EXPLICIT STRONG SOLUTIONS OF MULTIDIMENSIONAL STOCHASTIC DIFFERENTIAL EQUATIONS

BY MICHAEL A. KOURITZIN AND BRUNO RÉMILLARD

University of Alberta and HEC Montréal

ABSTRACT. Herein, we characterize strong solutions of multidimensional stochastic differential equations $dX_t^{x_0} = b(X_t^{x_0}, t)dt + \sigma(X_t^{x_0}, t)dW_t$, $X_0^{x_0} = x_0$ that can be represented locally as $X_t^{x_0} = \phi^{x_0}\left(t, \int_0^t U^{x_0}(u)dW_u\right)$, where W is an multidimensional Brownian motion and U, ϕ are continuous functions. Assuming that σ is continuously differentiable, we find that σ must satisfy a commutation relation for such explicit solutions to exist and we identify all drift terms b as well as U and ϕ that will allow X to be represented in this manner. Our method is based on the existence of a local change of coordinates in terms of a diffeomorphism between the solutions X and the strong solutions to a simpler Itô integral equation.

1. INTRODUCTION

Inasmuch as computability can be of utmost importance, one often confines selection of stochastic differential equation models to those facilitating calculation and simulation. This is best exemplified in mathematical finance, where the popularity of the inaccurate Black-Scholes model is only justifiable through the evaluation ease of the resulting derivative product formulae. Indeed, Kunita (1984, p. 272) writes in his notes on stochastic differential equations that "It is an important problem in applications that we can compute the output from the input explicitly". Although pedagogical considerations initially prompted our classifications of which Itô processes $X_t^{x_0}$, starting at x_0, are representable as a time-dependent function of a simple stochastic integral $\phi^{x_0}\left(t, \int_0^t U^{x_0}(u)dW_u\right)$, our determination of ϕ^{x_0}, U^{x_0} executes an effective means of calculation and simulation. To simulate, one merely needs to compute the Gauss-Markov process $\int_0^t U^{x_0}(u)dW_u$ at discrete times and substitute these samples into ϕ^{x_0}, which is often known in closed form. Our work also makes properties of certain stochastic differential equations readily discernible and simplifies some filtering calculations. Finally, as demonstrated in Karatzas and Shreve (1987), page 295 ff., explicit solutions can be useful in establishing convergence for solutions of stochastic differential equations.

Doss (1977) and Sussman (1978) 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 Lie algebra generated by the

2000 *Mathematics Subject Classification.* Primary 60H20, Secondary 60H10, 60H35.

Key words and phrases. Diffeomorphism, Itô processes, explicit solutions.

Partial funding in support of this work was provided by the Natural Sciences and Engineering Research Council of Canada, and by the Prediction in Interacting Systems project within MITACS Network of Centre of Excellence.

vector fields of coefficients and showed, in this case, that strong solutions, $X_t^{x_0}$, of Fisk-Stratonovich equations are representable as $X_t^{x_0} = \rho(\Phi(x_0, W.)_t, W_t)$, for some continuous ρ, Φ solving differential equations. Under the restriction of C^∞ coefficients, Yamato (1979) extended Doss by dispensing with the Abelian assumption in favour of less restrictive q step nilpotency, whilst also introducing a simpler form for his explicit solutions $X_t^{x_0} = u(x, t, (W_t^I)_{I\in F})$. 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 (1984), 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 similar to ours, and, more generally, under solvability of the underlying Lie algebra. Kunita's work therefore generalizes Yamato (1979). Perhaps, the two most distinguishing features of our work are: We allow time-dependent coefficients and utilize a different representation. We compare our results to these works in Subsection 2.2.

In order to describe our method, we mention that the hitherto rather ad hoc state space diffeomorphism mapping method can be 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 corresponding to a more complicated Itô process via $\mathcal{L}f(x) = \{L(f\circ\psi)\}\circ\psi^{-1}(x)$. For related examples, we refer the reader to the problems on page 126 of Friedman (1975) or page 303 of Ethier and Kurtz (1986). This corresponds to using Ito's formula on $X_t = \psi(\xi_t)$ for some continuously differentiable, injective ψ, where ξ is a diffusion process with infinitesimal generator L. Motivated by applications in filtering, Kouritzin and Li (1999) and Kouritzin (2000) used differential equation methods to study: "When can global, time-dependent diffeomorphism be used to construct solutions to Ito 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

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t,\ X_0 = x_0, \tag{1.1}$$

which are of the form $\phi^{x_0}\left(t, \int_0^t U(u)dW_u\right)$, and showed that all nonsingular solutions of this form were actually (time-dependent) diffeomorphisms $\psi_t(\xi_t)$ with ξ satisfying

$$d\xi_t = (\chi - \kappa\xi_t)dt + dW_t,\ \xi_0 = \psi_0^{-1}(x_0).$$

Nonsingular 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 non-stochastic differential equations continued to hold in the singular situation when global diffeomorphisms fail.

Herein, we consummate work on resolving the question: "When can we explicitly solve Itô equations like (1.1) through a representation of the form $X_t^{x_0} = \phi^{x_0}\left(t, \int_0^t U^{x_0}(u)dW_u\right)$?", now concentrating on the rich vector-valued case, with the dimensions of X_t, W_t being p, d respectively. Moreover, motivated by the aforementioned applications, we consider solutions starting from every point in a convex domain, that is a non-empty convex open set. In order to include as many interesting examples as possible we will only require *local* representation

$X_t = \phi\left(t, \int_0^t U(u)dW_u\right)$ and allow σ to have rank less than $\min(p,d)$. The first opportunity borne out of allowing the rank of $\sigma(x)$ to be less than p is the ability to 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: The diffeomorphism solutions are not just transformed Gaussian processes but rather are constructed via $X_t = \psi_t(Z_t)$, with $Z_t = \begin{bmatrix} Z_t^{(1)} \\ Z_t^{(2)} \end{bmatrix}$, where $Z_t^{(2)} \in \mathbb{R}^{p-r}$ is deterministic, and $Z^{(1)}$ is a Gauss-Markov process satisfying

$$dZ_t^{(1)} = \{\alpha_t - \beta_t Z_t^{(1)}\}dt + \left(I_r \middle| \gamma_t\right) dW_t, \tag{1.2}$$

for some coefficients α_t, β_t, and γ_t, depending on $Z_t^{(2)}$. The exact forms of $Z_t^{(2)}$, α_t, β_t, and γ_t will be given after Theorem 2.

In the next section, we introduce notation and state the main results. Section 3 is devoted to applications, while the proofs of the theorems are postponed to Section 4.

2. Notation and main results

We assume throughout that $D \subset \mathbb{R}^p$ is a convex domain, $T > 0$, and define

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

$(x,s) \in D_T$ means $x \in D$ when $D_T = D$. To deal with derivatives on D_T, we define:

Definition 1. *Suppose $O \subset D_T$ is relatively open. For functions $g \in \mathcal{C}(O, \mathbb{R}^p)$,*

$$\frac{d}{dt} g(x,t) = \lim_{h \searrow 0} \frac{g(x,t+h) - g(x,t)}{h}$$

for all $(x,t) \in O$ such that the limit exists. We define $\mathcal{C}^1(O;\mathbb{R}^p)$ to be the functions $g \in \mathcal{C}(O;\mathbb{R}^p)$ such that $\{\frac{d}{dx_i} g(\cdot,\cdot)\}_{i=1}^p$, $\frac{d}{dt} g(\cdot,\cdot)$, exist and are in $\mathcal{C}(O;\mathbb{R}^p)$. Moreover, we define $\mathcal{C}^r(O;\mathbb{R}^p)$ recursively to be the $g \in \mathcal{C}(O;\mathbb{R}^p)$ such that $\frac{d}{dx_i} g(\cdot,\cdot)$ $\frac{d}{dt} g(\cdot,\cdot)$ exist and are in $\mathcal{C}^{r-1}(O;\mathbb{R}^p)$ for each $i = 1,...,p$. For such functions of both x and t, $\nabla_x g$ is the Jacobian matrix of vector function g that is $(\nabla g)_{i,j} = \partial_{x_j} g_i$ while ∇g will include the time derivative as the last column.

