.12) holds is to have each column a constant multiple of another $\sigma_j = c_j \sigma_1$ for all j say. However, there are other possibilities. In general, we suppose Theorem 1 (a) hence (b) and solve for σ, h in (c).

Example 2. Let $p = d = 2$ and $D \subset \mathbb{R}$ be a domain. Suppose a, e, f, g, m, n are $C^2(D)$-functions and our Fisk–Stratonovich equation has time-invariant coefficients:

$$h(\varphi_1, \varphi_2) = \begin{pmatrix} f(\varphi_1)\, g(\varphi_2) \\ m(\varphi_1)\, n(\varphi_2) \end{pmatrix}, \ \sigma(\varphi_1, \varphi_2) = \begin{pmatrix} a(\varphi_1) & 0 \\ e(\varphi_2) & e(\varphi_2) \end{pmatrix}. \tag{2.16}$$

Moreover, suppose $a(\varphi_1)$ and $e(\varphi_2)$ are never 0. Then, σ is always non-singular and it follows by (2.7) as well as the mean value theorem that for any $u \in [s, t]$

$$\phi^{x,s}(y, u) - \phi^{x,s}(\hat{y}, u) = \sigma(\phi^{x,s}(y^*, u)) U_{s,u}^{-1} \phi^{x,s} \cdot (y - \hat{y})$$

with $y^* \in \mathcal{N}^{x,s}$ for $y, \hat{y} \in \mathcal{N}^{x,s}$ and any possible representation $\phi^{x,s}, U_{s,t}$. Hence, $\phi^{x,s}(y, u) = \phi^{x,s}(\hat{y}, u) \leftrightarrow y = \hat{y}$. Therefore, it follows from (2.3) that $U_{s,u} = V_{s,u}$ can not depend upon $\phi^{x,s}(y, u)$ for any $u \in [s, t]$ and B in (2.14) is constant by (2.11). Now,

$$\nabla_\varphi h = \begin{pmatrix} f'(\varphi_1) g(\varphi_2) & f(\varphi_1) g'(\varphi_2) \\ m'(\varphi_1) n(\varphi_2) & m(\varphi_1) n'(\varphi_2) \end{pmatrix} \tag{2.17}$$

and

$$\nabla_\varphi \sigma_2 = \begin{pmatrix} 0 & 0 \\ 0 & e'(\varphi_2) \end{pmatrix}, \ \nabla_\varphi \sigma_1 = \begin{pmatrix} a'(\varphi_1) & 0 \\ 0 & e'(\varphi_2) \end{pmatrix} \tag{2.18}$$

so the first commutator condition (2.12) is fine since

$$\nabla_\varphi \sigma_1 \sigma_2 = \begin{pmatrix} 0 \\ e'(\varphi_2) e(\varphi_2) \end{pmatrix} = \nabla_\varphi \sigma_2 \sigma_1. \tag{2.19}$$

Moreover,

$$\nabla_\varphi h \sigma_2 - \nabla_\varphi \sigma_2 h = \begin{pmatrix} e(\varphi_2) f(\varphi_1) g'(\varphi_2) \\ m(\varphi_1)(e(\varphi_2) n'(\varphi_2) - e'(\varphi_2) n(\varphi_2)) \end{pmatrix} \tag{2.20}$$

and

$$\nabla_\varphi h \sigma_1 - \nabla_\varphi \sigma_1 h = \begin{pmatrix} a f' g + e f g' - a' f g \\ a m' n + e m n' - e' m n \end{pmatrix}. \tag{2.21}$$

On the other hand, denoting $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$, we have

$$\sigma B = \begin{pmatrix} ab_{11} & ab_{12} \\ eb_{11} + eb_{21} & eb_{12} + eb_{22} \end{pmatrix}. \tag{2.22}$$

Hence, by (2.14) there is an explicit solution if and only if

$$\begin{pmatrix} af'g + efg' - a'fg & efg' \\ am'n + emn' - e'mn & m(en' - e'n) \end{pmatrix} = \begin{pmatrix} ab_{11} & ab_{12} \\ eb_{11} + eb_{21} & eb_{12} + eb_{22} \end{pmatrix} \tag{2.23}$$

for constants $b_{11}, b_{12}, b_{21}, b_{22}$. If $f = c_1 a$, $n = c_2 e$, $eg' = c_3$ and $m'a = c_4$ for some constants c_1, c_2, c_3, c_4, then it is easy to show that this condition is met with $b_{22} = -c_1c_3$, $b_{21} = c_2c_4 - c_1c_3$ and $b_{11} = b_{12} = c_1c_3$ so the representation holds for

$$h(\varphi_1, \varphi_2) = \begin{pmatrix} \alpha \frac{g(\varphi_2)}{m'(\varphi_1)} \\ \beta \frac{m(\varphi_1)}{g'(\varphi_2)} \end{pmatrix}, \ \sigma(\varphi_1, \varphi_2) = \begin{pmatrix} \frac{\gamma}{m'(\varphi_1)} & 0 \\ \frac{\delta}{g'(\varphi_2)} & \frac{\delta}{g'(\varphi_2)} \end{pmatrix}, \tag{2.24}$$

where $\alpha = c_1c_4, \beta = c_2c_3, \gamma = c_4, \delta = c_3$ are any constants and g, m are C^2-functions with $\frac{1}{m'(\varphi_1)}, \frac{1}{g'(\varphi_1)} \in C^1(D)$.

Example 3. In a similar manner, it follows that

$$h(\varphi_1, \varphi_2) = \begin{pmatrix} \alpha \frac{g(\varphi_2)}{m'(\varphi_1)} \\ \beta \frac{m(\varphi_1)}{g'(\varphi_2)} \end{pmatrix}, \ \sigma(\varphi_1, \varphi_2) = \begin{pmatrix} \frac{\gamma}{m'(\varphi_1)} & 0 \\ 0 & \frac{\delta}{g'(\varphi_2)} \end{pmatrix}, \tag{2.25}$$

for any constants $\alpha, \beta, \gamma, \delta$, also has a representation.

2.2. A simple (x, s, σ, h)-representation example

There was significant work done in the previous examples and we still did not have a (x, s, σ, h)-representation. The next example is the key to solving for complete representations and will be used in the following section. It will be worth observing in this next example that $V_{s,t} = U_{s,t}\widetilde{\phi} = U_{s,t}\widetilde{X}$ (with the notation defined within the example) so the operators $U_{s,t}$ act on the deterministic part of X.

Example 4. Suppose $\sigma(\varphi, t) = \begin{pmatrix} I_r & \overline{\kappa}(\varphi, t) \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{p\times d}$ satisfies (2.12). We will find the possible h, b satisfying (2.13) and the corresponding representations $U_{s,t}, \phi^{x,s}$ by Theorem 1.
Notation: As always, φ is a variable and ϕ is the representation function. Further, let $\overline{x} = (x_1, ..., x_r)$, $\widetilde{x} = (x_{r+1}, ..., x_d)$, $\overline{\varphi} = (\varphi_1, ..., \varphi_r)$, $\widetilde{\varphi} = (\varphi_{r+1}, ..., \varphi_d)$, $\widetilde{D} = \{\widetilde{\varphi} : (\overline{\varphi}, \widetilde{\varphi}) \in D \text{ for some } \overline{\varphi}\}$, $\widetilde{D}_T = \widetilde{D} \times [0, T)$,

$$\phi^{x,s}(y, t) = \begin{pmatrix} \overline{\phi}^{x,s}(y, t) \\ \widetilde{\phi}^{x,s}(y, t) \end{pmatrix}, \ h = \begin{pmatrix} \overline{h} \\ \widetilde{h} \end{pmatrix} \text{ and } A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \tag{2.26}$$

where $A_{11} \in \mathbb{R}^{r\times r}$. Finally, we let

$$\overline{\beta}(\varphi, t) = -A_{11}(\varphi, t) - \overline{\kappa}(\varphi, t)A_{21}(\varphi, t), \tag{2.27}$$

which will appear often below.

Step 1: Interpret (2.7) and the $\mathcal{C}_3$ condition (2.3) on $U_{s,t}, A$.
Suppose $u \in [s,t]$. By (2.7) as well as the mean value theorem

$$\begin{pmatrix} \overline{\phi}^{x,s}(y,u) - \overline{\phi}^{x,s}(\hat{y},u) \\ \widetilde{\phi}^{x,s}(y,u) - \widetilde{\phi}^{x,s}(\hat{y},u) \end{pmatrix} = \begin{pmatrix} I_r & \overline{\kappa}(\phi^{x,s}(y^*,u),u) \\ 0 & 0 \end{pmatrix} U_{s,u}^{-1}\phi^{x,s} \cdot (y-\hat{y}) \tag{2.28}$$

with $y^* \in \mathcal{N}^{x,s}$ for $y, \hat{y} \in \mathcal{N}^{x,s}$ and any possible representation $\phi^{x,s}$. Hence, $\overline{\phi}^{x,s}(y,u) \neq \overline{\phi}^{x,s}(\hat{y},u)$ implies $y \neq \hat{y}$. Therefore, it follows from (2.3) that $U_{s,t}\phi^{x,s}$ can not depend upon $\overline{\phi}^{x,s}(y,u)$ for any $u \in [s,t]$, which implies $U_{s,t}\phi^{x,s} \doteq U_{s,t}\widetilde{\phi}^{x,s}$ only depends on $\widetilde{\phi}^{x,s}, t$. This also means by (2.11) that

$$A(\varphi,t) = \frac{d}{dt}U_{u,t}\widetilde{\phi}^{\varphi,u}\Big|_{u=t}. \tag{2.29}$$

Step 2: Interpret commutator conditions on $\overline{\kappa}, h$.
Let e_i denote the i^{th} column of I_p so $\sigma_i = e_i$ for $i \leq r$. We have by (2.12), that

$$\begin{pmatrix} \nabla_\varphi \overline{\kappa}_{j-r} \\ 0 \end{pmatrix} e_i = 0, \quad \forall i \in \{1,2,...,r\}, j \in r+1,...,d, \tag{2.30}$$

which establishes that $\overline{\kappa}(\widetilde{\varphi},t)$ can only depend upon $\widetilde{\varphi}, t$. This is the only restriction on $\overline{\kappa}$ from (2.12). By (2.13), we find

$$\nabla_\varphi \begin{pmatrix} \overline{h} \\ \widetilde{h} \end{pmatrix} \sigma_j - \nabla_\varphi \sigma_j \begin{pmatrix} \overline{h} \\ \widetilde{h} \end{pmatrix} = \begin{pmatrix} \overline{\beta} & \partial_t \overline{\kappa} - A_{12} - \overline{\kappa}A_{22} \\ 0 & 0 \end{pmatrix}_j \tag{2.31}$$

so $\nabla_{\overline{\varphi}}\widetilde{h} = 0$, implying $\widetilde{h}(\varphi) \in C^1(\widetilde{D}_T, \mathbb{R}^{p-r})$ only depends upon $\widetilde{\varphi}, t$, and

$$\nabla_{\overline{\varphi}}\overline{h} = \overline{\beta}, \tag{2.32}$$

$$(\nabla_{\overline{\varphi}}\overline{h})\,\overline{\kappa}_{j-r} - (\nabla_{\widetilde{\varphi}}\overline{\kappa}_{j-r})\widetilde{h} = \partial_t\overline{\kappa}_{j-r} - (A_{12} - \overline{\kappa}A_{22})_{j-r}. \tag{2.33}$$

Now, it follows from (2.32), (2.33) that

$$\overline{\beta}\,\overline{\kappa} = [(\nabla_{\widetilde{\varphi}}\overline{\kappa}_1)\widetilde{h}, ..., (\nabla_{\widetilde{\varphi}}\overline{\kappa}_{d-r})\widetilde{h}] + \partial_t\overline{\kappa} - A_{12} - \overline{\kappa}A_{22}. \tag{2.34}$$

Hence, it follows from (2.7), (2.8), (2.2) that $\widetilde{\phi}^{x,s}$ satisfies

$$\nabla_y \widetilde{\phi}^{x,s}(y,t) = 0, \tag{2.35}$$

$$\partial_t \widetilde{\phi}^{x,s}(y,t) = \widetilde{h}(\widetilde{\phi}^{x,s}(y,t),t), \tag{2.36}$$

$$\widetilde{\phi}^{x,s}(0,s) = \widetilde{x}, \tag{2.37}$$

which implies that $\widetilde{\phi}^{x,s}$ does not depend upon $\overline{\phi}^{x,s}$ nor y. Moreover, by (2.29) and (2.27), we conclude that $A(\varphi,t) \doteq A(\widetilde{\varphi},t)$ and $\overline{\beta}(\varphi,t) \doteq \overline{\beta}(\widetilde{\varphi},t)$ only depend on $\widetilde{\varphi}, t$.
Step 3: Determine possible h, b.
By (2.32), we find

$$\overline{h}(\overline{\varphi},\widetilde{\varphi},t) = \overline{\beta}(\widetilde{\varphi},t)\overline{\varphi} + \overline{\theta}(\widetilde{\varphi},t) \tag{2.38}$$

for some C^1-function $\overline{\theta}$. Hence, the possible $h(\overline{\varphi},\widetilde{\varphi},t) = \begin{pmatrix} \overline{h}(\overline{\varphi},\widetilde{\varphi},t) \\ \widetilde{h}(\widetilde{\varphi},t) \end{pmatrix}$ are:

$$\begin{gathered}\widetilde{h} \in C^1(\widetilde{D}_T, \mathbb{R}^{p-r}),\\ \overline{h} \in \left\{\overline{\theta}(\widetilde{\varphi}, t) + \overline{\beta}(\widetilde{\varphi}, t)\overline{\varphi} : \overline{\beta} \in C^1(\widetilde{D}_T, \mathbb{R}^{r\times r}); \overline{\theta} \in C^1(\widetilde{D}_T, \mathbb{R}^r)\right\}.\end{gathered} \tag{2.39}$$

From (1.5) and fact $\overline{\kappa}(\widetilde{\varphi}, t)$ only depends on $\widetilde{\varphi}, t$, we find that

$$b = h + \frac{1}{2}\sum_{j=1}^{d}\{\nabla_\varphi \sigma_j\}\sigma_j = h. \tag{2.40}$$

Free parameters: A_{21}, A_{22}, $\overline{\kappa}$, $\overline{\beta}$, $\overline{\theta}$ and $\widetilde{h}$ can be anything (subject to dimensionality and dependency on only $\widetilde{\varphi}, t$). A_{12} is then determined by (2.34) and A_{11} by (2.27). $\overline{\beta}$ and $\overline{\theta}$ also determine the possible $\overline{h}$ above and $\phi^{x,s}$ below. Different choices of $\overline{\kappa}$, $\overline{\beta}$, $\overline{\theta}$ and $\widetilde{h}$ will result in different solutions. However, there is no loss in generality in taking A_{21}, A_{22} to be zero.