We suppose throughout that

A1: $\sigma \in \mathcal{C}^1(D_T;\mathbb{R}^{p\times d})$ and $b \in \mathcal{C}^1(D_T;\mathbb{R}^p)$.

Now, we let σ_j denote the jth column of the matrix σ, define

$$h = b - \frac{1}{2}\sum_{j=1}^d \{\nabla_x \sigma_j\}\sigma_j \text{ on } D_T \tag{2.1}$$

and assume

A2: $h \in \mathcal{C}^1(D_T;\mathbb{R}^p)$.

Next, we consider functions $\phi^{x_0,s}$ for each $x_0, s \in D_T$ such that

A3: $\phi^{x_0,s} \in \mathcal{C}^{1,2}((s,t_0)\times\mathcal{N}_{x_0,s};\mathbb{R}^p)$ satisfies $\lim_{t\searrow s}\phi(t,0) = x_0$, where $t_0(x_0,s) > s$, and $\mathcal{N}_{x_0,s} \subset \mathbb{R}^p$ is a neighbourhood of 0 that can depend on x_0, s, 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, and define

$$X_t^{x_0,s} = \phi(t, Y_t) = \phi^{x_0,s}(t, Y_t^{x_0,s}) \text{ on } [s, \tau^0_{x_0,s}). \tag{2.2}$$

Here, $Y_t = \displaystyle\int_s^t U(u)dW_u$, and $U = U^{x_0,s} \in \mathcal{C}^1([s,T];\mathbb{R}^{d\times d})$ is such that $U^{x_0,s}(s) = I_d$ for all $(x_0, s) \in D_T$, and

$$\tau^0_{x_0,s} = \min\left(T, \inf\{t > s : \ \phi^{x_0,s}(t, Y_t^{x_0,s}) \notin D \text{ or } \mathrm{Det}(U^{x_0,s}(t)) = 0\}\right).$$

To simplify notation, the dependence of ϕ and U on x_0, s will be often omitted.

Our first main result establishes necessary and sufficient conditions on σ, and h for existence of b, ϕ, and U so that $X^{x_0,s}$, defined in (2.2), is a strong solution to

$$dX_t = b(X_t, t)dt + \sigma(X_t, t)dW_t, \qquad X_s = x_0 \tag{2.3}$$

on $[s, \tau_{x_0,s})$, where $\tau_{x_0,s}$ is $\mathcal{F}_t$-stopping time, satisfying $s < \tau_{x_0,s} \leq \tau^0_{x_0,s}$. By continuity of ϕ, and Y, $P(\tau^0_{x_0,s} > s) = 1$ for all $x_0 \in D$. We note that b, σ are Lipschitz on any compact, convex subset of D_T by our $\mathcal{C}^1$-conditions and use the proof for Theorem II.5.2 of Kunita (1984) for existence and uniqueness of (strong) local solutions to the SDE in the following theorem until they leave such a compact subset. h is always related to b through (2.1).

Theorem 1. *Suppose [A1,A2,A3] hold. Then, a local solution $X_t^{x_0,s}$ to $dX_t = b(X_t,t)dt + \sigma(X_t,t)dW_t$, $X_s = x_0$ has an explicit form $\phi(t,Y_t)$ on $[s,\tau_{x_0,s})$, for some stopping time $s < \tau_{x_0,s} \leq \tau^0_{x_0,s}$ and each $(x_0,s) \in D_T$ if and only if*

$$(\nabla_x\sigma_k)\sigma_j = (\nabla_x\sigma_j)\sigma_k \text{ on } D_T, \text{ for all } j,k \in \{1,...,d\}, \tag{2.4}$$

and there exist $x \to A(x) \in \mathbb{R}^{d\times d}$, $\{\epsilon_s(x_0)\}_{s\geq 0} \subset (0,\infty)$ such that

$$(\sigma A)_j - \partial_t\sigma_j = \{\nabla_x\sigma_j\}h - \{\nabla_x h\}\sigma_j, \text{ for all } 1 \leq j \leq d, \tag{2.5}$$

on D_T and the following condition holds for all $t \in [s, s+\epsilon_s)$, y in a neighborhood of 0

$$\sigma(\phi(t,y),t)\{A(\phi(t,y),t) - U^{-1}(t)\dot{U}(t)\} = 0. \tag{2.6}$$

Then, for each fixed $(x_0,s) \in D_T$, there is a neighborhood $\mathcal{N}_{x_0,s}$ of $0 \in \mathbb{R}^d$ such that ϕ satisfies the following system of differential equations

$$\begin{aligned} \nabla_y\phi(t,y)U(t) &= \sigma(\phi(t,y),t), && (2.7)\\ \partial_t\phi(t,y) &= h(\phi(t,y),t), && (2.8)\\ \phi(s,0) &= x_0 && \end{aligned}$$

for all $t \in (s,\tau_{x_0,s})$ and $y \in \mathcal{N}_{x_0,s}$.

Remark 1. *The combination (2.5,2.6) constrain the possible h. For example, when $p = d$ and σ is nonsingular on D_T, A and U can not depend on x and (2.6) becomes equivalent to $A(s) = \dot{U}^s(s)$. Naturally, the general case is much richer and will require stronger conditions to study further. This is done in Theorem 2.*

When one explicit solution exists, there will be a whole class of such solutions corresponding to distinct b's. We now embark on identifying the b's, ϕ's and U's for these solutions. This necessitates introducing local diffeomorphisms.

Definition 2. *Suppose $\bar{x}=(x_0,0)\in D_T$. Then, a $\bar{x}$-local diffeomorphism $(O^{\bar{x}},\tilde{\Lambda})$ is a bijection $\tilde{\Lambda}:O^{\bar{x}}\to\tilde{\Lambda}(O^{\bar{x}})$ such that $\tilde{\Lambda}\in\mathcal{C}^1(O^{\bar{x}};\mathbb{R}^{(p+1)\times d})$, where $O^{\bar{x}}\subset D_T$ is a (relatively open) neighbourhood of $\bar{x}$. We define $\nabla\tilde{\Lambda}^{-1}(\tilde{\Lambda}(x,t))$ to be $\left[\nabla\tilde{\Lambda}(x,t)\right]^{-1}$ for $(x,t)\in O^{\bar{x}}$.*

Now, we can introduce our basic set of parameters for $\bar{x}=(x_0,0)$:

Definition 3. *Let $\mathcal{P}=\mathcal{P}^{\bar{x}}_{r,p,\sigma}$ be the set of all $(\tilde{\Lambda},\kappa,B,\theta)$ such that*

P1) *$\tilde{\Lambda}(x,t)=\begin{bmatrix}\Lambda(x,t)\\ t\end{bmatrix}$ if σ depends on time or $\tilde{\Lambda}(x)=\Lambda(x)$ otherwise is a $\bar{x}$-local diffeomorphism from $O\subset D_T$ onto $\mathcal{D}_T=\tilde{\Lambda}(O)$, where $\Lambda=\begin{bmatrix}\Lambda^{(1)}\\ \Lambda^{(2)}\end{bmatrix}$ with $\Lambda^{(1)}\in\mathbb{R}^r$;*

P2) *$\kappa\in\mathcal{C}^1(\mathcal{D}_T;\mathbb{R}^{r\times(d-r)})$ depends only on $x_{r+1},\dots,x_p$, and t;*

P3) *$\{(\nabla_x\Lambda)\sigma\}\circ(\Lambda)^{-1}=\begin{pmatrix}I_r & \kappa\\ 0 & 0\end{pmatrix}$ on $\mathcal{D}_T$;*

P4) *$B\in\mathcal{C}^1(\mathcal{D}_T;\mathbb{R}^{r\times r})$ depends only on $x_{r+1},\dots,x_p$, and t;*

P5) *$\theta=\begin{bmatrix}\theta_1\\ \theta_2\end{bmatrix}\in\mathcal{C}^1(\mathcal{D}_T;\mathbb{R}^r)\times\mathcal{C}^1(\mathcal{D}_T;\mathbb{R}^{p-r})$ depends only on $x_{r+1},\dots,x_p,t$.*

To each $(\Lambda,\kappa,B,\theta)\in\mathcal{P}_{r,p,\sigma,D_T}$, we extend θ,κ,B to $(\bar{\mathbb{R}})^p\times(-T,T)$, where $\bar{\mathbb{R}}$ is the extended real line, by defining them to be 0 off of $\mathcal{D}_T$ and associate the following functions:

$$
(2.9)\quad
\begin{cases}
c_2\in\mathbb{R}^{p-r}\text{ uniquely solves }\partial_t c_2=\theta_2(c_2,t),\ c_2(0)=\Lambda^{(2)}(x,0);\\
G(t)=\Big(I_r\Big|\ \kappa_t\circ c_2(t)\Big)\in\mathbb{R}^{r\times d};\\
R(t)=B_t\circ c_2(t)\in\mathbb{R}^{r\times r};\\
Q\in\mathbb{R}^{r\times r}\text{ is the unique solution of }\dot{Q}=QR,\ Q(0)=I_r;\\
c=\begin{bmatrix}c_1\\ c_2\end{bmatrix},\text{ with }c_1(t)=Q^{-1}(t)\Big\{\Lambda^{(1)}(x_0,0)+\int_0^t Q(u)\theta_1(c_2(u),u)du\Big\}.
\end{cases}
$$

These objects are well defined for $t<T$. With these definitions in hand, we characterize all solutions $X^{x_0}_t=\phi^{x_0}(t,Y_t)$ to (2.3) with $s=0$. Accordingly, we must strengthen our assumptions on σ. For the following theorem, we recall that b and h are still related through (2.1) and assume:

C1: $b,h\in\mathcal{C}^1(D_T;\mathbb{R}^p)$.

∂_r: $\sigma\in\mathcal{C}^{\bar{r}}(D_T;\mathbb{R}^{p\times d})$, where $\bar{r}=r+1$ and $r\in\{1,2,...\}$.

H_r: The rank of σ is r and its first r columns are linearly independent on D_T.

If σ has rank r yet H_r is not satisfied, we can simply permute the indices.

Theorem 2. *Suppose that [C1,H_r,∂_r] hold, and $X^{x_0,s}_t=\phi^{x_0,s}\left(t,\int_0^t U^{x_0,s}(u)dW_u\right)$ solves (2.3) up to some stopping time $\tau_{x_0,s}$ satisfying $0<\tau_{x_0,s}\le\tau^0_{x_0,s}$, for all $(x_0,s)\in D_T$. Then, for any $x_0\in D$ there exists $((O^{x_0},\tilde{\Lambda}),\kappa,B,\theta)\in\mathcal{P}^{x_0}_{r,p,\sigma}$, and related functions c,G,R,Q defined by (2.9), such that*

$$
(2.10)\qquad h_t=[\nabla_x\Lambda_t]^{-1}\times\left\{\theta_t\circ\Lambda^{(2)}_t-\partial_t\Lambda_t-\begin{bmatrix}(B_t\circ\Lambda^{(2)}_t)\Lambda^{(1)}_t\\ 0\end{bmatrix}\right\}\ \text{on } O^{x_0},
$$

$$\phi(t,y) = \phi_{(\Lambda,\kappa,B,\theta)}(t,y) = \Lambda_t^{-1}\left(c(t) + \begin{bmatrix} Q^{-1}(t)G(0)y \\ 0 \end{bmatrix}\right) \tag{2.11}$$

on $\mathcal{N}^{x_0} = \{(t,y) : c(t) + \begin{bmatrix} Q^{-1}(t)G(0)y \\ 0 \end{bmatrix} \in \Lambda_t(O^{x_0})\}$*, and* U *is any solution to* $G(0)U(t) = Q(t)G(t)$*. Moreover,*

$$X_t^{x_0,0} = \Lambda_t^{-1}\left(c(t) + \begin{bmatrix} Q^{-1}(t)\tilde{Y}_t \\ 0 \end{bmatrix}\right)$$

up to $\tau'_{x_0} = \min\left(\tau_{x_0,0}, \inf\{t > 0 : (t, Y_t) \notin \mathcal{N}^{x_0}\}\right)$*, where*

$$\tilde{Y}_t = G(0)Y_t = \int_0^t G(0)U(s)dW(s) = \int_0^t Q(s)G(s)dW(s).$$

Finally, if $\bar{\Lambda}$ *and* $\bar{\kappa}$ *satisfies P1–P3, then there exist* $\bar{B}, \bar{\theta}$ *such that* $(\bar{\Lambda}, \bar{\kappa}, \bar{B}, \bar{\theta}) \in \mathcal{P}$*,* $b_{(\bar{\Lambda},\bar{\kappa},\bar{B},\bar{\theta})} = b_{(\Lambda,\kappa,B,\theta)}$*, and* $\phi_{(\bar{\Lambda},\bar{\kappa},\bar{B},\bar{\theta})} = \phi_{(\Lambda,\kappa,B,\theta)}$*.*

The quantities $Z_t = \begin{bmatrix} Z_t^{(1)} \\ Z_t^{(2)} \end{bmatrix}$, α_t, β_t, and γ_t appearing in (1.2) are related to our parameters in the following way: $Z_t = \Lambda_t(X_t)$, implying that $Z_t^{(2)} = c_2(t)$, $\alpha_t = \theta_1(c_2(t), t)$, $\beta_t = B_t \circ c_2(t)$, and $\gamma_t = \kappa(c_2(t), t)$.

Remark 2. *There is no loss of generality in setting*

$$U^{x_0}(t) = \begin{pmatrix} Q(t) & Q(t)\kappa \circ c_2(t) - \kappa \circ c_2(0) \\ 0 & I_{d-r} \end{pmatrix}.$$

It yields $G(0)U(t) = Q(t)G(t)$ *and* $U(0) = I_d$*. Moreover, its inverse always exists and is given by* $\begin{pmatrix} Q^{-1}(t) & Q^{-1}(t)\kappa \circ c_2(0) - \kappa \circ c_2(t) \\ 0 & I_{d-r} \end{pmatrix}$*. When* $\bar{U}$ *also satisfies* $Q(t)G(t) = G(0)\bar{U}(t)$ *we have that*

$$G(0)Y_t = \int_0^t G(0)\bar{U}(s)dW(s) = \int_0^t Q(s)G(s)dW(s) = \int_0^t G(0)U(s)dW(s).$$

Suppose we set $A = \dot{U}(0)$ *and* $\begin{bmatrix} B & B' \\ 0 & 0 \end{bmatrix} = \begin{pmatrix} I_r & \kappa \\ 0 & 0 \end{pmatrix} A$ *(so* B *does not depend on* $\Lambda^{(1)}$*). Then,* ϕ *and* h *are uniquely determined by* Λ_t*,* θ*, and* B *through (2.11) and (2.10).*

Remark 3. *To illustrate the need of the final statement of Theorem 2, we take for example,* $\sigma(x) = x \in \mathbb{R}^p$*. Then, any* $L \in \mathcal{C}^1(\mathbb{R}^p)$ *depending on* $x_2/x_1, \ldots, x_p/x_1$ *satisfies* $(\nabla L)\sigma = 0$*. Therefore,* Λ *and hence the parameter set is not unique but we can create the same* b, ϕ *from any consistent* κ, Λ*.*

In the next two subsections, we compare our framework to Fisk-Stratonovich equations, and our results with those appearing in Yamato (1979) and Kunita (1984).