Step 4: Interpret differential system for $\phi^{x,s}$.

Since $\phi^{x,s} = \begin{pmatrix} \overline{\phi}^{x,s} \\ \widetilde{\phi}^{x,s}\end{pmatrix}$ satisfies (2.8), (2.2), $\widetilde{\phi}^{x,s}$ must be of the form

$$\partial_t \widetilde{\phi}^{x,s} = \widetilde{h}(\widetilde{\phi}^{x,s}, t), \quad \text{s.t. } \widetilde{\phi}^{x,s}(s) = \widetilde{x}. \tag{2.41}$$

We let $\widetilde{X}_t$ denote the solution of this differential equation. Next, since $\phi^{x,s}$ satisfies (2.7), $\overline{\phi}^{x,s}$ must be of the form

$$\overline{\phi}^{x,s}(y,t) = \overline{c}(t) + \left[I_r \;\; \overline{\kappa}(\widetilde{X}_t, t)\right] U_{s,t}^{-1}\widetilde{\phi}^{x,s} y, \tag{2.42}$$

for some $\overline{c} \in C^1([0,T); \mathbb{R}^r)$. Differentiating in t, noting by (2.29) (with $\widetilde{\varphi} = \widetilde{X}_t$) that

$$A(\widetilde{X}_t, t) = \frac{d}{dt} U_{u,t}\widetilde{\phi}^{\widetilde{X}_u,u}\big|_{u=t}, \tag{2.43}$$

and using (2.42), (2.43), (2.4), (2.41), (2.34), (2.27), one has (with $U_{s,t}^{-1} = U_{s,t}^{-1}\widetilde{\phi}^{x,s}$) that

$$\begin{aligned}&\partial_t \overline{\phi}^{x,s}(y,t) \\ &= \overline{c}'(t) - \left[I \;\; \overline{\kappa}(\widetilde{X}_t, t)\right] A(\widetilde{X}_t, t) U_{s,t}^{-1} y \\ &+ \left[0 \;\; \partial_t \overline{\kappa}(\widetilde{X}_t, t) + \nabla_{\widetilde{\varphi}}\overline{\kappa}_1(\widetilde{X}_t, t)\,\widetilde{h}(\widetilde{X}_t, t), ..., \nabla_{\widetilde{\varphi}}\overline{\kappa}_{d-r}(\widetilde{X}_t, t)\,\widetilde{h}(\widetilde{X}_t, t)\right] U_{s,t}^{-1} y \\ &= \overline{c}'(t) + \overline{\beta}(\widetilde{X}_t, t)[I \;\; \overline{\kappa}(\widetilde{X}_t, t)] U_{s,t}^{-1} y \\ &= \overline{c}'(t) + \overline{\beta}(\widetilde{X}_t, t)(\overline{\phi}^{x,s}(y,t) - \overline{c}(t)).\end{aligned} \tag{2.44}$$

On the other hand, by (2.8) and (2.38)

$$\partial_t \overline{\phi}^{x,s}(y,t) = \overline{\theta}(\widetilde{X}_t, t) + \overline{\beta}(\widetilde{X}_t, t)\overline{\phi}^{x,s}(y,t). \tag{2.45}$$

Comparing (2.44) and (2.45), one has that

$$\overline{c}'(t) = \overline{\theta}(\widetilde{X}_t, t) + \overline{\beta}(\widetilde{X}_t, t)\overline{c}(t) \;\; \text{subject to } \overline{c}(s) = \overline{x}. \tag{2.46}$$

Step 5: Determine U in terms of $\overline{\kappa}$, $\overline{\beta}$ and θ.

We just need A to satisfy (2.27), (2.34) so there is no loss of generality in taking

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}(\widetilde{\varphi},t) = \begin{pmatrix} -\overline{\beta} & [(\nabla_{\widetilde{\varphi}}\overline{\kappa}_1)\widetilde{h},...,(\nabla_{\widetilde{\varphi}}\overline{\kappa}_{d-r})\widetilde{h}] + \partial_t\overline{\kappa} - \overline{\beta}\,\overline{\kappa} \\ 0 & 0 \end{pmatrix}(\widetilde{\varphi},t). \tag{2.47}$$

By (2.43), (2.4) and (2.47), we know

$$\begin{aligned} \partial_t U_{s,t}\widetilde{X} &= (U_{s,t}\widetilde{X})A(\widetilde{X}_t,t) \\ &= U_{s,t}\widetilde{X}\begin{pmatrix} -\overline{\beta} & \{[(\nabla_{\widetilde{\varphi}}\overline{\kappa}_1)\widetilde{h},...,(\nabla_{\widetilde{\varphi}}\overline{\kappa}_{d-r})\widetilde{h}] + \partial_t\overline{\kappa} - \overline{\beta}\,\overline{\kappa}\} \\ 0 & 0 \end{pmatrix}(\widetilde{X}_t,t) \end{aligned} \tag{2.48}$$

subject to $U_{s,s}\widetilde{X} = U_{s,s}\widetilde{x} = I_d$. Now, suppose that $T_{u,t}$ is the two parameter semigroup:

$$\frac{d}{dt}T_{u,t} = -T_{u,t}\,\overline{\beta}(\widetilde{X}_t,t), \quad \forall t \geq u \ \text{ subject to } T_{u,u} = I_r. \tag{2.49}$$

Then, the solution of (2.48) is

$$U_{s,t}\widetilde{X} = \begin{pmatrix} T_{s,t} & T_{s,t}\overline{\kappa}(\widetilde{X}_t,t) - \overline{\kappa}(\widetilde{X}_s,s) \\ 0 & I_{d-r} \end{pmatrix}, \tag{2.50}$$

and so

$$U_{s,t}^{-1}\widetilde{X} = \begin{pmatrix} T_{s,t}^{-1} & T_{s,t}^{-1}\overline{\kappa}(\widetilde{X}_s,s) - \overline{\kappa}(\widetilde{X}_t,t) \\ 0 & I_{d-r} \end{pmatrix}. \tag{2.51}$$

Moreover, it follows by (2.46) that $\overline{c}$ can also be expressed in terms of $T_{s,t}^{-1}$.
Step 6: Solution Algorithm.

- a: Check $\overline{\kappa}$ only depends upon $\widetilde{\varphi}, t$. This must be true by Step 2.
- b: Choose any functions $\overline{\beta} \in C^1(\widetilde{D}_T, \mathbb{R}^{r\times r}); \overline{\theta} \in C^1(\widetilde{D}_T, \mathbb{R}^r)$ and $\widetilde{h} \in C^1(\widetilde{D}_T, \mathbb{R}^{p-r})$ for drift of the form $b(\overline{\varphi},\widetilde{\varphi},t) = h(\overline{\varphi},\widetilde{\varphi},t) = \begin{pmatrix} \overline{\theta}(\widetilde{\varphi},t) + \overline{\beta}(\widetilde{\varphi},t)\overline{\varphi} \\ \widetilde{h}(\widetilde{\varphi},t) \end{pmatrix}$. These are the only possible drifts by Step 3.
- c: Solve

$$\widetilde{X}_t' = \widetilde{h}(\widetilde{X}_t,t) \ \text{ subject to } \widetilde{X}_s = \widetilde{x}$$

- d: Solve

$$\frac{d}{dt}T_{s,t} = -T_{s,t}\,\overline{\beta}(\widetilde{X}_t,t), \quad \forall t \geq s \ \text{ subject to } T_{s,s} = I_r. \tag{2.52}$$

Then, set

$$U_{s,t}\widetilde{X} = \begin{pmatrix} T_{s,t} & T_{s,t}\overline{\kappa}(\widetilde{X}_t,t) - \overline{\kappa}(\widetilde{X}_s,s) \\ 0 & I_{d-r} \end{pmatrix}, \tag{2.53}$$

$$U_{s,t}^{-1}\widetilde{X} = \begin{pmatrix} T_{s,t}^{-1} & T_{s,t}^{-1}\overline{\kappa}(\widetilde{X}_s,s) - \overline{\kappa}(\widetilde{X}_t,t) \\ 0 & I_{d-r} \end{pmatrix}, \tag{2.54}$$

$$\overline{c}(t) = T_{s,t}^{-1}\overline{x} + T_{s,t}^{-1}\int_s^t T_{s,u}\overline{\theta}(\widetilde{X}_u,u)du. \tag{2.55}$$

e: Divide $\phi^{x,s} = \begin{pmatrix} \overline{\phi}^{x,s} \\ \widetilde{\phi}^{x,s} \end{pmatrix}$ and set $\widetilde{\phi}^{x,s}(t) = \widetilde{X}_t$,

$$\overline{\phi}^{x,s}(y,t) = \overline{c}(t) + \left[I_r \quad \overline{\kappa}(\widetilde{X}_t, t) \right] (U_{s,t}^{-1} \widetilde{X}) y.$$

The preceding example was intuitively pleasing: We showed you could indeed represent *linear* SDEs using a single Gaussian stochastic integral. Further, we showed that we could *append* an ordinary differential equation ($d\widetilde{X}_t = \widetilde{h}(\widetilde{X}_t)dt$) and use its solution within the coefficients of the stochastic differential equation. Finally, we showed how to construct the solution. While none of this is surprising, it does explain our necessary and sufficient conditions. In the next section, we will show how to combine this example with diffeomorphisms to handle the general case with nonlinear coefficients.

3. Construction results and examples

When one explicit solution exists, there will be a whole class of such solutions corresponding to distinct b's. We now identify the b's, ϕ's and U's for these solutions corresponding to a given σ. This is done by using local diffeomorphisms to convert the general case to the case of Example 4. The idea is based upon the following simple lemma.

Lemma 1. *Suppose $D \subset \mathbb{R}^p$ is a domain, $T > 0$, $D_T = D \times [0,T)$, $\widehat{\Lambda} \doteq \begin{pmatrix} \Lambda_t \\ t \end{pmatrix} : D_T \to \widehat{\Lambda}(D_T) \subset \mathbb{R}^{p+1}$ is a C^2-diffeomorphism and $\sigma, b, h, \{\phi^{x,s}\}_{(x,s)\in D_T}$, $\{U_{s,t}\phi^{x,s}\}_{(x,s)\in D_T, s\le t<T}$, A satisfy Conditions $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$ as well as equations (1.5), (2.11). Let $\widehat{D}_T = \widehat{\Lambda}(D_T)$,*

$$\widehat{\sigma} = \{(\nabla_\varphi \Lambda_t)\sigma\} \circ \widehat{\Lambda}^{-1}, \quad \widehat{h} = \{(\nabla_\varphi \Lambda_t)h\} \circ \widehat{\Lambda}^{-1},$$

$$\widehat{b} = \left\{ (\nabla_\varphi \Lambda_t)b + \frac{1}{2}\sum_{j=1}^{d}\sum_{i,k=1}^{p} (\partial_{\varphi_i}\partial_{\varphi_k}\Lambda_t)\sigma_{i,j}\sigma_{k,j} \right\} \circ \widehat{\Lambda}^{-1},$$

$$\widehat{\phi}^{x,s}(y,t) = \Lambda_t \circ \phi^{\widehat{\Lambda}^{-1}(x,s)}(y,t),$$

$$\widehat{U}_{s,t}\widehat{\phi}^{x,s} = U_{s,t}\phi^{\widehat{\Lambda}^{-1}(x,s)},$$

$$\widehat{A} = A \circ \widehat{\Lambda}^{-1}.$$

Then, $\widehat{\sigma}, \widehat{b}, \widehat{h}, \{\widehat{\phi}^{x,s}\}_{(x,s)\in \widehat{D}_T}, \widehat{U}, \widehat{A}$ satisfy Conditions $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$ as well as equations (1.5), (2.11) on $\widehat{D}_T$. Moreover,

i) *$\widehat{\phi}^{x,s}, \widehat{U}_{s,t}$ is a $(x,s,\widehat{\sigma},\widehat{h})$-representation for each $(x,s) \in \widehat{D}_T$ if and only if $\phi^{x,s}, U_{s,t}$ is a (x,s,σ,h)-representation for each $(x,s) \in D_T$.*
ii) *Equation (2.12) holds if and only if*

$$(\nabla_\varphi \widehat{\sigma}_k)\widehat{\sigma}_j = (\nabla_\varphi \widehat{\sigma}_j)\widehat{\sigma}_k, \text{ on } \widehat{D}_T \text{ for all } j,k \in \{1,...,d\}. \tag{3.1}$$

iii) *Equation (2.13) holds if and only if*

$$(\nabla_\varphi \widehat{h})\widehat{\sigma}_j = (\nabla_\varphi \widehat{\sigma}_j)\widehat{h} + \partial_t \widehat{\sigma}_j - \widehat{\sigma}\widehat{A}_j, \text{ on } \widehat{D}_T \text{ for all } 1 \le j \le d. \tag{3.2}$$

Remark 5. In the time-homogeneous case, we can deal with B instead of A and set $\widehat{B} = B \circ \Lambda_0^{-1}$. The notation $\widehat{\Lambda} \doteq \begin{pmatrix} \Lambda_t \\ t \end{pmatrix}$ just means that $\widehat{\Lambda}$ is a diffeomorphism with the constraint that the last component is the identity map.