2.1. **Relation to Fisk-Stratonovich equations.** It follows from, for example, Kunita (1984) p. 239 that the unique local solutions of our Itô equation (1.1) and of the Fisk-Stratonovich equation

$$dX_t^{x_0} = h(X_t^{x_0}, t)dt + \sigma(X_t^{x_0}, t) \bullet dW_t \tag{2.12}$$

are equal if (2.1) holds and σ is twice continuously differentiable or satisfies the Fisk-Stratonovich acceptable condition in D. We refer the reader to Chapter 5 of Protter (1995) for conditions that should be placed on the coefficients of Stratonovich equations when they are not $\mathcal{C}^2$. Therefore, irrespective of whether σ satisfies such a condition or not we will always relate b and h through (2.1) in the sequel. To avoid making a $\mathcal{C}^2$ or like assumption on σ, we will work with the slightly more cumbersome Itô equations.

2.2. **Comparison with the works of Yamato and Kunita.** In Section III.3 of Kunita's (1984) treatise, he considers representations of time-homogeneous Fisk-Stratonovich equations

$$dX_t^{x_0} = h(X_t^{x_0})dt + \sigma(X_t^{x_0}) \bullet dW_t \tag{2.13}$$

in terms of the flows generated by the vector fields $\mathfrak{X}_0(y) = \sum_{i=1}^{p} h_i(y)\frac{\partial}{\partial y_i}$ and $\mathfrak{X}_k(y) = \sum_{i=1}^{p} \sigma_{ik}(y)\frac{\partial}{\partial y_i}$, $k = 1, ..., d$, 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 $\mathcal{C}^3_\alpha$, $\mathcal{C}^4_\alpha$ (the locally four times continuously differentiable functions whose fourth derivative is α-Hölder continuous), his work gives rise to the composition formula

$$\begin{aligned}(X_t^{x_0})_i &= Exp\,(t\mathfrak{X}_0) \circ Exp\left(W_t^1\mathfrak{X}_1\right) \circ \cdots \circ Exp\left(W_t^d\mathfrak{X}_d\right) \circ \chi_i(x), \\ &= \phi_i(t, W_t)\end{aligned} \tag{2.14}$$

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

$$\frac{d}{ds}(f \circ \varphi_s) = \mathfrak{X}_k f(\varphi_s),\ \varphi_0 = x \qquad \forall f \in \mathcal{C}^\infty. \tag{2.15}$$

In fact, to use (2.14), one must solve (2.15) for $k = 0, ..., d$ and $f = \chi_i$, $i = 1, ..., d$. Kunita also goes beyond commutability, even surpassing Yamato (1979) in generality by considering the situation where $L_0(\mathfrak{X}_0, ..., \mathfrak{X}_d)$ is only solvable, but the expression replacing (2.14) necessarily becomes more unwieldy.

Our characterization of ϕ provided by Theorem 2 provides an alternative to (2.14) that is much more amenable to direct calculation. Corollary 1 (to follow) supplies a converse to (2.14) in the sense that if $X_t^{x_0}$ were to have such a functional representation $\phi^{x_0}(t, W_t)$ in terms of Brownian motions only, then the vector fields must commute. This was previously established in Theorem 4.1 of Yamato (1979) under $\mathcal{C}^\infty$ conditions on both ϕ and the coefficients. Moreover, 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 favour of Itô processes.
- We remove the nilpotency assumptions (for our representations).

To validate the final claim, we take $p = 2$, $d = 1$, $\mathfrak{X}_0 = \{\theta_1(x_2) - B(x_2)x_1\}\partial_{x_1} + \theta_2(x_2)\partial_{x_2}$, and $\mathfrak{X}_1 = \partial_{x_1}$. Then $[\mathfrak{X}_0, \mathfrak{X}_1] = B\partial_{x_1}$. Moreover, if $\mathfrak{X}_k = [\mathfrak{X}_0, \mathfrak{X}_{k-1}]$, $k \geq 2$, then $\mathfrak{X}_k = a_k(x_2)\partial_{x_1}$, where $a_{k+1} = \theta_2(\partial_{x_2} a_k) + Ba_k$, $k \geq 1$, where $a_1 = 1$. 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 (1984), that is valid under the mild regularity on b, σ, h given at the beginning of the section.

Corollary 1. *Suppose that there is a domain $\widetilde{D}$ such that the coefficients σ and h are time-homogeneous and Fisk-Stratonovich acceptable on $\widetilde{D}_T = \widetilde{D} \times (0, T)$ and that the solution to the Fisk-Stratonovich equation (2.13) has a unique local solution*

$$(X_t^{x_0})_i = Exp\,(t\mathfrak{X}_0) \circ Exp\left(W_t^1 \mathfrak{X}_1\right) \circ \cdots \circ Exp\left(W_t^d \mathfrak{X}_d\right) \circ \chi_i(x)$$

on $0 \leq t < \tau_x$ 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 immediately following (2.13). Then,

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

Proof. We find that $X_t^{x_0} = \phi(t, Y_t)$ with $U(t) = I$ so it follows from Theorem 1 that $\sigma A = 0$. The condition $[\mathcal{X}_k, \mathcal{X}_j] = 0$ then follows from (2.4,2.5). $\square$

3. Examples of applications

3.1. **The square case.** Suppose that $\sigma = \sigma(x, t)$ is a $d \times d$ non singular continuously differentiable matrix satisfying (2.4). It follows from Theorem 2 that there exists a local diffeomorphism Λ_t such that $\nabla_x \Lambda_t = [\sigma(x, t)]^{-1}$, and all explicit solutions are of the form $\phi(t, y) = \Lambda_t^{-1}\left(c(t) + Q^{-1}(t)y\right)$, where $Q(t) = \int_0^t Q(s)B(s)ds + I$ and $c(t) = Q^{-1}(t)\left\{\Lambda_0(x) + \int_0^t Q(s)\theta(s)ds\right\}$ for some $\theta \in \mathcal{C}([0, T); \mathbb{R}^d)$, and some $B \in \mathcal{C}^1([0, T), \mathbb{R}^{d \times d})$. In this case, the corresponding diffusion drift b is given by

$$b_t(x) = \sigma(x, t) \times \{\theta(t) - B(t)\Lambda_t(x) - \partial_t \Lambda_t\} + \frac{1}{2}\sum_{j=1}^d (\nabla_x \sigma_j(x, t))\sigma_j(x, t).$$

In particular, if $d = 1$ and $x_0 \in \mathbb{R}$, then $\Lambda_t(x) = \int_{x_0}^x \frac{1}{\sigma(y,t)} dy$ is one solution.

If σ does not depend on t, then Λ need not either, both θ and B are constant, $Q(t) = e^{tB}$, b is given by

$$b(x) = \sigma(x) \times \{\theta - B\Lambda(x)\} + \frac{1}{2}\sum_{j=1}^d (\nabla \sigma_j(x))\sigma_j(x),$$

and $\phi(t, y) = \Lambda^{-1}\left(c(t) + Q^{-1}(t)y\right)$.

Example 1. *For example, take $\sigma_{ij}(x,t) = x_i\gamma_{ij}(t)$ and $D = (0,\infty)^d$. Then σ satisfies condition (2.4) since $[(\nabla_x\sigma_j)\sigma_k]_i = x_i\gamma_{ij}\gamma_{ik}$, and the diffeomorphism can be chosen as $\Lambda_t(x) = \gamma_t^{-1}\begin{bmatrix}\log x_1\\ \vdots \\ \log x_d\end{bmatrix}$. Note that the image is $\mathbb{R}^d$, so $\Lambda_t^{-1} = \begin{bmatrix} e^{(\gamma_t x)_1} \\ \vdots \\ e^{(\gamma_t x)_d}\end{bmatrix}$ is defined everywhere. In this case, for $1 \le i \le d$,*

$$b_i(x,t) = x_i\left\{\frac{1}{2}[\gamma_t\gamma_t^\top]_{ii} + [\gamma(t)\theta(t)]_i - \sum_{j=1}^d \left([\gamma_t B_t\gamma_t^{-1}]_{ij} - [\dot\gamma_t\gamma_t^{-1}]_{ij}\right)\log x_j\right\},$$

and $\phi_i(t,y) = \exp\left[\gamma_t\{c(t)+Q^{-1}(t)y\}\right]_i$, where $Q(t)$ is the (fundamental matrix) solution to $\dot Q(t) = Q(t)B(t)$, $B(0)=I$, and $c(t) = Q^{-1}(t)\left\{\Lambda_0(x) + \int_0^t Q(s)\theta(s)ds\right\}$. Note also that $b_i(x,t)$ can also be written as $x_i\left\{\alpha_i(t) - \sum_{j=1}^d \beta_{ij}(t)\log x_j\right\}$, where $\alpha_i(t) = \frac{1}{2}[\gamma_t\gamma_t^\top]_{ii} + [\gamma(t)\theta(t)]_i$, $1\le i\le d$, and $B_t = \gamma_t^{-1}\beta_t\gamma_t - \gamma_t^{-1}\dot\gamma_t$.

Example 2. *Another example is provided by the so-called linear case where σ depends only on t. In that case, $\Lambda_t = [\sigma(t)]^{-1}x$ satisfies the conditions, $\Lambda_t^{-1} = \sigma(t)x$, b is given by*

$$b(x,t) = \sigma(t)\times\left\{\theta(t) - B(t)[\sigma(t)]^{-1}x + [\sigma(t)]^{-1}\dot\sigma(t)[\sigma(t)]^{-1}\right\},$$

and $\phi(t,y) = \sigma(t)\times\left\{c(t)+Q^{-1}(t)y\right\}$, where $Q(t) = \exp\left\{\int_0^t B(s)ds\right\}$, and $c(t) = Q^{-1}(t)\left\{[\sigma(0)]^{-1}(x) + \int_0^t Q(s)\theta(s)ds\right\}$. Therefore, the stochastic differential equation $dX_t = (\alpha(t) - \beta(t)X_t)dt + \sigma(t)dW_t$ corresponds to $B(t) = [\sigma(t)]^{-1}\beta(t)\sigma(t)$ and $\theta(t) = [\sigma(t)]^{-1}\alpha(t) + \partial_t[\sigma(t)]^{-1}$.

Example 3. *Suppose*

$$\sigma(x_1,x_2) = \begin{pmatrix} \frac{x_1}{2}\sqrt{2\log x_1x_2 - (\log x_1x_2)^2} & -\frac{x_1}{2x_2}\sqrt{x_2(x_1-x_2)} \\ \\ \frac{x_2}{2}\sqrt{2\log x_1x_2 - (\log x_1x_2)^2} & -\frac{1}{2}\sqrt{x_2(x_1-x_2)}\end{pmatrix},$$

on $1 < x_1x_2 < e$, $x_2 \le x_1$. Then σ satisfies condition (2.4), and $\Lambda(x_1,x_2) = \begin{bmatrix}\frac{\pi}{2} + \arcsin(\log x_1x_2 - 1) \\ \frac{\pi}{2} + \arcsin(\frac{2x_2}{x_1} - 1)\end{bmatrix}$ satisfies $(\nabla\Lambda)\sigma = I_2$. It follows that

$$\Lambda^{-1}(x_1,x_2) = \begin{bmatrix} 2^{1/2}\exp\{(1-\cos x_1)/2\}\Big/(1-\cos x_2) \\ \exp\{(1-\cos x_1)/2\}((1-\cos x_2)/2)^{1/2}\end{bmatrix}, \quad (x_1,x_2)\in(0,\pi)^2,$$

and

$$(\nabla\sigma_1)\sigma_1 + (\nabla\sigma_2)\sigma_2 = \frac{1}{4}\begin{bmatrix}\frac{x_1(2x_1+x_2-x_2(\log x_1x_2)^2)}{x_2} \\ x_2(1-(\log x_1x_2)^2)\end{bmatrix}.$$

3.2. **Heisenberg group.** Let $A = A(t)$ be a continuously differentiable $\mathbb{R}^{d\times d}$ matrix and set $\sigma(x,t) = \sigma(y,z,t) = \begin{bmatrix} I_d \\ (Ay)^\top \end{bmatrix}$, where $y \in \mathbb{R}^d$, z is real, and $t \geq 0$. Then σ has rank d and $\nabla_x \sigma_j = \begin{pmatrix} 0 & 0 \\ (A^\top)_j & 0 \end{pmatrix}$. Hence

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

Therefore condition (2.4) holds true if and only if A is symmetric. In that case, using Theorem 2, one diffeomorphism $\tilde{\Lambda}(y,z,t) = \begin{bmatrix} \Lambda(y,z,t) \\ t \end{bmatrix}$ is found by setting $\Lambda(y,z,t) = \begin{bmatrix} y \\ g \end{bmatrix}$, where $g(y,z,t) = z - \frac{1}{2} y^\top A(t) y$.

Since $(\nabla \sigma_j)\sigma_j = \begin{bmatrix} 0 \\ A_{jj} \end{bmatrix}$, it follows that $\frac{1}{2}\sum_{j=1}^d (\nabla\sigma_j)\sigma_j = \frac{1}{2}\begin{bmatrix} 0 \\ \mathrm{Tr}(A) \end{bmatrix}$. Next, $\partial_t \Lambda = -\frac{1}{2}\begin{bmatrix} 0 \\ y^\top \dot{A} y \end{bmatrix}$, $[\nabla_x \Lambda]^{-1} = \begin{pmatrix} I_d & 0 \\ (Ay)^\top & 1 \end{pmatrix}$. Using Theorem 2, b must be of the form

$$b(y,z,t) = \begin{bmatrix} \theta_1(g,t) - B(g,t)y \\ \theta_2(g,t) + \theta_1(g,t)^\top A(t) y - \frac{1}{2} y^\top \dot{A}(t) y - y^\top A(t)^\top B_t \circ g y + \frac{1}{2}\mathrm{Tr}\{A(t)\} \end{bmatrix},$$

where $B \in \mathbb{R}^{d\times d}$, $\theta_1 \in \mathbb{R}^d$ and $\theta_2 \in \mathbb{R}$ all depend on (z,t) and are continuously differentiable. Moreover, the corresponding ϕ is given by

$$\phi(t,y) = \begin{bmatrix} c_1(t) + Q^{-1}(t) y \\ c_2(t) + \frac{1}{2}\{c_1(t) + Q^{-1}(t)y\}^\top A(t) \{c_1(t) + Q^{-1}(t) y\} \end{bmatrix},$$

where $\dot{c}_2 = \theta_2(c_2, t)$, $c_2(0) = g(y,z,0) = z - \frac{1}{2} y^\top A(0) y$, $R(t) = B_t \circ c_2(t)$, Q solves $\dot{Q} = Q(t)R(t)$, $Q(0) = I_d$, and

$$c_1(t) = Q^{-1}(t)\left\{ y + \int_0^t Q(s)\theta_1(c_2(s), s) ds \right\}.$$

Finally, $\tilde{Y}_t = Y_t = \int_0^t Q(s) dW_s$.