Proof. This lemma follows by direct calculation. Perhaps, the fastest way to verify the commutator conditions is to think of (1.4) as a time-homogeneous equation

$$d\begin{bmatrix} X_t \\ t \end{bmatrix} = \begin{bmatrix} h(X_t,t) \\ 1 \end{bmatrix} dt + \begin{bmatrix} \sigma(X_t,t) \\ 0 \end{bmatrix} \bullet dW_t, \qquad \begin{bmatrix} X_s \\ s \end{bmatrix} = \begin{bmatrix} x \\ s \end{bmatrix}$$

on $[s, \tau^{x,s})$, by appending the trivial equation $t = t$ and thinking of t as an additional state variable. Then, verifying (2.13) is equivalent to (3.2) is the same as verifying

$$\begin{aligned} &\left(\nabla \begin{bmatrix} h \\ 1 \end{bmatrix}\right) \begin{bmatrix} \sigma_j \\ 0 \end{bmatrix} = \left(\nabla \begin{bmatrix} \sigma_j \\ 0 \end{bmatrix}\right) \begin{bmatrix} h \\ 1 \end{bmatrix} - \begin{bmatrix} \sigma \\ 0 \end{bmatrix} A_j \\ \leftrightarrow &\left(\nabla \begin{bmatrix} \widehat{h} \\ 1 \end{bmatrix}\right) \begin{bmatrix} \widehat{\sigma}_j \\ 0 \end{bmatrix} = \left(\nabla \begin{bmatrix} \widehat{\sigma}_j \\ 0 \end{bmatrix}\right) \begin{bmatrix} \widehat{h} \\ 1 \end{bmatrix} - \begin{bmatrix} \widehat{\sigma} \\ 0 \end{bmatrix} \widehat{A}_j, \end{aligned}$$

which avoids $\partial_t \sigma_j$ and Λ_t if we express $(\widehat{h}^T, 1)^T$ and $(\widehat{\sigma}_j^T, 0)^T$ in terms of $\widehat{\Lambda}$. □

The idea behind this lemma is that σ gets changed into $\widehat{\sigma} = \begin{pmatrix} I_r & \overline{\kappa} \\ 0 & 0 \end{pmatrix}$ with some diffeomorphism and we can use Example 4 to solve for the possible $\hat{h}$ and the representations $\widehat{\phi}^{x,s}$, $\widehat{U}^{x,s}$. Unfortunately, it is sometimes impossible to have a single diffeomorphism for all of D_T and, even when it is possible, we may not know that until after local diffeomorphisms are constructed and one of them is extendable to all of D_T.

Definition 2. Suppose $(x, s) \in D_T$. Then, an (x, s)-local diffeomorphism $(O^{x,s}, \widehat{\Lambda}^{x,s})$ is a bijection $\widehat{\Lambda}^{x,s} : O^{x,s} \to \widehat{\Lambda}^{x,s}(O^{x,s})$ such that $\widehat{\Lambda}^{x,s} \in C^2(O^{x,s}; \mathbb{R}^{p+1})$, where $O^{x,s} \subset D_T$ is a (relatively open) neighborhood of x, s. We define $\nabla\widehat{\Lambda}^{-1}(\widehat{\Lambda}(\varphi, t))$ to be $\left[\nabla\widehat{\Lambda}(\varphi, t)\right]^{-1}$ for $(\varphi, t) \in O^{x,s}$.

We imposed sufficient differentiability on our local diffeomorphisms for our uses to follow. Our (x, s)-local diffeomorphisms will take the form $\widehat{\Lambda} = \begin{pmatrix} \Lambda_t \\ t \end{pmatrix}$ with Λ_t being constructed from σ under the conditions:

- $[D]$: Let $D \subset \mathbb{R}^p$ be a bounded convex domain, $T > 0$ and $D_T = D \times [0, T)$.
- $[H_r]$: The rank of σ is r on D_T with the first r rows having full row rank.
- $[\partial]$: $\sigma \in C^{r+1}(D_T; \mathbb{R}^{p\times d})$.
- $[B]$: $(\nabla_\varphi \sigma_j)\sigma_k - (\nabla_\varphi \sigma_k)\sigma_j = 0$ on D_T, for $1 \le j, k \le d$ and $(x, s) \in D_T$.

To ensure the row rank part of H_r, we can just permute the rows of (1.3), amounting to relabeling the $\{X_t^i\}_{i=1}^p$.

Proposition 1. *Suppose [D, H_r, ∂, B] hold. Then for any $(x, s) \in D_T$, there exists an (x, s)-local diffeomorphism $(O^{x,s}, \widehat{\Lambda}^{x,s})$ and a constant permutation matrix π such that*

$$\widehat{\sigma} \doteq \{(\nabla_\varphi \Lambda_t)\sigma\pi\} \circ \widehat{\Lambda}^{-1} = \begin{pmatrix} I_r & \overline{\kappa} \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{p\times d} \ \textit{on} \ \widehat{\Lambda}(O^{x,s}),$$

where $\overline{\kappa} \in C^1(\widehat{\Lambda}(O^{x,s}); \mathbb{R}^{r\times(d-r)})$ *does not depend on* $\varphi_1, \dots, \varphi_r$.

Proof. Provided in Section 5. □

Remark 6. The permutation matrix π permutes the columns of σ. We label the permuted diffusion coefficient $\sigma^\pi = \sigma\pi$ and note that

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t = b(X_t)dt + \sigma^\pi(X_t)dW_t^\pi,$$

where $W^\pi = \pi^{-1}W$ is a permutation of the Brownian motions W. Also, the Stratonovich drift h remains the same by (1.5).

Remark 7. It follows from the proof in Section 5 that the diffeomorphism can have the form $\widehat{\Lambda} = \widehat{\Lambda}_r \circ \cdots \circ \widehat{\Lambda}_2 \circ \widehat{\Lambda}_1$ for any diffeomorphisms $\widehat{\Lambda}_i : \widehat{\Lambda}_{i-1} \circ \cdots \circ \widehat{\Lambda}_2 \circ \widehat{\Lambda}_1(D_T) \to \mathbb{R}^{p+1}$ satisfying $\{\nabla\widehat{\Lambda}_i \cdots \nabla\widehat{\Lambda}_2\nabla\widehat{\Lambda}_1\sigma_i^\pi\} \circ \widehat{\Lambda}_1^{-1} \circ \widehat{\Lambda}_2^{-1} \circ \cdots \circ \widehat{\Lambda}_i^{-1} = e_i$, where $(e_1\, e_2\, \dots\, e_p\, e_{p+1}) = I_{p+1}$ is the identity matrix. However, as will be seen below in Remark 9, this does not uniquely define the diffeomorphism.

Proposition 1 immediately provides us our second main theorem.

Theorem 2. *Suppose [D, H_r, ∂, B] hold, $h \in C^1(D_T; \mathbb{R}^p)$, $(x, s) \in D_T$ and W is an $\mathbb{R}^d$-valued standard Brownian motion. Then, there exists a stopping time $\tau > s$, a permutation matrix π and an (x, s)-local diffeomorphism $(O^{x,s}, \widehat{\Lambda}^{x,s})$, as in Proposition 1 and Remark 7, such that*

$$i)\ \ \widehat{\sigma} \doteq \{(\nabla_\varphi \Lambda_t)\sigma^\pi\} \circ \widehat{\Lambda}^{-1} = \begin{pmatrix} I_r & \overline{\kappa} \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{p\times d} \ \textit{on} \ \widehat{\Lambda}(O^{x,s}),$$

with $\overline{\kappa} \in C^1(\Lambda(O^{x,s}); \mathbb{R}^{r\times(d-r)})$ not depending on $\varphi_1, \dots, \varphi_r$ and ii) the Stratonovich SDE $dX_t = h(X_t)dt + \sigma(X_t) \bullet dW_t,\ X_s = x$ has a solution $X_t = \Lambda_t^{-1}\begin{pmatrix} \overline{X}_t \\ \widetilde{X}_t \end{pmatrix}$ on $[0, \tau]$ if and only if the simpler SDE

$$d\begin{bmatrix} \overline{X}_t \\ \widetilde{X}_t \end{bmatrix} = \widehat{h}\begin{pmatrix} \overline{X}_t \\ \widetilde{X}_t \end{pmatrix} dt + \begin{pmatrix} I_r & \overline{\kappa} \\ 0 & 0 \end{pmatrix} dW_t^\pi,\ \begin{bmatrix} \overline{X}_s \\ \widetilde{X}_s \end{bmatrix} = \Lambda_s(x) \tag{3.3}$$

has a solution on $[0, \tau]$, where $\widehat{h} = (\nabla_\varphi \Lambda_t h + \partial_t \Lambda_t) \circ \widehat{\Lambda}^{-1}$.

We stated the simpler SDE in terms of Itô integration. However, it follows by (1.5) and the nature of $\overline{\kappa}$ that (3.3) would have exactly the same form in terms of Stratonovich integration.

In this theorem we do not have a commutator condition for h so we can not guarantee the simple form of $\widehat{h}$ as in Example 4. This means that $\widetilde{X}$ is not in general deterministic nor is $\overline{X}$ necessarily Gaussian. We also impose slightly stronger conditions on σ compared to Theorem 1 but gain information about the representation as local diffeomorphisms.

For our final main result, we add back the commutator condition for h, and characterize all the solutions $X_t^{x,s} = \phi^{x,s}(Y_t^s, t)$ to (2.10) via Example 4. We do this through our basic set of parameters for (x, s):

Definition 3. Suppose [D, H_r, ∂] hold, D_T is as defined above and $O = O^{x,s} \subset D_T$ below. Let $\mathcal{P} = \mathcal{P}_\sigma^{x,s}$ be the set of all $(\widehat{\Lambda}, \overline{\kappa}, \overline{\beta}, \overline{\theta}, \widetilde{h}, \pi)$ such that

P0) π is a constant permutation matrix.

P1) $(O^{x,s}, \widehat{\Lambda}^{x,s})$ is a (x,s)-local diffeomorphism, where $\widehat{\Lambda}(\varphi,t) = \begin{bmatrix}\Lambda_t(\varphi)\\ t\end{bmatrix}$. For convenience, we let $\Lambda_t = \begin{bmatrix}\overline{\Lambda}_t\\ \widetilde{\Lambda}_t\end{bmatrix}$ with $\overline{\Lambda}_t \in \mathbb{R}^r$;

P2) $\overline{\kappa} \in C^1(\widehat{\Lambda}(O); \mathbb{R}^{r\times(d-r)})$ depends only on $\varphi_{r+1}, \ldots, \varphi_p$, and t;

P3) $\{(\nabla_\varphi \Lambda_t)\sigma^\pi\} \circ (\widehat{\Lambda})^{-1} = \begin{pmatrix} I_r & \overline{\kappa}\\ 0 & 0\end{pmatrix}$ on $\widehat{\Lambda}(O)$;

P4) $\overline{\beta} \in C^1(\widehat{\Lambda}(O); \mathbb{R}^{r\times r})$ depends only on $\varphi_{r+1}, \ldots, \varphi_p$, and t;

P5) $\overline{\theta} \in C^1(\widehat{\Lambda}(O); \mathbb{R}^{r})$ depends only on $\varphi_{r+1}, \ldots, \varphi_p, t$;

P6) $\widetilde{h} \in C^1(\widehat{\Lambda}(O); \mathbb{R}^{p-r})$ depends only on $\varphi_{r+1}, \ldots, \varphi_p, t$.

To each $(\widehat{\Lambda}, \overline{\kappa}, \overline{\beta}, \overline{\theta}, \widetilde{h}, \pi) \in \mathcal{P}$, we associate the following functions:

$$\begin{cases} \widetilde{X} = \widetilde{X}^{x,s} \in \mathbb{R}^{p-r} \text{ uniquely solves } \frac{d}{dt}\widetilde{X}_t = \widetilde{h}(\widetilde{X}_t, t),\ \widetilde{X}_s = \widetilde{\Lambda}_s(x);\\ G(t) = \left(I_r \middle|\ \overline{\kappa}(\tilde{X}_t, t)\right) \in \mathbb{R}^{r\times d};\\ \frac{d}{du}T_{s,u} = -T_{s,u}\,\overline{\beta}(\widetilde{X}_u, u),\quad \forall u \ge s \text{ subject to } T_{s,s} = I_r;\\ U_{s,u}\widetilde{X} = \begin{pmatrix} T_{s,u} & T_{s,u}\overline{\kappa}(\widetilde{X}_u, u) - \overline{\kappa}(\widetilde{X}_s, s)\\ 0 & I_{d-r}\end{pmatrix};\\ U^{-1}_{s,u}\widetilde{X} = \begin{pmatrix} T^{-1}_{s,u} & T^{-1}_{s,u}\overline{\kappa}(\widetilde{X}_s, s) - \overline{\kappa}(\widetilde{X}_u, u)\\ 0 & I_{d-r}\end{pmatrix};\\ \overline{c}_s(t) = T^{-1}_{s,t}\overline{\Lambda}_s(x) + T^{-1}_{s,t}\int_s^t T_{s,u}\overline{\theta}(\widetilde{X}_u, u)du. \end{cases} \tag{3.4}$$

The following theorem follows from Theorem 2, Theorem 1 (so the explicit solution implies $[B]$ above) and Example 4. In particular, we must have

$$(\nabla_\varphi \Lambda_t h + \partial_t \Lambda_t) \circ \widehat{\Lambda}^{-1} = \begin{pmatrix} \overline{h}(\overline{\varphi}, \widetilde{\varphi}, t)\\ \widetilde{h}(\widetilde{\varphi}, t)\end{pmatrix} = \begin{pmatrix} \overline{\theta}(\widetilde{\varphi}, t) + \overline{\beta}(\widetilde{\varphi}, t)\overline{\varphi}\\ \widetilde{h}(\widetilde{\varphi}, t)\end{pmatrix}, \tag{3.5}$$

which gives our possible drifts h in the following theorem.

Theorem 3. *Suppose [D, H_r, ∂] hold, $(x,s) \in D_T$ and $X^{x,s}_t = \phi^{x,s}\left(\int_s^t U_{s,u}\phi^{x,s}dW^\pi_u, t\right)$, with ϕ, U satisfying $\mathcal{C}_1$, $\mathcal{C}_2$, $\mathcal{C}_3$, solves (2.10) up to some stopping time $\tau^{x,s} > s$. Then, there exists $((O^{x,s}, \widehat{\Lambda}^{x,s}), \overline{\kappa}, \overline{\beta}, \overline{\theta}, \widetilde{h}, \pi) \in \mathcal{P}^{x,s}_\sigma$, and related functions $\widetilde{X}, G, U, \overline{c}$ defined by (3.4), such that*

$$h = [\nabla_\varphi \Lambda_t]^{-1}\left\{\begin{bmatrix}\overline{\theta}(\widetilde{X}_t, t)\\ \widetilde{h}(\widetilde{X}_t, t)\end{bmatrix} - \partial_t\Lambda_t + \begin{bmatrix}\overline{\beta}(\widetilde{X}_t, t)\ \overline{\Lambda}_t\\ 0\end{bmatrix}\right\} \text{ on } O^x, \tag{3.6}$$

$$\phi^{x,s}(y,t) = \phi_{(\widehat{\Lambda}, \overline{\kappa}, \overline{\beta}, \overline{\theta}, \widetilde{h})}(y,t) = \Lambda_t^{-1}\left(\begin{bmatrix}\overline{c}_s(t) + G(t)(U^{-1}_{s,t}\widetilde{X})y\\ \widetilde{X}_t\end{bmatrix}\right) \tag{3.7}$$

on $\mathcal{N}^x = \left\{(y,t) : \begin{bmatrix}\overline{c}_s(t) + G(t)U^{-1}_{s,t}\widetilde{X}y\\ \widetilde{X}_t\end{bmatrix} \in \Lambda_t(O^{x,s})\right\}$. Finally, if $\breve{\pi}$, $\breve{\Lambda}$ and $\breve{\kappa}$ also satisfies P0–P3, then there exist $\breve{\beta}, \breve{\theta}, \breve{h}$ such that $(\breve{\Lambda}, \breve{\kappa}, \breve{\beta}, \breve{\theta}, \breve{h}, \breve{\pi}) \in \mathcal{P}$, $b_{(\breve{\Lambda}, \breve{\kappa}, \breve{\beta}, \breve{\theta}, \breve{h}, \breve{\pi})} = b_{(\widehat{\Lambda}, \overline{\kappa}, \overline{\beta}, \overline{\theta}, \widetilde{h})}$, and $\phi_{(\breve{\Lambda}, \breve{\kappa}, \breve{\beta}, \breve{\theta}, \breve{h}, \breve{\pi})} = \phi_{(\widehat{\Lambda}, \overline{\kappa}, \overline{\beta}, \overline{\theta}, \widetilde{h})}$.

Remark 8. For the sake of brevity in the examples below, we will just give local diffeomorphisms satisfying P3) above. However, as is shown in our companion paper Kouritzin [14], it is often possible to solve for them using the technique used in the proof of Proposition 1 herein.

Remark 9. To illustrate the need of the final statement of Theorem 3, we take for example, $\sigma(x) = x \in \mathbb{R}^p$. Then, any $L \in C^1(\mathbb{R}^p)$ depending on $x_2/x_1, \ldots, x_p/x_1$ satisfies $(\nabla L)\sigma = 0$. Therefore, $\widehat{\Lambda}$ and hence the parameter set is not unique but we can create the same b, ϕ from any consistent $\overline{\kappa}, \widehat{\Lambda}$.

3.1. One dimensional case

Suppose $d = p = r = 1$, $D \subset \mathbb{R}$ and $x \in D$. Then, $\overline{\kappa}, \widetilde{h}$ do not exist and $\overline{\beta}, \overline{\theta}$ only depend on t. Moreover, $U_{s,t} = T_{s,t} = e^{-\int_s^t \overline{\beta}(u)du}$, $\overline{c}_s(t) = T_{s,t}^{-1}\overline{\Lambda}_s(x) + T_{s,t}^{-1}\int_s^t T_{s,u}\overline{\theta}(u)du$ and the diffeomorphism can be taken as $\Lambda_t(\varphi) = \int \frac{1}{\sigma(\varphi,t)}d\varphi$. One then finds by (1.5), (3.4), (3.6), (3.7) that the corresponding diffusion drift b and explicit solutions are

$$b(\varphi, t) = \sigma(\varphi, t)\left\{\overline{\theta}(t) + \overline{\beta}(t)\Lambda_t(\varphi) - \partial_t\Lambda_t\right\} + \frac{1}{2}\sigma(\varphi, t)\partial_\varphi\sigma(\varphi, t) \tag{3.8}$$

$$X_t = \Lambda_t^{-1}\left[\left\{\Lambda_s(x) + \int_s^t T_{s,u}\overline{\theta}(u)du + \int_s^t T_{s,u}dW_u\right\}\Big/ T_{s,t}\right]. \tag{3.9}$$

Example 5 *(Time-varying Cox–Ingersoll–Ross model).* Suppose $\overline{\theta}, \overline{\beta}$ and continuously differentiable $s(t) > 0$ are chosen and $\sigma(\varphi, t) = s(t)\sqrt{\varphi}$. Then, $\Lambda_t(\varphi) = \frac{2\sqrt{\varphi}}{s(t)}$, $\Lambda_t^{-1}(z) = \left(\frac{zs(t)}{2}\right)^2$ and the possible Itô drifts are

$$b(\varphi, t) = \overline{\theta}(t)s(t)\sqrt{\varphi} + 2\left(\overline{\beta}(t) + \frac{\dot{s}(t)}{s(t)}\right)\varphi + \frac{s^2(t)}{4}.$$

The explicit solutions are then

$$\begin{aligned} X_t^{x,s} = \Bigg| &\frac{s(t)}{s(s)}e^{\int_s^t \overline{\beta}(v)dv}\sqrt{x} \\ &+ \frac{s(t)}{2}\left\{\int_s^t e^{\int_u^t \overline{\beta}(v)dv}\overline{\theta}(u)du + \int_s^t e^{\int_u^t \overline{\beta}(v)dv}dW_u\right\}\Bigg|^2. \end{aligned} \tag{3.10}$$

In the case $s(t) = \sigma, \overline{\theta}$ and $\overline{\beta}$ are taken constant, we get

$$X_t^{x,s} = \frac{1}{4}\left\{2e^{\overline{\beta}(t-s)}\sqrt{x} + \frac{\overline{\theta}\sigma}{\overline{\beta}}(e^{\overline{\beta}(t-s)} - 1) + \sigma\int_s^t e^{\overline{\beta}(t-u)}dW_u\right\}^2$$

solves

$$dX_t^{x,s} = \left(\sigma^2/4 + 2\overline{\beta}X_t^{x,s} + \sigma\overline{\theta}\sqrt{X_t^{x,s}}\right)dt + \sigma\sqrt{X_t^{x,s}}dW_t,\ X_s = x$$

as long as $X_t^{x,s} > 0$. This solves the usual CIR model

$$dX_t = \alpha\left(\beta - X_t\right)dt + \sigma\sqrt{X_t}dW_t, \tag{3.11}$$

when $\overline{\theta} = 0$, $\alpha = 2\overline{\beta}$, $\beta = \sigma^2/(8\overline{\beta})$. Now, set $Y_t = \sqrt{X_t}$, where X solves (3.11) with $\sigma^2 = 4\alpha\beta$, and $\tau = \inf\{t > 0; X_t = 0\}$. It is well known that $P(\tau < \infty) = 1$. Then,

$$\begin{aligned} dY_t &= \frac{1}{8Y_t}\left(4\alpha\beta-\sigma^2\right)dt - \frac{\alpha}{2}Y_t dt + \frac{\sigma}{2}dW_t \\ &= -\frac{\alpha}{2}Y_t dt + \frac{\sigma}{2}dW_t, \end{aligned} \tag{3.12}$$

by Itô's formula. However, since (3.12) defines a Gaussian process and Y must be non-negative, one cannot have Y_t defined by (3.12) unless $t<\tau$. This explains why we first look for explicit *local* solutions.

3.2. Square non-singular case

Suppose that $d=p=r$, $\sigma=\sigma(\varphi,t)$ is a $d\times d$ non-singular continuously-differentiable matrix satisfying (2.12), $D\subset\mathbb{R}^p$ and $x\in D$. Again, we apply Theorem 3 and find $\overline{\kappa},\widetilde{h}$ do not exist while $\overline{\beta},\overline{\theta}$ only depend on t. Also, there is a local diffeomorphism $\widehat{\Lambda}=\begin{pmatrix}\Lambda_t\\ t\end{pmatrix}$ such that $\nabla_\varphi\Lambda_t(\varphi)=[\sigma(\varphi,t)]^{-1}$, and all explicit solutions are of the form $\phi^{x,s}(t,y)=\Lambda_t^{-1}\left(\overline{c}_s(t)+U_{s,t}^{-1}y\right)$, where

$$U_{s,t}=-\int_s^t U_{s,u}\overline{\beta}(u)du+I \text{ and } \overline{c}_s(t)=U_{s,t}^{-1}\left\{\Lambda_s(x)+\int_s^t U_{s,u}\overline{\theta}(u)du\right\}$$

for some $\overline{\theta}\in C([0,T);\mathbb{R}^d)$ and $\overline{\beta}\in C^1([0,T),\mathbb{R}^{d\times d})$. The resulting drift is

$$b(\varphi,t)=\sigma(\varphi,t)\left\{\overline{\theta}(t)+\overline{\beta}(t)\Lambda_t(\varphi)-\partial_t\Lambda_t(\varphi)\right\}+\frac{1}{2}\sum_{j=1}^d(\nabla_\varphi\sigma_j(\varphi,t))\sigma_j(\varphi,t).$$

Example 6. *Geometric Brownian motions:* Take $\sigma_{ij}(\varphi)=\varphi_i\gamma_{ij}$ with γ non-singular and $D=(0,\infty)^d$. Then, σ satisfies the commutation condition (2.12) since $[(\nabla_\varphi\sigma_j)\sigma_k]_i=\varphi_i\gamma_{ij}\gamma_{ik}$, and the diffeomorphism can be chosen as $\Lambda(\varphi)=\Lambda_t(\varphi)=\gamma^{-1}\begin{bmatrix}\log\varphi_1\\ \vdots\\ \log\varphi_d\end{bmatrix}$. Λ's image is $\mathbb{R}^d$, so $\Lambda^{-1}(z)=\begin{bmatrix}e^{(\gamma z)_1}\\ \vdots\\ e^{(\gamma z)_d}\end{bmatrix}$ is defined everywhere and $\phi_i^{x,s}(y,t)=\exp\left[\gamma\{\overline{c}_s(t)+U_{s,t}^{-1}y\}\right]_i$. The possible drifts satisfy

$$b_i(\varphi,t)=\varphi_i\left\{\alpha_i(t)-\sum_{j=1}^d B_{ij}(t)\log\varphi_j\right\},$$

for $1\le i\le d$, where $B(t)=\gamma\overline{\beta}(t)\gamma^{-1}$, and $\alpha_i(t)=\frac{1}{2}[\gamma\gamma^\top]_{ii}+[\gamma\overline{\theta}(t)]_i$.

Example 7. *Diffeomorphism example:* In the previous examples, we started with σ. Suppose instead we had a diffeomorphism

$$\Lambda(\varphi_1,\varphi_2)=\Lambda_t(\varphi_1,\varphi_2)=\begin{bmatrix}\frac{\pi}{2}+\arcsin(\log\varphi_1\varphi_2-1)\\ \frac{\pi}{2}+\arcsin(\frac{2\varphi_2}{\varphi_1}-1)\end{bmatrix}$$

on $1<\varphi_1\varphi_2<e$, $\varphi_2\le\varphi_1$. Then, the possible full rank σ's satisfy $\sigma=(\nabla_\varphi\Lambda)^{-1}$ i.e.

$$\sigma(\varphi_1,\varphi_2)=\begin{pmatrix}\frac{\varphi_1}{2}\sqrt{2\log\varphi_1\varphi_2-(\log\varphi_1\varphi_2)^2} & -\frac{\varphi_1}{2\varphi_2}\sqrt{\varphi_2(\varphi_1-\varphi_2)}\\ \frac{\varphi_2}{2}\sqrt{2\log\varphi_1\varphi_2-(\log\varphi_1\varphi_2)^2} & -\frac{1}{2}\sqrt{\varphi_2(\varphi_1-\varphi_2)}\end{pmatrix} \tag{3.13}$$

so $(\nabla\Lambda)\sigma = I_2$ and σ satisfies (2.12) by Lemma 1 ii). The possible Stratonovich (time-dependent) drifts $h(\varphi_1, \varphi_2, t)$ are

$$\sigma(\varphi_1,\varphi_2)\begin{pmatrix} \overline{\theta}_1(t) + \overline{\beta}_{11}(t)(\frac{\pi}{2} + \arcsin(\log\varphi_1\varphi_2 - 1)) - \overline{\beta}_{12}(t)(\frac{\pi}{2} + \arcsin(\frac{2\varphi_2}{\varphi_1} - 1)) \\ \overline{\theta}_2(t) + \overline{\beta}_{21}(t)(\frac{\pi}{2} + \arcsin(\log\varphi_1\varphi_2 - 1)) - \overline{\beta}_{22}(t)(\frac{\pi}{2} + \arcsin(\frac{2\varphi_2}{\varphi_1} - 1)) \end{pmatrix} \tag{3.14}$$

while $U_{s,t}, \overline{c}_s$ satisfy the equations at the start of Subsection 3.2.

3.3. Non-square case

The *Extended Heston* model of our companion paper [14] is an important non-square example. We provide a second interesting non-square example herein.