The interesting case of skew-symmetric matrices A (e.g. Rémillard, 1994) is not covered as expected since the solution to $X_t = \int_0^t \sigma(X_s) dW_s = \begin{bmatrix} W_t \\ \int_0^t \{A(s)W_s\}^\top dW_s \end{bmatrix}$ has one component that is an iterated stochastic integral. In the case $A_t = A$ does not depend on time, the process X_t is known as the Brownian motion on the Heisenberg group, where the group operation is defined on $\mathbb{R}^{d+2} \times \mathbb{R}^{d+2}$ by

$$(y,z,t) \circ (y',z',t') = (y + y', z + z' + \frac{1}{2} < Ay, y' >, t + t').$$

Note that the group is commutative if and only if A is symmetric.

4. Proofs of the main results

4.1. **Proof of Theorem 1 and Corollary 1.**

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

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

Since continuous finite variation martingales are necessarily constant, the (continuous) coefficients of the Itô process $\phi(t, Y_t)$ match those of (2.3) on $(s, \tau_{x_0,s})$ if and only if there is a neighbourhood $\mathcal{N}_{x_0,s}$ of 0 such that

$$\sigma_{ij}(\phi, t) = \sum_{m=1}^d \partial_{y_m}\phi_i\; U_{mj}(t),\ 1 \le i \le p,\ 1 \le j \le d,$$

or in matrix form

$$\sigma(\phi, t) = \{\nabla_y \phi\} U(t),$$

for all $t \in (s, \tau_{x_0,s})$, $y \in \mathcal{N}_{x_0,s}$ proving (2.7), and

$$b_i(\phi,t) = \partial_t\phi_i + \frac{1}{2}\sum_{j=1}^d\sum_{k=1}^d \partial_{y_j}\partial_{y_k}\phi_i\; (U(t)U^\top(t))_{jk},\ 1 \le i \le p. \tag{4.1}$$

Now, using (2.7), one obtains

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

Multiplying the last equality by U_{mk}, summing over m and using (2.7) again, one finds that

$$\sum_{l=1}^p \{\partial_{x_l}\sigma_{ij}\}(\phi,t)\sigma_{lk}(\phi,t) = \sum_{m=1}^d\sum_{l=1}^d \partial_{y_m}\partial_{y_l}\phi_i\; U_{lj}(t)U_{mk}(t), \tag{4.2}$$

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

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

Hence under (2.7), (4.1) is equivalent to

$$\partial_t\phi = b(\phi,t) - \frac{1}{2}\sum_{k=1}^d \{\nabla_x\sigma_k\}(\phi,t)\; \sigma_k(\phi,t) = h(\phi,t), \tag{4.3}$$

using (2.1). This proves (2.8).

Next, it follows by [A2] that there is a unique continuous solution on $[s, t_0]$ for some $t_0 = t_0(x_0, s) > s$ to $\partial_t\phi(t, 0) = h(\phi(t, 0), t)$ such that $\phi(s, 0) = x_0$. Therefore, existence of our function $\phi^{x_0,s}$ follows from exactness of differential 1-forms, (2.7), and (4.3) if we can show that under (2.4), $\partial_{y_j}\{\sigma(\phi,t)(U^{-1}(t))_k\} =$

$\partial_{y_k}\{\sigma(\phi,t)(U^{-1}(t))_j\}$ and $\frac{d}{dt}\{\sigma(\phi,t)(U^{-1}(t))_k\} = \partial_{y_k}h(\phi,t)$. However, it follows by (2.7) and (2.4) that

$$\begin{aligned} (4.4)\quad \partial_{y_j}\{\sigma(\phi,t)(U^{-1}(t))_k\} &= \sum_m \nabla_\phi\sigma_m(\phi,t)\sigma(\phi,t)(U^{-1}(t))_j(U^{-1}(t))_{mk} \\ &= \sum_n \nabla_\phi\sigma_n(\phi,t)\sigma(\phi,t)(U^{-1}(t))_k(U^{-1}(t))_{nj} \\ &= \partial_{y_k}\{\sigma(\phi,t)(U^{-1}(t))_j\}. \end{aligned}$$

Conversely, since the righthand side of (4.2) is symmetric is j and k, it follows that for all $1 \le j,k \le d$,

$$\{\nabla_\phi\sigma_j\}(\phi,t)\sigma_k(\phi,t) - \{\nabla_\phi\sigma_k\}(\phi,t)\sigma_j(\phi,t) = 0.$$

Since $\phi^{x_0}(s,0) = x_0$, (2.4) must hold when our representation does. Next, turning to the necessity of (2.5,2.6), one gets by (2.7) and (2.8) that

$$\frac{d}{dt}\{\sigma_j(\phi,t)\} = \{\nabla_\phi\sigma_j\}(\phi,t)\partial_t\phi + \partial_t\sigma_j(\phi,t) = \{\nabla_\phi\sigma_j\}(\phi,t)h(\phi,t) + \partial_t\sigma_j(\phi,t),$$

and

$$\begin{aligned} \frac{d}{dt}\{\sigma_j(\phi,t)\} &= \{\nabla_y(\partial_t\phi)\}U_j(t) + \{\nabla_y\phi\}\dot U_j(t) \\ &= \{\nabla_\phi h\}(\phi,t)\nabla_y\phi U_j(t) + \{\nabla_y\phi\}\dot U_j(t) \\ &= \{\nabla_\phi h\}(\phi,t)\sigma_j(\phi,t) + \{\nabla_y\phi\}\dot U_j(t). \end{aligned}$$

Hence

$$\{\nabla_\phi\sigma_j\}(\phi,t)h(\phi,t) - \{\nabla_\phi h\}(\phi,t)\sigma_j(\phi) = \{\nabla_y\phi\}\dot U_j(t) - \partial_t\sigma_j(\phi,t).$$

Putting $(t,y) = (s,0)$, using the identity $\sigma(x,s) = lim_{t\searrow s}\nabla_y\phi^{s,x}(t,0)$ (from $U(s) = I_d$ and (2.7)), one obtains (2.5), that is

$$\{\nabla_x\sigma_j\}h - \{\nabla_x h\}\sigma_j = (\sigma A)_j - \partial_t\sigma_j \text{ on } D_T, \text{ for all } 1 \le j \le d,$$

where $A(s,x) = \dot U^{s,x}(s)$. Finally, using the last two identities as well as (2.7), and recalling the fact that U has an inverse for $t < \tau^0_{x_0,s}$, one gets (2.6), that is

$$\sigma(\phi,t)\{A(\phi,t) - U^{-1}(t)\dot U(t)\} = 0.$$

Conversely, using (2.5) and (2.6), we get that

$$\begin{aligned} \frac{d}{dt}\{\sigma(\phi,t)(U^{-1}(t))_k\} &= \nabla_\phi\{\sigma(\phi,t)(U^{-1}(t))_k\}\partial_t\phi + \{\partial_t\sigma(\phi,t)\}(U^{-1}(t))_k \\ &\quad +\sigma(\phi,t)(\dot{U^{-1}}(t))_k \\ &= \sum_n\{\nabla_\phi\sigma_n(\phi,t)\}h(\phi,t)(U^{-1}(t)_{nk} + \{\partial_t\sigma(\phi,t)\}(U^{-1}(t))_k \\ &\quad -\sigma(\phi,t)(A(\phi,t)U^{-1}(t))_k \\ &= \{\nabla_\phi h(\phi,t)\}\sigma(\phi,t)(U^{-1}(t))_k \\ &= \partial_{y_k}h(\phi,t) \end{aligned}$$

so by (4.4) and the previous equations (2.4), (2.5), and (2.6) are also sufficient. □