Example 8 *(Heisenberg group).* Let $\overline{x} \in \mathbb{R}^d$ and $\widetilde{x} \in \mathbb{R}$ be the components of the starting point, $A = A(t)$ be a $\mathbb{R}^{d\times d}$ continuously differentiable matrix function and $\sigma(\varphi, t) = \sigma(\xi, z, t) = \begin{bmatrix} I_d \\ (A(t)\xi)^\top \end{bmatrix}$, where $\xi \in \mathbb{R}^d$, $z \in \mathbb{R}$. Then, σ has rank $r = d$. The solution to $dX_t = \sigma(X_t, t)dW_t$ is known as the Brownian motion on the Heisenberg group. Moreover,

$$(\nabla_\varphi \sigma_j)\sigma_k - (\nabla_\varphi \sigma_k)\sigma_j = \begin{bmatrix} 0 \\ A_{jk} - A_{kj} \end{bmatrix}.$$

Therefore, (2.12) holds true if and only if A is symmetric. In this case, one can solve for an explicit solution for an arbitrary starting point $(\overline{x}, \widetilde{x}, s)$. The diffeomorphism $\widehat{\Lambda}(\xi, z, t) = \begin{bmatrix} \Lambda_t(\xi, z) \\ t \end{bmatrix}$ is solved $\Lambda_t(\xi, z) = \begin{bmatrix} \xi \\ g \end{bmatrix}$ with $g(\xi, z, t) = z - \frac{1}{2}\xi^\top A(t)\xi$ following the proof of Proposition 1 in Section 5 (see [14] for details on a more involved example). Hence, $\pi = I_d$, $\widehat{\sigma} = \begin{bmatrix} I_d \\ 0 \end{bmatrix}$, $\overline{\kappa}$ does not exist so $G(t) = I_d$ and $[\nabla\Lambda_t]^{-1} = \begin{bmatrix} I_d & 0 \\ \xi^\top A(t) & 1 \end{bmatrix}$. Now, we can take any functions $\overline{\theta} \in \mathbb{R}^d$, $\overline{\beta} \in \mathbb{R}^{d\times d}$, $\widetilde{h} \in \mathbb{R}$ satisfying the differentiability conditions in Definition 3 and let $\widetilde{X}_t$, $U_{s,t}\widetilde{X}$, $\overline{c}_s(t)$ satisfy:

$$\frac{d}{dt}\widetilde{X}_t = \widetilde{h}(\widetilde{X}_t, t) \text{ s.t. } \widetilde{X}_s = \widetilde{x} - \frac{1}{2}\overline{x}^\top A(s)\overline{x}$$

$$\frac{d}{du}U_{s,u}\widetilde{X} = -(U_{s,u}\widetilde{X})\,\overline{\beta}(\widetilde{X}_u, u) \text{ s.t. } U_{s,s}\widetilde{X} = I_d$$

$$\overline{c}_s(t) = U_{s,t}^{-1}\left\{\overline{x} + \int_0^t U_{s,u}\overline{\theta}(\widetilde{X}_u, u)du\right\}.$$

From Theorem 3 and (1.5), drift b must be of the (quadratic) form

$$b(\xi, z, t) = \begin{bmatrix} \overline{\theta}(\widetilde{X}_t, t) - \overline{\beta}(\widetilde{X}_t, t)\xi \\ \widetilde{h}(\widetilde{X}_t, t) + \xi^\top A(t)\overline{\theta}(\widetilde{X}_t, t) - \xi^\top A(t)\overline{\beta}(\widetilde{X}_t, t)\xi + \frac{1}{2}\xi^\top \frac{d}{dt}A(t)\xi + \frac{1}{2}\mathrm{Tr}\{A(t)\} \end{bmatrix}$$

for some $\overline{\theta}$, $\overline{\beta}$, $\widetilde{h}$. Finally, the corresponding ϕ is given by

$$\phi(y, t) = \begin{bmatrix} \overline{c}_s(t) + (U_{s,t}^{-1}\widetilde{X})y \\ \widetilde{X}_t + \frac{1}{2}(\overline{c}_s(t) + (U_{s,t}^{-1}\widetilde{X})y)^\top A(t)(\overline{c}_s(t) + (U_{s,t}^{-1}\widetilde{X})y) \end{bmatrix}.$$

4. Comparison with the works of Yamato and Kunita

Now, we compare our existence results to those appearing in Yamato [22] and Kunita [18]. In Section III.3 of Kunita's treatise, he considers representations of time-homogeneous Fisk–Stratonovich equations

$$dX_t^x = h(X_t^x)dt + \sigma(X_t^x) \bullet dW_t \tag{4.1}$$

in terms of the flows generated by the vector fields

$$\mathfrak{X}_0(y) = \sum_{i=1}^p h_i(y)\frac{\partial}{\partial y_i} \text{ and } \mathfrak{X}_k(y) = \sum_{i=1}^p \sigma_{ik}(y)\frac{\partial}{\partial y_i}, k = 1, ..., d, \tag{4.2}$$

under conditions imposed on the Lie algebra $L_0(\mathfrak{X}_0, \mathfrak{X}_1, \ldots, \mathfrak{X}_d)$ generated by $\mathfrak{X}_k$, $0 \le k \le d$. In the special case where these vector fields commute, i.e. the Lie bracket $[\mathfrak{X}_k, \mathfrak{X}_j] = 0$ for each $j, k = 0, ..., d$, and the coefficients h_i, σ_{ik} are respectively in C_α^3, C_α^4 (the locally four times continuously differentiable functions whose fourth derivative is α-Hölder continuous), his work gives rise to the composition formula

$$\begin{aligned}(X_t^x)_i &= Exp\,(t\mathfrak{X}_0) \circ Exp\left(W_t^1\mathfrak{X}_1\right) \circ \cdots \circ Exp\left(W_t^d\mathfrak{X}_d\right) \circ \chi_i(x), \\ &= \phi_i(W_t, t)\end{aligned} \tag{4.3}$$

locally. Here, χ_i is the function taking x to its i^{th} component and $Exp\,(u\mathfrak{X}_k)$ is the one parameter group of transformations generated by vector field $\mathfrak{X}_k$, i.e. the unique solution to

$$\frac{d}{du}(f \circ \varphi_u) = \mathfrak{X}_k f(\varphi_u),\ \varphi_0 = x, \qquad \forall f \in C^\infty. \tag{4.4}$$

In fact, to use (4.3), one must solve (4.4) for $k = 0, ..., d$ and $f = \chi_i$, $i = 1, ..., d$. Kunita also goes beyond commutability, even surpassing Yamato [22] in generality by considering the situation where $L_0(\mathfrak{X}_0, ..., \mathfrak{X}_d)$ is only solvable, but the expression replacing (4.3) necessarily becomes more unwieldy.

Our characterization of $\phi^{x,s}$ provided by Theorem 3 provides an alternative to (4.3) that is more amenable to direct calculation. Corollary 1 (to follow) supplies a converse to (4.3) in the sense that if $X_t^{x,s}$ were to have such a functional representation $\phi^{x,s}(W_t, t)$ in terms of Brownian motions only, then the vector fields must commute. This was previously established in Theorem 4.1 of Yamato [22] under C^∞ conditions on both ϕ and the coefficients.

The other advantages of our representations over Kunita's results are:

- We allow time dependent vector fields.
- We decrease the regularity assumptions by imposing weaker differentiability on h and on σ when r is small. The looser regularity on the coefficients requires eschewing Fisk–Stratonovich equations in favor of Itô processes.
- We remove the nilpotency assumptions (for our representations).

To validate the final claim, we take $p = 2$, $d = 1$,

$$\mathfrak{X}_0 = \{\overline{\theta}(x_2) - B(x_2)x_1\}\partial_{x_1} + \widetilde{\theta}(x_2)\partial_{x_2},$$

and $\mathfrak{X}_1 = \partial_{x_1}$. Then $[\mathfrak{X}_0, \mathfrak{X}_1] = B\partial_{x_1}$. Moreover, if $\mathfrak{X}_k = [\mathfrak{X}_0, \mathfrak{X}_{k-1}]$, $k \ge 2$, then $\mathfrak{X}_k = a_k(x_2)\partial_{x_1}$, where $a_{k+1} = \widetilde{\theta}(\partial_{x_2}a_k) + Ba_k$, $k \ge 1$ and $a_1 = 1$. In general, the a_k's will not vanish and thereby the Lie algebra contains an infinite number of linearly independent vector fields. This algebra is solvable but is not nilpotent.

Using Theorem 1, we can also give the converse to Kunita's result, Example III.3.5 in Kunita [18], that is valid under the mild regularity on b, σ, h given at the beginning of the section.

Corollary 1. *Suppose that there exists a domain $\widetilde{D}$ such that the coefficients σ and h are time-homogeneous and Fisk–Stratonovich acceptable on $\tilde{D}_T = \tilde{D} \times (0, T)$. Further, assume that the solution to the Fisk–Stratonovich equation (4.1) has a unique local solution*

$$(X_t^x)_i = Exp\,(t\mathfrak{X}_0) \circ Exp\left(W_t^1 \mathfrak{X}_1\right) \circ \cdots \circ Exp\left(W_t^d \mathfrak{X}_d\right) \circ \chi_i(x)$$

on $0 \le t < \tau_x$ for some positive stopping time τ_x and each $x \in \widetilde{D}$, where $\mathfrak{X}_k$, $k = 0, 1, \ldots, d$ are the vector fields defined in (4.2). Then,

$$[\mathfrak{X}_k, \mathfrak{X}_j] = 0 \text{ on } \widetilde{D} \text{ for each } j, k = 0, \ldots, d.$$

Proof. We find that $X_t^x = \phi(Y_t, t)$ with $U_{s,t} = I$ so $A = 0$ from (2.11) and $\sigma A = 0$. It now follows from Theorem 1 and (2.12), (2.13) that $[\mathcal{X}_k, \mathcal{X}_j] = 0$. □

5. Proofs of the main results

Note: For notational simplicity, we will drop superscripts s and x in the proofs as they are just fixed starting points.

5.1. Uniqueness in Theorem 1 a)

The closure $\overline{D_T}$ of D_T is convex and compact. Further, b, σ can be extended to Lipschitz continuous functions on $\overline{D_T}$ by our C^1-conditions in a) of Theorem 1. Now, we use the proof of Kunita [18, Theorem II.5.2] for uniqueness of (strong) local solutions to the SDE until they leave $\overline{D_T}$.

5.2. Proof of Theorem 1 a) is equivalent to b)

Proof. Using (2.1) and Itô's formula for $X_t = \phi(Y_t, t)$, one finds that for any $1 \le i \le p$,

$$\begin{aligned} d(X_t)_i &= \sum_{m=1}^d \sum_{j=1}^d \partial_{y_m} \phi_i(Y_t, t)(U_{s,t}\phi)_{mj} dW_t^j \\ &+ \left[\partial_t \phi_i(Y_t, t) + \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d \partial_{y_j} \partial_{y_k} \phi_i(Y_t, t)(U_{s,t}\phi\ (U_{s,t}\phi)^\top)_{jk}\right] dt. \end{aligned} \tag{5.1}$$

Now, starting with b) implies a), we have a (x, s, σ, h)-representation $\phi^{x,s}, U_{s,t}$ (that satisfies $\mathcal{C}_1$, $\mathcal{C}_2$, $\mathcal{C}_3$). Using (2.7), (2.8) on (5.1), we find

$$\begin{aligned} d(X_t)_i &= \sigma_i(\phi(Y_t, t), t) dW_t + h_i(\phi(Y_t, t), t) dt \\ &+ \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d \partial_{y_j} \partial_{y_k} \phi_i(Y_t, t)(U_{s,t}\phi\ (U_{s,t}\phi)^\top)_{jk} dt. \end{aligned} \tag{5.2}$$

Moreover,

$$\partial_{y_m}\{\sigma_{ij}(\phi, t)\} = \sum_{n=1}^p \{\partial_{\varphi_n} \sigma_{ij}\}(\phi, t) \partial_{y_m} \phi_n$$

and if (2.7) is true, one obtains

$$\partial_{y_m}\{\sigma_{ij}(\phi,t)\} = \sum_{l=1}^{d} \partial_{y_m}\partial_{y_l}\phi_i \ (U_{s,t}\phi)_{lj}.$$

Abbreviating notation $U_{mk}(\phi,t) = (U_{s,t}\phi)_{mk}$, multiplying the last two equalities by $U_{mk}(\phi,t)$, summing over m and using (2.7) again, one finds that

$$\sum_{n=1}^{p}\{\partial_{\varphi_n}\sigma_{ij}\}(\phi,t)\sigma_{nk}(\phi,t) = \sum_{m=1}^{d}\sum_{l=1}^{d}\partial_{y_m}\partial_{y_l}\phi_i \ U_{lj}(\phi,t)U_{mk}(\phi,t), \tag{5.3}$$

and, taking $k=j$ and summing over j, one has that

$$\sum_{j=1}^{d}\{\nabla_\varphi\sigma_j\}(\phi,t)\sigma_j(\phi,t) = \sum_{l=1}^{d}\sum_{m=1}^{d} (U(\phi,t)U^\top(\phi,t))_{lm}\partial_{y_m}\partial_{y_l}\phi. \tag{5.4}$$

Therefore, if (2.7), (2.8), (2.2) are satisfied, then clearly X_t is a local strong solution to (2.10) by (1.5). Moreover, letting $t \searrow s$, we find by (2.7), (2.8), (2.2) that

$$\sigma(x,s) = \nabla_y\phi^{x,s}(0,s) \ \text{ and } h(x,s) = \partial_t\phi^{x,s}(0,s)$$

so $\sigma, h \in C^1$ by the last part of Definition 1.

To show a) implies b), we suppose X_t is a strong solution to (2.10) on $(s,\tau^{x,s})$. Then, since continuous finite-variation martingales are constant, the (continuous) Itô process $\phi(Y_t,t)$ from (5.1) matches (2.10) if and only if

$$\sigma_{ij}(\phi,t) = \sum_{m=1}^{d}\partial_{y_m}\phi_i \ (U_{s,t}\phi)_{mj}, \quad \forall 1 \le i \le p, \ 1 \le j \le d, \tag{5.5}$$

and

$$b_i(\phi,t) = \partial_t\phi_i + \frac{1}{2}\sum_{j=1}^{d}\sum_{k=1}^{d}\partial_{y_j}\partial_{y_k}\phi_i \ (U_{s,t}\phi(U_{s,t}\phi)^\top)_{jk} \ \forall 1 \le i \le p \tag{5.6}$$

for all $t \in (s,\tau^{x,s})$. Rewriting (5.5) in matrix form, one finds

$$\sigma(\phi(Y_t,t),t) = \{\nabla_y\phi(Y_t,t)\}U_{s,t}\phi, \tag{5.7}$$

and (2.7) is true. Now, we can use (5.4) (which was just shown to be a consequence of (2.7)) to find (5.6) is equivalent to

$$\partial_t\phi = b(\phi,t) - \frac{1}{2}\sum_{k=1}^{d}\{\nabla_\varphi\sigma_k\}(\phi,t)\ \sigma_k(\phi,t) = h(\phi,t), \tag{5.8}$$

using (1.5). Now, (2.8) follows by continuity and (2.6). Letting $t \searrow s$ in (5.7) and (5.8), one finds

$$\sigma(x,s) = \nabla_y\phi^{x,s}(0,s) \ \text{ and } h(x,s) = \partial_t\phi^{x,s}(0,s)$$

so the last part of Definition 1 follows from the C^1 property of h, σ. $\square$

5.3. Switching paths lemma

We will use the following lemma within the proof of Theorem 1 b) and c) equivalence. It is related to the question of when a vector field is the gradient of a scalar field, integration over different paths and exactness of one forms. The spirit of this Lemma is well known. It is stated and proved in the exact manner needed below. There are two reasons why this lemma is necessarily more complicated than one might first expect: i) It is the $\{\widehat{\sigma}_j\}$ not the $\{\widehat{\beta}V_j\}$ (which also involve the function U_t) that commute via (5.9) below. ii) The right hand side of the other commutator condition (5.10) is not zero.