Proof of Corollary 1. The corollary follows directly from Theorem 1, (2.4) and (2.5) by noting that having a function of Brownian motion W corresponds to taking $U(t) = I_d$ in our representations. □

4.2. **Proof of Theorem 2 in the time-independent case.** To prove Theorem 2, we change coordinates using a local time-dependent diffeomorphism. In this new coordinate system, the problem of representing $Z_t = \Lambda_t(X_t)$ as $\tilde{\phi}(t, Y_t) = \Lambda_t \circ \phi(t, Y_t)$ is much simpler. The existence of such a diffeomorphism follows from results in differential geometry given in the appendix. We will first prove all the results for diffusion coefficients b and σ not depending on time.

Proof of Theorem 2. We show that if, $X_t^{x_0} = \phi^{x_0}(t, Y_t)$ is a solution of the stochastic differential equation $dX_t = b(X_t)dt + \sigma(X_t)dW_t$, $X_0 = x_0$ up to τ_{x_0}, then, there exists $(\tilde{\Lambda}, \kappa, B, \theta) \in \mathcal{P} = \mathcal{P}_{r,p,\sigma,D_T}$ such that $b = b_{(\Lambda,\kappa,B,\theta)}$, and $\phi = \phi_{(\Lambda,\kappa,B,\theta)}$.

By Theorem 1, (2.4), and Proposition 2 in the appendix there exist x_0-local diffeomorphism (O, Λ) and κ satisfying P1), P2) and P3). In particular,

$$\tilde{\sigma} = \{(\nabla_x \Lambda)\sigma\} \circ \Lambda^{-1} = \begin{pmatrix} I_r & \kappa \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{p\times d} \text{ on } \mathcal{D} = \Lambda(O),$$

where $\kappa \in \mathbb{R}^{r\times(d-r)}$ does not depend on $x_1, \ldots, x_r$. If h is the the function defined by (2.1), set $\tilde{h} = \{(\nabla_x \Lambda)h\} \circ \Lambda^{-1}$. Then, it follows from Lemma 1 in the appendix that $\tilde{\sigma}$ satisfies (2.4), rank$(\tilde{\sigma})$ = r and $(\tilde{\sigma}, \tilde{h})$ satisfies (2.5). One finds that $Z_t = \Lambda(X_t)$ is the unique local solution to $dZ_t = \tilde{b}(Z_t)dt + \tilde{\sigma}(Z_t)dW_t$, $Z_0 = \Lambda(x_0)$, where

$$\tilde{b} = \tilde{h} + \frac{1}{2}\sum_{j=1}^{d}(\nabla_z \tilde{\sigma}_j)\tilde{\sigma}_j.$$

Moreover X_t has representation $\phi(t, Y_t)$ if and only if Z_t has representation $\tilde{\phi}(t, Y_t) = \Lambda \circ \phi(t, Y_t)$. Therefore, one only has to prove that there exists B and θ satisfying P4) and P5), such that $\tilde{h}(u, v) = \theta(v) - \begin{bmatrix} B(v)u \\ 0 \end{bmatrix}$, and $\phi(t, y) = c(t) + \begin{bmatrix} Q^{-1}(t)G(0)y \\ 0 \end{bmatrix}$, where $z = \begin{bmatrix} u \\ v \end{bmatrix} \in \mathcal{D}$, $u \in \mathbb{R}^r$, and c, Q, R, and G satisfy (2.9). Thus, until stated otherwise, we set $\sigma = \begin{pmatrix} I_r & \kappa \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{p\times d}$, omitting the tilde.

It follows from Theorem 1 that ϕ satisfies (2.7), that is $\nabla_y \phi = \sigma(\phi)U^{-1}$. Then, ϕ must be of the form

$$\phi(t, y) = c(t) + \begin{pmatrix} I_r & \kappa \circ c_2(t) \\ 0 & 0 \end{pmatrix} U^{-1}(t)y = \begin{bmatrix} c_1(t) + G(t)U^{-1}(t)y \\ c_2(t) \end{bmatrix},$$

for some $c = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \in \mathcal{C}^1(\mathcal{D}; \mathbb{R}^p)$, $c_2 \in \mathbb{R}^{p-r}$, $c(0) = \Lambda(x_0)$, and $G(t) = \left(I_r \middle| \kappa \circ c_2(t)\right)$. Also, $\sigma \circ \phi(t, y) = \begin{pmatrix} I_r & \kappa \circ c_2(t) \\ 0 & 0 \end{pmatrix} = \begin{bmatrix} G(t) \\ 0 \end{bmatrix}$. Set $\tilde{B} = \left(B \middle| \bar{B}\right) = \left(I_r \middle| \kappa\right) A$, where $B \in \mathbb{R}^{r\times r}$. It follows from (2.6) that $\tilde{B}(\phi) = GU^{-1}\dot{U}$. Since the right hand-side does not depend on y, one obtains

$$0 = \nabla_y\{\tilde{B}_j \circ \phi\}U = \{\nabla \tilde{B}_j\}(\phi)(\nabla_y \phi)U = \{\nabla \tilde{B}_j\}(\phi)\sigma(\phi) = (\nabla \tilde{B}_j)(\phi) \begin{bmatrix} G \\ 0 \end{bmatrix},$$

for any $1 \le j \le d$. Then, putting $(t, y) = (0, 0)$ in the last equation, one obtains

$$0 = \left(\nabla_u \tilde{B}_j \middle| \nabla_v \tilde{B}_j\right) \begin{pmatrix} I_r & \kappa \\ 0 & 0 \end{pmatrix},$$

which entails that $\tilde{B}$ does not depend on u. Setting $\tilde{R} = \tilde{B} \circ c_2$, and $R = B \circ c_2$, we write (2.6) as

$$\tilde{R}V + G\dot{V} = 0, \tag{4.5}$$

where $V = U^{-1}$. Recall from Theorem 1 that h must solve

$$(\nabla\sigma_j)h - (\nabla h)\sigma_j = \sigma A_j = \begin{bmatrix} \tilde{B}_j \\ 0 \end{bmatrix}, \; 1 \le j \le d.$$

In particular, from $1 \le j \le r$, one finds that

$$h(z) = h(u, v) = \theta(v) - \begin{bmatrix} B(v)u \\ 0 \end{bmatrix},$$

for some $\theta \in \mathcal{C}^1(\mathcal{D}; \mathbb{R}^p)$ depending on v only. Next, taking into account the indices $j+r$, $1 \le j \le d-r$, one also has $(\nabla h)\sigma_{j+r} = -\begin{bmatrix} B\kappa_j \\ 0 \end{bmatrix}$, where κ_j denotes the j-th column of κ. Therefore

$$\begin{bmatrix} \tilde{B}_{j+r} \\ 0 \end{bmatrix} = \begin{bmatrix} \bar{B}_j \\ 0 \end{bmatrix} = (\nabla\sigma_{j+r})h - (\nabla h)\sigma_{j+r} = \begin{bmatrix} (\nabla_v \kappa_j)\theta_2 \\ 0 \end{bmatrix} + \begin{bmatrix} B\kappa_j \\ 0 \end{bmatrix}.$$