Lemma 2. *Suppose that $\mathcal{N} \subset (-1,\infty)\times\mathbb{R}^d$, $\Delta \subset \mathbb{R}^{p+1}$ are bounded domain with closures $\overline{\mathcal{N}}$, $\overline{\Delta}$; $(0,x)\in\mathcal{N}$ with $x\in\mathbb{R}^p$; $\widehat{h}\in C^1(\overline{\mathcal{N}};\mathbb{R}^{p+1})$, $\widehat{\sigma}\subset C^1(\overline{\mathcal{N}};\mathbb{R}^{(p+1)\times d})$, $A\in C^1(\overline{\mathcal{N}};\mathbb{R}^{d\times d})$ satisfy*

$$(\nabla\widehat{\sigma}_j)\widehat{\sigma}_k - (\nabla\widehat{\sigma}_k)\widehat{\sigma}_j = 0,\ \forall j,k\in\{1,...,d\} \tag{5.9}$$

$$(\nabla\widehat{h})\widehat{\sigma}_k - (\nabla\widehat{\sigma}_k)\widehat{h} = -\widehat{\sigma}A_k,\ \forall k\in\{1,...,d\}; \tag{5.10}$$

and $\widehat{\phi}$ is a solution to

$$\widehat{\phi}(y) = (0,x) + \sum_{i=0}^{d}\int_0^{y_i}\widehat{\beta}(\widehat{\phi}(y_0,...,y_{i-1},u,0_{d-i}))(V_{y_0})_i du,\ \forall y\in\mathcal{N}, \tag{5.11}$$

where $\widehat{\beta} = [\widehat{h}\ \widehat{\sigma}]$ and $V_t = \begin{bmatrix} 1 & 0_d \\ 0^d & U_t^{-1}\end{bmatrix}$ with U_t being the $d\times d$ solution to the linear equation

$$U_t = I + \int_0^t U_u A(\widehat{\phi}(u,0_d))du.$$

Then, $\widehat{\phi}$ also solves

$$\widehat{\phi}(y) = (0,x) + \sum_{i=0}^{d}\int_0^{y_{\pi(i)}}\widehat{\beta}(\widehat{\phi}(\pi^{-1}(y_{\pi(0)},...,y_{\pi(i-1)},u,0_{d-i}))(V_{y_\pi^i(u)})_{\pi(i)}du, \tag{5.12}$$

where

$$y_\pi^i(u) = \begin{cases} y_0 & 0\in\{\pi(0),...,\pi(i-1)\} \\ u & 0=\pi(i) \\ 0 & \text{otherwise}\end{cases}$$

for any $y\in\mathcal{N}$ and permutation π of $\{0,1,...,d\}$ so the integration order does not matter. Here, π^{-1} is an operator re-ordering the arguments to undo the permutation, i.e. to move $y_{\pi(j)}$ from the jth to the $\pi(j)$th position.

Remark 10. In the statement and proof of this lemma we have made time as the first rather than last variable at the request of readers as it seems to be more natural for them in this type of result. It also causes the notation to simplify slightly.

Proof. It follows from its definition that

$$V_t = I - \int_0^t \widehat{A}(\widehat{\phi}(u, 0_d))V_u du \quad \text{with } \widehat{A} = \begin{bmatrix} 0 & 0_d \\ 0^d & A \end{bmatrix}. \tag{5.13}$$

The permutations of $\{0, 1, 2, ..., d\}$ is a symmetric group and any permutation is the composition of at most $\frac{d(d+1)}{2}$ elementary permutations. Hence, we take a permutation π and consider a further elementary permutation $(r, r+1)$ for some $r \in \{0, ..., d-1\}$. The result follows by induction once we show that

$$\begin{aligned} &\int_0^{y_{\pi(r)}} \widehat{\beta}(\widehat{\phi}(\pi^{-1}(y_{\pi^1(0)}, ..., y_{\pi(r-1)}, u, 0_{d-r}))(V_{y_\pi^r(u)})_{\pi(r)} du \\ &+ \int_0^{y_{\pi(r+1)}} \widehat{\beta}(\widehat{\phi}(\pi^{-1}(y_{\pi^1(0)}, ..., y_{\pi(r-1)}, y_{\pi(r)}, u, 0_{d-r-1}))(V_{y_\pi^{r+1}(u)})_{\pi(r+1)} du \\ &= \int_0^{y_{\pi(r+1)}} \widehat{\beta}(\widehat{\phi}(\pi^{-1}(y_{\pi(0)}, ..., y_{\pi(r-1)}, 0, u, 0_{d-r-1}))(V_{y_\pi^{r+1}(u)})_{\pi(r+1)} du \\ &+ \int_0^{y_{\pi(r+1)}} \widehat{\beta}(\widehat{\phi}(\pi^{-1}(y_{\pi^1(0)}, ..., y_{\pi(r-1)}, u, y_{\pi(r+1)}, 0_{d-r-1}))(V_{y_\pi^r(u)})_{\pi(r)} du. \end{aligned} \tag{5.14}$$

(5.14) can be divided into three cases: a) $\pi(0) > r+1$ (when $r < d-1$), b) $\pi(0) < r$ and c) $\pi(0) \in r, r+1$. To ease the notation, we note that showing these three cases is equivalent to assuming that $\widehat{\phi}$ satisfies:

$$\textbf{a:} \qquad \widehat{\phi}(w, z) = \int_0^w \widehat{\sigma}_j(\widehat{\phi}(u, 0))du + \int_0^z \widehat{\sigma}_k(\widehat{\phi}(w, v))dv,$$

$$\textbf{b:} \qquad \widehat{\phi}(w, z) = \int_0^w \widehat{\sigma}(\widehat{\phi}(u, 0))(U_t^{-1})_j du + \int_0^z \widehat{\sigma}(\widehat{\phi}(w, v))(U_t^{-1})_k dv,$$

$$\textbf{c:} \qquad \widehat{\phi}(t, z) = \int_0^t \widehat{h}(\widehat{\phi}(u, 0))du + \int_0^z \widehat{\sigma}(\widehat{\phi}(t, v))(U_t^{-1})_j dv,$$

for $j, k \in \{1, ..., d\}$ and showing the corresponding integration order switch:

$$\textbf{a:} \qquad \widehat{\phi}(w, z) = \int_0^w \widehat{\sigma}_j(\widehat{\phi}(u, z))du + \int_0^z \widehat{\sigma}_k(\widehat{\phi}(0, v))dv,$$

$$\textbf{b:} \qquad \widehat{\phi}(w, z) = \int_0^w \widehat{\sigma}(\widehat{\phi}(u, z))(U_t^{-1})_j du + \int_0^z \widehat{\sigma}(\widehat{\phi}(0, v))(U_t^{-1})_k dv,$$

$$\textbf{c:} \qquad \widehat{\phi}(t, z) = \int_0^t \widehat{h}(\widehat{\phi}(u, z))du + \int_0^z \widehat{\sigma}_j(\widehat{\phi}(0, v))dv.$$

Here, w, z represent two of $\{y_1, ..., y_d\}$ with the others fixed. t is used for y_0.

Case a: This case is subsumed by case b (with $U_t^{-1} = I$) proved below.

Case b: It follows from hypothesis b that $\frac{\partial}{\partial v}\widehat{\phi}(w,v) = \widehat{\sigma}(\widehat{\phi}(w,v))(U_t^{-1})_k$ and

$$\frac{\partial}{\partial w}\widehat{\phi}(w,z) = \widehat{\sigma}(\widehat{\phi}(w,0))(U_t^{-1})_j + \int_0^z \frac{\partial}{\partial w}\widehat{\sigma}(\widehat{\phi}(w,v))(U_t^{-1})_k dv \tag{5.15}$$
$$= \widehat{\sigma}(\widehat{\phi}(w,0))(U_t^{-1})_j + \int_0^z \sum_{i=1}^d \nabla\widehat{\sigma}_i(\widehat{\phi}(w,v))\frac{\partial}{\partial w}\widehat{\phi}(w,v)(U_t^{-1})_{ik}dv.$$

Moreover, it follows from the commutator condition (5.9) that

$$\widehat{\sigma}(\widehat{\phi}(w,z))(U_t^{-1})_j - \widehat{\sigma}(\widehat{\phi}(w,0))(U_t^{-1})_j \tag{5.16}$$
$$= \int_0^z \sum_l \nabla\widehat{\sigma}_l(\widehat{\phi}(w,v))\frac{\partial}{\partial v}\widehat{\phi}(w,v)(U_t^{-1})_{lj}dv$$
$$= \int_0^z \sum_l \nabla\widehat{\sigma}_l(\widehat{\phi}(w,v))\sum_i \widehat{\sigma}_i(\widehat{\phi}(w,v))(U_t^{-1})_{ik}(U_t^{-1})_{lj}dv$$
$$= \int_0^z \sum_i \sum_l \nabla\widehat{\sigma}_i(\widehat{\phi}(w,v))\widehat{\sigma}_l(\widehat{\phi}(w,v))(U_t^{-1})_{ik}(U_t^{-1})_{lj}dv$$
$$= \int_0^z \sum_i \nabla\widehat{\sigma}_i(\widehat{\phi}(w,v))\widehat{\sigma}(\widehat{\phi}(w,v))(U_t^{-1})_j(U_t^{-1})_{ik}dv.$$

Therefore, it follows by (5.15), (5.16) and Gronwall's inequality that

$$\frac{\partial}{\partial w}\widehat{\phi}(w,z) = \widehat{\sigma}(\widehat{\phi}(w,z))(U_t^{-1})_j. \tag{5.17}$$

Finally, by the commutator condition (5.9) again and (5.17)

$$\int_0^w [\widehat{\sigma}(\widehat{\phi}(u,z)) - \widehat{\sigma}(\widehat{\phi}(u,0))](U_t^{-1})_j du \tag{5.18}$$
$$= \int_0^w\int_0^z \sum_i \nabla\widehat{\sigma}_i(\widehat{\phi}(u,v))\frac{\partial}{\partial v}\widehat{\phi}(u,v)(U_t^{-1})_{ij}dvdu$$
$$= \int_0^w\int_0^z \sum_{i,l} \nabla\widehat{\sigma}_i(\widehat{\phi}(u,v))\widehat{\sigma}_l(\widehat{\phi}(u,v))(U_t^{-1})_{lk}(U_t^{-1})_{ij}dvdu$$
$$= \int_0^z\int_0^w \sum_{i,l} \nabla\widehat{\sigma}_l(\widehat{\phi}(u,v))\widehat{\sigma}_i(\widehat{\phi}(u,v))(U_t^{-1})_{ij}(U_t^{-1})_{lk}dudv$$
$$= \int_0^z\int_0^w \sum_l \nabla\widehat{\sigma}_l(\widehat{\phi}(u,v))\frac{\partial}{\partial u}\widehat{\phi}(u,v)(U_t^{-1})_{lk}dudv$$

$$= \int_0^z [\widehat{\sigma}(\widehat{\phi}(w,v)) - \widehat{\sigma}(\widehat{\phi}(0,v))](U_t^{-1})_k dv$$

and Case b follows by rearrangement.

Case c: It follows from hypothesis c and (5.13) that $\frac{\partial}{\partial v}\widehat{\phi}(t,v) = \widehat{\sigma}(\widehat{\phi}(t,v))(U_t^{-1})_j$ and

$$\begin{aligned}\frac{\partial}{\partial t}\widehat{\phi}(t,z) &= \widehat{h}(\widehat{\phi}(t,0)) + \int_0^z \frac{\partial}{\partial t}\left\{\widehat{\sigma}(\widehat{\phi}(t,v))(U_t^{-1})_j\right\} dv \\ &= \widehat{h}(\widehat{\phi}(t,0)) + \int_0^z \sum_i \nabla\widehat{\sigma}_i(\widehat{\phi}(t,v))\frac{\partial}{\partial t}\widehat{\phi}(t,v)(U_t^{-1})_{ij} dv \\ &\quad - \sum_i \int_0^z \widehat{\sigma}_i(\widehat{\phi}(t,v))dv \sum_l A_{il}(\widehat{\phi}(t,0))(U_t^{-1})_{lj} dv.\end{aligned} \tag{5.19}$$

Moreover, it follows from the commutator condition (5.10) that

$$\begin{aligned}\widehat{h}(\widehat{\phi}(t,z)) - \widehat{h}(\widehat{\phi}(t,0)) &= \int_0^z \nabla\widehat{h}(\widehat{\phi}(t,v))\frac{\partial}{\partial v}\widehat{\phi}(t,v)dv \\ &= \int_0^z \sum_l \nabla\widehat{h}(\widehat{\phi}(t,v))\widehat{\sigma}_l(\widehat{\phi}(t,v))(U_t^{-1})_{lj} dv \\ &= \int_0^z \sum_l \nabla\widehat{\sigma}_l(\widehat{\phi}(t,v))\widehat{h}(\widehat{\phi}(t,v))(U_t^{-1})_{lj} dv \\ &\quad - \int_0^z \sum_l \widehat{\sigma}(\widehat{\phi}(t,v))A(\widehat{\phi}(t,0))_l(U_t^{-1})_{lj} dv.\end{aligned} \tag{5.20}$$

Therefore, it follows by (5.19), (5.20) and Gronwall's inequality that

$$\frac{\partial}{\partial t}\widehat{\phi}(t,z) = \widehat{h}(\widehat{\phi}(t,z)). \tag{5.21}$$

Finally, by the commutator condition (5.10) again, (5.13) and (5.21)

$$\begin{aligned}&\int_0^t \widehat{h}(\widehat{\phi}(u,z)) - \widehat{h}(\widehat{\phi}(u,0))du \\ &= \int_0^t\int_0^z \nabla\widehat{h}(\widehat{\phi}(u,v))\frac{\partial}{\partial v}\widehat{\phi}(u,v)dvdu \\ &= \int_0^t\int_0^z \sum_k \nabla\widehat{h}(\widehat{\phi}(u,v))\widehat{\sigma}_k(\widehat{\phi}(u,v))(U_u^{-1})_{kj} dvdu \\ &= \int_0^z\int_0^t \sum_k \nabla\widehat{\sigma}_k(\widehat{\phi}(u,v))\widehat{h}(\widehat{\phi}(u,v))(U_u^{-1})_{kj} dudv\end{aligned} \tag{5.22}$$

$$
\begin{aligned}
&- \int_0^z \int_0^t \sum_k \widehat{\sigma}(\widehat{\phi}(u,v)) A_k(\widehat{\phi}(u,0))(U_u^{-1})_{kj} du dv \\
&= \int_0^z \int_0^t \sum_k \nabla \widehat{\sigma}_k(\widehat{\phi}(u,v)) \frac{\partial}{\partial u} \widehat{\phi}(u,v)(U_u^{-1})_{kj} du dv \\
&+ \int_0^z \int_0^t \widehat{\sigma}(\widehat{\phi}(u,v)) \frac{\partial}{\partial u}(U_u^{-1})_j du dv \\
&= \int_0^z [\widehat{\sigma}(\widehat{\phi}(t,v))(U_t^{-1})_j - \widehat{\sigma}_j(\widehat{\phi}(0,v))] dv
\end{aligned}
$$

and Case c follows by rearrangement. □

5.4. Proof of Theorem 1 b) is equivalent to c)

Proof. Step 1: Show that (c) implies (b).
Let $\mathcal{N}^{x,s}$ be a open ball centered at $0 \in \mathbb{R}^d$ (whose radius can depend upon starting point (x,s)), $t_0 > s$ and $\mathcal{N} = \mathcal{N}^{x,s} \times (s,t_0)$. Next, we define successive approximations to $\phi^{x,s}$, $U_{s,t}\phi$ by the path integral and linear equation

$$
\begin{aligned}
\phi^{n+1}(y,t) = x + \int_s^t h(\phi^n(0_d,u),u) du \qquad (5.23)\\
+ \sum_{i=0}^d \int_0^{y_i} \sigma(\phi^n(y_1,...,y_{i-1},u,0_{d-i},t),t)(U^n_{s,t})_i^{-1} du,
\end{aligned}
$$

$$
U_{s,t}^{n+1} = I + \int_s^t U_{s,u}^{n+1}\ A(\phi^n(0_d,u),u) du, \qquad (5.24)
$$

starting with $\phi^0(y,t) = x$. (Note that $U^n_{s,t}$'s inverse $B^n_{s,t}$ exist and satisfies $B^n_{s,t} = I - \int_s^t A(\phi^{n-1}(0,u),u) \times B^n_{s,u} du$.) Let L be the integral operator corresponding to (5.23), (5.24) so that

$$
(\phi^{n+1}, U^{n+1}) = L(\phi^n, U^n).
$$

Then, it is well known and easy to verify that the iterated operator L^m is a contraction on $C(\mathcal{N};\mathbb{R}^p) \times C([s,t_0);\mathbb{R}^{d\times d})$ with supremum norm for some $m \in \mathbb{N}$. Hence, (ϕ^n, U^n) converges as $n \to \infty$ to some unique fixed point $(\phi^{x,s}, U_s)$ satisfying

$$
(\phi^{x,s}, U_s) = L(\phi^{x,s}, U_s),
$$

i.e. for each $(y,t) \in \mathcal{N}$

$$
\phi^{x,s}(y,t) = x + \int_s^t h(\phi^{x,s}(0_d,u),u) du \qquad (5.25)
$$

$$+\sum_{i=0}^{d}\int_0^{y_i}\sigma(\phi^{x,s}(y_1,...,y_{i-1},u,0_{d-i},t),t)(U_{s,t})_i^{-1}du,$$

$$U_{s,t}=I+\int_s^t U_{s,u}\ A(\phi^{x,s}(0_d,u),u)du. \tag{5.26}$$

Hence,

$$\frac{\partial}{\partial t}U_{s,t}^{-1}=-A(\phi^{x,s}(0_d,t),t)U_{s,t}^{-1}. \tag{5.27}$$

$\mathcal{C}_2$, $\mathcal{C}_3$ are true by our construction. Moreover, $\mathcal{C}_1$ and the continuity of $\partial_s\nabla_y\phi^{x,s}(0_d,s)$, $\partial_s\partial_t\phi^{x,s}(0_d,s)$, $\partial_{x_i}\nabla_y\phi^{x,s}(0_d,s)$ and $\partial_{x_i}\partial_t\phi^{x,s}(0_d,s)$ will follow from (5.25) and the conditions on h and σ once we have established $\phi^{x,s}$ satisfies (2.7), (2.8).

It remains to show that $\phi^{x,s}$ satisfies (2.7), (2.8). The fundamental theorem of calculus immediately tells us that $\frac{\partial}{\partial y_d}\phi^{x,s}(y,t)=\sigma(\phi^{x,s}(y,t))(U_{s,t})_d^{-1}$. We use a different path to have access to the other partial derivatives. Clearly, (5.25) is equivalent to

$$\begin{pmatrix}\phi^{x,s}(y,t)\\ t\end{pmatrix}=\begin{pmatrix}x\\ s\end{pmatrix}+\int_s^t\begin{pmatrix}h(\phi^{x,s}(0_d,u),u)\\ 1\end{pmatrix}du \tag{5.28}$$
$$+\sum_{i=0}^{d}\int_0^{y_i}\begin{pmatrix}\sigma(\phi^{x,s}(y_1,...,y_{i-1},u,0_{d-i},t),t)(U_{s,t})_i^{-1}\\ 0_d\end{pmatrix}du$$

and we can define new coefficients corresponding to this enlarged equation:

$$\widehat{h}(\widehat{\varphi})=\begin{pmatrix}h(\varphi,t)\\ 1\end{pmatrix},\quad \widehat{\sigma}_j(\widehat{\varphi})=\begin{pmatrix}\sigma(\varphi,t)\\ 0_d\end{pmatrix}, \tag{5.29}$$

where $\widehat{\varphi}=\begin{pmatrix}\varphi\\ t\end{pmatrix}$. One finds the commutator conditions (2.12), (2.13) are equivalent to

$$(\nabla\widehat{\sigma}_k)\widehat{\sigma}_j-(\nabla\widehat{\sigma}_j)\widehat{\sigma}_k=0,\ \forall j,k\in\{1,...,d\} \tag{5.30}$$

$$(\nabla\widehat{\sigma}_0)\widehat{\sigma}_j-(\nabla\widehat{\sigma}_j)\widehat{h}=-\widehat{\sigma}A_j,\ \forall j\in\{1,...,d\}, \tag{5.31}$$

which means we can use Lemma 2 (with time shifted functions $\widehat{h}(\cdot,s+\cdot)$, $\widehat{\sigma}(\cdot,s+\cdot)$, $A(\cdot,s+\cdot)$, $U_{s,s+\cdot}$ and $\widehat{\phi}^{x,s}(\cdot,s+\cdot)$) to move the path segment of the desired partial derivative to the end and find

$$\frac{\partial}{\partial y_i}\phi^{x,s}(y,t)=\sigma(\phi^{x,s}(y,t),t)(U_{s,t})_i^{-1},\ \frac{\partial}{\partial t}\phi^{x,s}(y,t)=h(\phi^{x,s}(y,t),t). \tag{5.32}$$

Step 2: Show that (2.7) implies (2.12).
By $\mathcal{C}_1$ and (2.7), one has that

$$\partial_{y_j}\{\sigma(\phi,t)(U_{s,t}^{-1}\phi)_k\}=\partial_{y_j}\partial_{y_k}\phi(y,t)=\partial_{y_k}\partial_{y_j}\phi(y,t)=\partial_{y_k}\{\sigma(\phi,t)(U_{s,t}^{-1}\phi)_j\}. \tag{5.33}$$

However, it follows by (2.3) of $\mathcal{C}_3$ and then (2.7) that

$$\begin{aligned}&\partial_{y_j}\{\sigma(\phi,t)(U_{s,t}^{-1}\phi)_k\} \\ &=\sum_m \{\partial_{y_j}\sigma_m(\phi,t)\}(U_{s,t}^{-1}\phi)_{mk} \\ &=\sum_m \nabla_\phi\sigma_m(\phi,t)\sigma(\phi,t)(U_{s,t}^{-1}\phi)_j(U_{s,t}^{-1}\phi)_{mk} \\ &=\sum_m\sum_n \nabla_\phi\sigma_m(\phi,t)\sigma_n(\phi,t)(U_{s,t}^{-1}\phi)_{nj}(U_{s,t}^{-1}\phi)_{mk}\end{aligned} \tag{5.34}$$

and similarly

$$\begin{aligned}&\partial_{y_k}\{\sigma(\phi,t)(U_{s,t}^{-1}\phi)_j\} \\ &=\sum_n\sum_m \nabla_\phi\sigma_n(\phi,t)\sigma_m(\phi,t)(U_{s,t}^{-1}\phi)_{mk}(U_{s,t}^{-1}\phi)_{nj}\,.\end{aligned} \tag{5.35}$$

Letting $t \searrow s$ in (5.34) and (5.35), one finds by (5.33) that for all $1 \le j,k \le d$,

$$\begin{aligned}&\sum_m\sum_n \nabla_x\sigma_m(x,s)\sigma_n(x,s)(U_{s,s}^{-1}\phi)_{nj}(U_{s,s}^{-1}\phi)_{mk} \\ &=\lim_{t\searrow s}\partial_{y_j}\{\sigma(\phi,t)(U_{s,t}^{-1}\phi)_k\} \\ &=\lim_{t\searrow s}\partial_{y_k}\{\sigma(\phi,t)(U_{s,t}^{-1}\phi)_j\} \\ &=\sum_m\sum_n \nabla_x\sigma_n(x,s)\sigma_m(x,s)(U_{s,s}^{-1}\phi)_{nj}(U_{s,s}^{-1}\phi)_{mk}.\end{aligned} \tag{5.36}$$

However, $U_{s,s}^{-1}\phi = I$ so we have that

$$(\nabla_x\sigma_q)(x,s)\sigma_p(x,s) = (\nabla_x\sigma_p)(x,s)\sigma_q(x,s).$$

Hence, (2.12) holds.
Step 3: Show that (2.7), (2.8) imply (2.13).
By $\mathcal{C}_1$ and (2.7), (2.8), one has that

$$\frac{d}{dt}\{\sigma(\phi,t)(U_{s,t}^{-1}\phi)_k\} = \partial_{y_k}h(\phi,t). \tag{5.37}$$

One gets by (2.7) that

$$\partial_{y_k}h(\phi,t) = \nabla_\phi h(\phi,t)\partial_{y_k}\phi(y,t) = \nabla_\phi h(\phi,t)\sigma(\phi,t)(U_{s,t}^{-1}\phi)_k \tag{5.38}$$

and by the chain rule, (5.37), (2.8) as well as the standard formula $\frac{d}{dt}B_t^{-1} = -B_t^{-1}\left(\frac{d}{dt}B_t\right)B_t^{-1}$ that

$$\begin{aligned}\partial_{y_k}h(\phi,t) &= \frac{d}{dt}\{\sigma(\phi,t)(U_{s,t}^{-1}\phi)_k\} \\ &=\sum_m \nabla_\phi\sigma_m(\phi,t)h(\phi,t)(U_{s,t}^{-1}\phi)_{mk} + \partial_t\sigma(\phi,t)(U_{s,t}^{-1}\phi)_k \\ &\quad -\sigma(\phi,t)U_{s,t}^{-1}\phi\sum_m\frac{d}{dt}(U_{s,t}\phi)_m(U_{s,t}^{-1}\phi)_{mk}.\end{aligned} \tag{5.39}$$

Combining (5.38), (5.39), multiplying by $(U_{s,t}\phi)_{kn}$ and summing, we get

$$\nabla_\phi h(\phi,t)\sigma_n(\phi,t) = \nabla_\phi \sigma_n(\phi,t)h(\phi,t) + \partial_t \sigma_n(\phi,t) - \sigma(\phi,t)U_{s,t}^{-1}\phi^{x,s}\frac{d}{dt}(U_{s,t}\phi^{x,s})_n \tag{5.40}$$

so, letting $t \searrow s$ and using (2.11), (2.2), one arrives at (2.13). □

5.5. Proof of Proposition 1

Our methods are motivated in part by Brickell and Clark [2, Propositions 8.3.2 and 11.5.2]. By reducing $T > 0$ if necessary, we can find a permutation π such that the first r columns of $\sigma^\pi = \sigma\pi$ are linearly independent on D_T.

Proof. Fix $(x, s) \in D_T$ and $t \in [s, T)$. Λ_t will have form:

$$\Lambda_t = \Lambda_t^{r,1}, \quad \text{where} \quad \Lambda_t^{i,1} = \Lambda_t^i \circ \Lambda_t^{i-1} \circ \cdots \circ \Lambda_t^2 \circ \Lambda_t^1, \tag{5.41}$$

$$\Lambda_t^i(\varphi) = \sum_{j=1}^{i-1} \varphi_j e_j + H^i(\varphi_i, ..., \varphi_p, t). \tag{5.42}$$

Here, Λ_t^i is a C^{r+2-i}-diffeomorphism on a neighborhood $O^{x_t^{i-1}}$ of $x_t^{i-1} = \Lambda_t^{i-1,1}(x)$ so $\Lambda_t : O_t^x \to \mathbb{R}^p$ for some neighborhood O_t^x of x.

To construct Λ_t^i recursively, we suppose $\widehat{\sigma}_j = e_j$ for $j < i$ and

$$\alpha_i \doteq \{\nabla \Lambda_t^{i-1,1} \sigma_i^\pi\} \circ (\Lambda_t^{i-1,1})^{-1} \tag{5.43}$$

does not depend upon $\varphi_1, ..., \varphi_{i-1}$, which are vacuously true when $i = 1$. Moreover, without loss of generality, we assume the i^{th} component of α_i satisfies $\alpha_{i,i} \neq 0$ (or else we change π by permuting columns $i, ..., d$ of σ^π). Set $\psi_t^i(\varphi) = \theta_t(\varphi_i - x_{t,i}^{i-1}; \varphi_1, ..., \varphi_{i-1}, x_{t,i}^{i-1}, \varphi_{i+1}, ..., \varphi_p)$, where θ satisfies $\theta_t(0; \varphi) = \varphi$ and $\frac{d}{du}\theta_t(u; \varphi) = \alpha_i(\theta_t(u; \varphi), t)$ for $u \in I^\varphi$, an open interval containing 0, and φ in a neighborhood containing x_t^{i-1}. Then, $\partial_{\varphi_i}\psi^i = \alpha_i(\psi^i)$. For $j \neq i$, we have $\partial_{\varphi_j}\psi_t^i(\varphi) = \partial_{\varphi_j}\theta_t(\varphi_i - x_{t,i}^{i-1}; \varphi_1, ..., \varphi_{i-1}, x_{t,i}^{i-1}, \varphi_{i+1}, ..., \varphi_p)$ and

$$\partial_u \partial_{\varphi_j}\theta_t(u; \varphi) = \partial_{\varphi_j}\alpha_i(\theta(u; \varphi), t) \ \text{ s.t. } \ \partial_{\varphi_j}\theta_t(0; \varphi) = e_j$$

so $\nabla\psi_t^i(x_t^{i-1})$ has determinant $\alpha_{i,i}(x_t^{i-1}, t) \neq 0$. Thus by the Inverse Function Theorem, ψ_t^i has inverse $\Lambda_t^i \in C^{r+2-i}(O^{x_t^{i-1}}, \mathbb{R}^p)$ and $\nabla\Lambda_t^i = [\nabla\psi_t^i]^{-1}(\Lambda_t^i)$ on neighborhood $O^{x_t^{i-1}} = \psi_t^i(U^{x_t^{i-1}})$ of x_t^{i-1} with

$$U^{x_t^{i-1}} = \left\{\varphi : \|\nabla\psi_t^i(\varphi) - \nabla\psi_t^i(x_t^{i-1})\| < \frac{1}{2\|(\nabla\psi_t^i(x_t^{i-1}))^{-1}\|}\right\}, \tag{5.44}$$

and $\|\cdot\|$ being Frobenius norm. Hence, $\nabla\Lambda_t^i((\Lambda_t^i)^{-1})\nabla\psi_t^i = I$ and

$$\widehat{\sigma}_i = \{\nabla\Lambda_t^i \alpha_i\}(\Lambda_t^i)^{-1} = e_i \in \mathbb{R}^p. \tag{5.45}$$

Moreover, Λ_t^i has the form (5.42) if ψ_t^i has similar form. ψ_t^i has this form by its definition as well as the facts α_i is locally Lipschitz and does not depend upon $\varphi_1, ..., \varphi_{i-1}$. (5.42) and induction then imply that

$$e_k = \widehat{\sigma}_k = \{\nabla\Lambda_t^{i,1}\sigma_k^\pi\} \circ (\Lambda_t^{i,1})^{-1} \ \forall k \leq i.$$

Next,

$$(\nabla\widehat{\sigma}_j)\widehat{\sigma}_k - (\nabla\widehat{\sigma}_k)\widehat{\sigma}_j = (\nabla\sigma^\pi_j)\sigma^\pi_k - (\nabla\sigma^\pi_k)\sigma^\pi_j = 0 \ \ \forall \ \ 1 \le k, j \le d \tag{5.46}$$

by Lemma 1. Now, since $\widehat{\sigma}_k = e_k \in \mathbb{R}^p$ for $1 \le k \le i$, (5.46) implies

$$(\nabla\widehat{\sigma}_j)e_k = (\nabla\widehat{\sigma}_j)e_k - (\nabla e_k)\widehat{\sigma}_j = 0 \ \ \forall\, 1 \le k \le i < j$$

on a neighborhood O of x. Therefore, $\widehat{\sigma}_j$ and (by a similar argument) α_{i+1} can not depend upon $\varphi_1, \ldots, \varphi_i$ so we can take $i = r$ by induction and

$$\widehat{\sigma} = \{(\nabla\Lambda_t)\sigma^\pi\} \circ \Lambda_t^{-1} = \begin{pmatrix} I_r & \overline{\kappa} \\ 0 & \widetilde{\kappa} \end{pmatrix} \in \mathbb{R}^{p+1\times d} \text{ on } \Lambda_t(O^x),$$

where $\overline{\kappa} \in \mathbb{R}^{r\times(d-r)}$ and $\widetilde{\kappa} \in \mathbb{R}^{(p-r)\times(d-r)}$ do not depend on the variables $\varphi_1, \ldots, \varphi_r$. Since $\widehat{\sigma}$ has also rank r, it follows that $\widetilde{\kappa} = 0$.

It remains to show there is a relatively open $O^{x,s} \subset D_T$ containing (x, s) such that $(\varphi, t) \to \Lambda_t(\varphi)$ is twice continuously differentiable on $O^{x,s}$. The desired differentiability of ψ^i, α_i and Λ^i follow from their definitions and [∂]. For $V \subset \mathbb{R}^p$ and $\gamma > 0$, we let $V^\gamma = \{v + \underline{\gamma} : v \in V, \underline{\gamma} \in \mathbb{R}^p \text{ with } |\underline{\gamma}| \le \gamma\}$.

We let $O^{x,s} = O^{x,s}_r$, $I^i_s = (s - t^i, s + t^i) \cap [0, T)$, where t^i, $O^{x,s}_i$ are found recursively, $O^{x,s}_i \subset D_T$ is relatively open, contains $\{x\} \times I^i_s$ and $\Lambda^{i,1} : O^{x,s}_i \to \mathbb{R}^p$. Let $t^0 = T$, $U^{x,s}_0 = B_x(1)$, $O^{x,s}_0 = B_x(1) \times I^0_s$ and for $i = 1, 2, ..., r$ define

$$\begin{aligned}
K^i &= \sup_{(\varphi,t)\in O^{x,s}_{i-1}} (|\Lambda^{i-1,1}_t(\varphi) - \Lambda^{i-1,1}_t(x)|/|\varphi - x|) \vee 1 \\
L^i &= \sup_{(\varphi,t)\in O^{x,s}_{i-1}} (\|\nabla\psi^i_t(\Lambda^{i-1,1}_t(\varphi)) - \nabla\psi^i_s(\Lambda^{i-1,1}_s(\varphi))\|/|t - s|) \vee 1 \\
M^i &= \sup_{(\varphi,t)\in O^{x,s}_{i-1}} (|\psi^i_t(\Lambda^{i-1,1}_t(\varphi)) - \psi^i_t(x^{i-1}_t)|/|\varphi - x|) \vee 1 \\
N^i &= \sup_{(\varphi,t)\in O^{x,s}_{i-1}, |\underline{\gamma}^i|\le\gamma^i} (|\Lambda^i_t(\psi^i_t(\Lambda^{i-1,1}_t(\varphi)) + \underline{\gamma}^i) - \Lambda^i_t(\psi^i_t(\Lambda^{i-1,1}_t(\varphi)))|/\gamma^i) \vee 1 \\
U^{x,s}_i &= \left\{\varphi \in U^{x,s}_{i-1} : \|\nabla\psi^i_s(\Lambda^{i-1,1}_s(\varphi)) - \nabla\psi^i_s(x^{i-1}_s)\| < \frac{1}{8\|(\nabla\psi^i_s(x^{i-1}_s))^{-1}\|}\right\}
\end{aligned}$$

recursively. ($\underline{\gamma}^i$ is a vector of size γ^i so the supremum in N^i is over vectors below this size.) $U^{x,s}_i$ must contain a ball $B_x((K^i + 1)\varepsilon)$ for some $\varepsilon = \varepsilon^i > 0$. Let $\gamma^i = \frac{\varepsilon}{N^i}$, $\delta^i = \min\left\{\frac{\gamma^1}{M^1}, ..., \frac{\gamma^i}{M^i}\right\}$, $O^{x,s}_i = B_x(\delta^i) \times I^i_s$ and $0 < t^i < t^{i-1}$ be such that $\sup_{t\in I^i_s} \|(\nabla\psi^i_t(x^{i-1}_t))^{-1}\| \le 2\|(\nabla\psi^i_s(x^{i-1}_s))^{-1}\|$ and $L^i t^i < \dfrac{1}{16\|(\nabla\psi^i_s(x^{i-1}_s))^{-1}\|}$. We need only show that $\psi^i_s(\Lambda^{i-1,1}_s(B_x(\delta^i))) \subset \psi^i_t(U^{x^{i-1}_t})$ for all $t \in I^i_s$.

First, $(\psi^i_t(\Lambda^{i-1,1}_t(B_x(\varepsilon))))^{\gamma^i} \subset \psi^i_t(\Lambda^{i-1,1}_t(B_x((K^i + 1)\varepsilon)))$ for $t \in I_s$ follows by considering $\varphi \in B_x(\varepsilon)$ and

$$\begin{aligned}
|\Lambda^i_t(\psi^i_t(\Lambda^{i-1,1}_t(\varphi)) + \underline{\gamma}^i) - x^{i-1}_t| &\le |\Lambda^{i-1,1}_t(\varphi) - x^{i-1}_t| \\
&\quad + |\Lambda^i_t(\psi^i_t(\Lambda^{i-1,1}_t(\varphi)) + \underline{\gamma}^i) - \Lambda^i_t(\psi^i_t(\Lambda^{i-1,1}_t(\varphi)))| \\
&< K^i\varepsilon + N^i\gamma^i < (K^i + 1)\varepsilon.
\end{aligned}$$

Now, $\psi^i_t(\Lambda^{i-1,1}_t(B_x(\delta^i))) \subset (\psi^i_t(\Lambda^{i-1,1}_t(B_x(\varepsilon))))^{\gamma^i}$ for $t \in I^i_s$ since $\varphi \in B_x(\delta^i)$ implies

$$|\psi_t^i(\Lambda_t^{i-1,1}(\varphi)) - \psi_t^i(x_t^{i-1})| < M^i\delta^i \le \gamma^i$$

and $\varphi \in (\psi_t^i(\Lambda_t^{i-1,1}(B_x(\varepsilon)))^{\gamma^i}$.

Finally, $\psi_t^i(\Lambda_t^{i-1,1}(U_i^{x,s})) \subset \psi_t^i(U^{x_t^{i-1}})$ for $t \in I_s^i$, since $\varphi \in U_i^{x,s}$ implies

$$\begin{aligned}\|\nabla\psi_t^i(\Lambda_t^{i-1,1}(\varphi)) - \nabla\psi_t^i(x_t^{i-1})\| &\le \|\nabla\psi_t^i(\Lambda_t^{i-1,1}(\varphi)) - \nabla\psi_s^i(\Lambda_s^{i-1,1}(\varphi))\| \\ &\quad + \|\nabla\psi_s^i(\Lambda_s^{i-1,1}(\varphi)) - \nabla\psi_s^i(x_s^{i-1})\| \\ &\quad + \|\nabla\psi_s^i(x_s^{i-1}) - \nabla\psi_t^i(x_t^{i-1})\| \\ &\le 2L^i|t-s| + \frac{1}{8\|(\nabla\psi_s^i(x_s^{i-1}))^{-1}\|} \\ &< \frac{1}{4\|(\nabla\psi_s^i(x_s^{i-1}))^{-1}\|} \\ &< \frac{1}{2\|(\nabla\psi_t^i(x_t^{i-1}))^{-1}\|}\end{aligned}$$

for $t \in I_s$ so $\varphi \in U^{x_t^{i-1}}$. $\square$

References

[1] G. Ben Arous, Flots et series de Taylor stochastiques, Probab. Theory Related Fields 81 (1989) 29–77.
[2] F. Brickell, R.S. Clark, Differentiable Manifolds, Van Nostrand Reinhold, London, 1970.
[3] F. Castell, Asymptotic expansion of stochastic flows, Probab. Theory Related Fields 96 (1993) 225–239.
[4] F. Castell, J. Gaines, The ordinary differential equation approach to asymptotically efficient schemes for solution of stochastic differential equations, Ann. Inst. Henri Poincaré Probab. Stat. 32 (2) (1996) 231–250.
[5] H. Doss, Liens entre équations différentielles stochastiques et ordinaires, Ann. Inst. Henri Poincaré Probab. Stat. 13 (1977) 99–125.
[6] S.N. Ethier, T.G. Kurtz, Markov Processes: Characterization and Convergence, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, 1986.
[7] A. Friedman, Stochastic Differential Equations and Applications, Dover Publications Inc., Mineola, NY, 2006, Reprint of the 1975 and 1976 original published in two volumes.
[8] S. Heston, A closed-form solutions for options with stochastic volatility with applications to bond and currency options, Rev. Financ. Stud. 6 (1993) 327–343.
[9] Y.-Z. Hu, Série de Taylor stochastique et formule de Campbell–Hausdorff, d'après Ben Arous, Lecture Notes in Math., vol. 1526, 1992, pp. 579–586.
[10] I. Karatzas, S.E. Shreve, Brownian Motion and Stochastic Calculus, Springer, New York, 1987.
[11] P.E. Kloeden, E. Platen, Numerical Solution of Stochastic Differential Equations, Springer, Berlin, 1992.
[12] M. Kouritzin, On exact filters for continuous signals with discrete observations, IEEE Trans. Automat. Control 43 (1998) 709–715.
[13] M. Kouritzin, Exact infinite dimensional filters and explicit solutions, in: L. Gorostiza, G. Ivanoff (Eds.), Stochastic Models, a Volume in Honour of Donald A. Dawson, in: Canadian Mathematical Society Conference Proceedings, vol. 26, American Mathematical Society, Providence, RI, 2000, pp. 265–282.
[14] M.A. Kouritzin, Explicit Heston solutions and stochastic approximation for path-dependent option pricing, Int. J. Theor. Appl. Finance 21 (2018) 1850006, https://doi.org/10.1142/S0219024918500061.
[15] M.A. Kouritzin, D. Li, On explicit solutions to stochastic differential equations, Stoch. Anal. Appl. 18 (4) (2000) 571–580.
[16] M.A. Kouritzin, A. MacKay, Branching Particle Pricers with Heston Examples, Working paper, 2018.
[17] M.A. Kouritzin, A. MacKay, VIX-linked fees for GMWBS via explicit solution simulation methods, Insurance Math. Econom. 81 (2018) 1–17, https://doi.org/10.1016/j.insmatheco.2018.04.001.
[18] H. Kunita, Stochastic differential equations and stochastic flows of diffeomorphisms, in: École d'été de probabilités de Saint-Flour, XII—1982, in: Lecture Notes in Math., vol. 1097, Springer, Berlin, 1984, pp. 143–303.
[19] S.J.A. Malham, A. Wiese, Stochastic expansions and Hopf algebras, 2009.
[20] P.E. Protter, Stochastic Integration and Differential Equations, in: Stochastic Modelling and Applied Probability, second edition, in: Applications of Mathematics (New York), vol. 21, Springer-Verlag, Berlin, 2004.
[21] H. Sussmann, On the gap between deterministic and stochastic differential equations, Ann. Probab. 6 (1978) 19–41.
[22] Y. Yamato, Stochastic differential equations and nilpotent Lie algebras, Z. Wahrsch. Verw. Gebiete 47 (2) (1979) 213–229.