Thus

$$\bar{B}_j = (\nabla_v \kappa_j)\theta_2 + B\kappa_j, \text{ for all } 1 \le j \le d-r. \tag{4.6}$$

Since $h \circ \phi(t, y) = \theta \circ c_2(t) - \begin{bmatrix} R(t)\{c_1(t) + G(t)V(t)y\} \\ 0 \end{bmatrix}$, the condition $\partial_t \phi = h(\phi)$ of Theorem 1 yields $\dot{c}_2 = \theta_2(c_2)$ and

$$\dot{c}_1 + \dot{G}Vy + G\dot{V}y = \theta_1(c_2(t)) - R(t)\{c_1(t) + G(t)V(t)y\}.$$

Therefore, c_2 is the unique solution of $\dot{c}_2 = \theta_2(c_2)$, $c_2(0) = \Lambda^{(2)}(x_0)$. One can now rewrite (4.6) as $\tilde{R} = \dot{G} + RG$. This equation, together with (4.5) implies $\partial_t(GV) + R(GV) = 0$, which can be written as $\partial_t(QGV) = 0$, where Q is the unique solution of $\dot{Q} = QR$, $Q(0) = I_r$. Therefore, $Q(t)G(t)U^{-1}(t) = G(0)$, $G(t)U^{-1}(t) = Q^{-1}(t)G(0)$, and $Q(t)G(t) = G(0)U(t)$. Hence, one can conclude that

$$\phi(t, y) = c(t) + \begin{bmatrix} Q^{-1}(t)G(0)y \\ 0 \end{bmatrix}.$$

Since $\partial_t(GV) + R(GV) = 0$, one obtains $\dot{c}_1 = \theta_1(c_2(t)) - R(t)c_1(t)$. Therefore

$$c_1(t) = Q^{-1}(t)\left\{\Lambda^{(1)}(x_0) + \int_0^t Q(s)\theta_1 \circ c_2(s)ds\right\}.$$

Finally, we show that any two pairs (Λ, κ) and $(\bar{\Lambda}, \bar{\kappa})$ satisfying P1), P2) and P3) generate the same class of solutions. More precisely, if $(\tilde{\Lambda}, \kappa, B, \theta) \in \mathcal{P}$, there exists $((\bar{\Lambda}\ t)^T, \bar{\kappa}, \bar{B}, \bar{\theta}) \in \mathcal{P}$ such that $h = h_{(\Lambda,\kappa,B,\theta)} = \bar{h}_{(\bar{\Lambda},\bar{\kappa},\bar{B},\bar{\theta})} = \bar{h}$, on $O^{x_0} \cap \bar{O}^{x_0}$ and $\phi = \phi_{(\Lambda,\kappa,B,\theta)} = \bar{\phi}_{(\bar{\Lambda},\bar{\kappa},\bar{B},\bar{\theta})} = \bar{\phi}$ on a neighbourhood of $(0,0)$.

$\hat{\Lambda} = \bar{\Lambda} \circ \Lambda^{-1}$ is a local diffeomorphism on $\mathcal{D}$ such that $\nabla\hat{\Lambda}\begin{pmatrix} I_r & \kappa \\ 0 & 0 \end{pmatrix} \circ \hat{\Lambda}^{-1} = \begin{pmatrix} I_r & \bar{\kappa} \\ 0 & 0 \end{pmatrix}$ so $\bar{\Lambda} \circ \Lambda^{-1}(u, v) = \begin{bmatrix} u + \psi_1(v) \\ \psi_2(v) \end{bmatrix}$, where ψ_2 is a diffeomorphism on a subset of $\mathbb{R}^{d-r}$ and $\kappa = \bar{\kappa} \circ \psi_2$. Therefore $\bar{\Lambda}^{-1}(\bar{u}, \bar{v}) = \Lambda^{-1}\left(\begin{bmatrix} \bar{u} - \psi_1 \circ \psi_2^{-1}(\bar{v}) \\ \psi_2^{-1}(v) \end{bmatrix}\right)$.

Moreover $\bar\Lambda = \begin{bmatrix} \Lambda^{(1)} + \psi_1 \circ \Lambda^{(2)} \\ \psi_2 \circ \Lambda^{(2)} \end{bmatrix}$. Suppose that θ and B are fixed and let c, G, R, and Q be the associated functions, as defined by (2.9). Set $\bar\theta_2 = \{(\nabla_v \psi_2)\theta_2\} \circ \psi_2^{-1}$ and $\bar B = B \circ \psi_2^{-1}$. Then $\bar c_2 = \psi_2 \circ c_2$ solves $\dot{\bar c}_2 = \bar\theta_2(\bar c_2)$, $\bar c_2(0) = \bar\Lambda^{(2)}(x_0)$. Moreover $\bar G = \left(I_r \middle| \ \bar\kappa \circ \bar c_2 \right) = \left(I_r \middle| \ \kappa \circ c_2 \right) = G(t)$, $\bar R = \bar B \circ \bar c_2 = B \circ c_2 = R$ and $\bar Q = Q$, where $\bar Q$ solves $\dot{\bar Q} = \bar Q \bar R$, $\bar Q(0) = I_r$. Finally, we set $\bar\theta_1 = \{(\nabla_v \psi_1)\theta_2 + B\psi_1 + \theta_1\} \circ \psi_2^{-1}$. Then

$$
\begin{aligned}
\bar c_1(t) &= (\bar Q)^{-1}(t) \left\{ (\bar\Lambda)^{(1)}(x_0) + \int_0^t \bar Q(s) \bar\theta_1 \circ \bar c_2(s) ds \right\} \\
&= Q^{-1}(t) \left\{ \Lambda^{(1)}(x_0) + \psi_1 \circ \Lambda^{(2)}(x_0) + \int_0^t Q(s)\theta_1 \circ c_2(s) ds \right\} \\
&\quad + Q^{-1}(t) \int_0^t Q(s)\{R(s)\psi_1 \circ c_2(s) + (\nabla_v \psi_1) \circ c_2(s)\theta_2 \circ c_2(s)\} ds \\
&= c_1(t) + \psi_1 \circ c_2(t).
\end{aligned}
$$

Hence

$$
\begin{aligned}
\bar\phi(t,y) &= (\bar\Lambda)^{-1} \left(\bar c(t) + \begin{bmatrix} (\bar Q)^{-1}(t) \bar G(0) y \\ 0 \end{bmatrix} \right) \\
&= (\bar\Lambda)^{-1} \left(\begin{bmatrix} c_1(t) + \psi_1 \circ c_2(t) + Q^{-1}(t) G(0) y \\ \psi_2 \circ c_2(t) \end{bmatrix} \right) \\
&= (\Lambda)^{-1} \left(\begin{bmatrix} c_1(t) + Q^{-1}(t) G(0) y \\ c_2(t) \end{bmatrix} \right) \\
&= \phi(t,y).
\end{aligned}
$$

It is also easy to check that

$$
\bar h(u,v) = (\nabla \bar\Lambda)^{-1} \times \left\{ \bar\theta - \begin{bmatrix} \bar B(v) u \\ 0 \end{bmatrix} \right\} = (\nabla \Lambda)^{-1} \times \left\{ \theta - \begin{bmatrix} B(v) u \\ 0 \end{bmatrix} \right\} = h(u,v).
$$

This completes the proof. □

4.3. **Proof of Theorem 2 in the time-dependent case.** By considering time as supplemental variable x_{p+1}, one can prove Theorem 2 when σ and b depend on time as a by-product of the time-independent case.

Proof. We first note that by hypotheses there is a unique strong solution to

$$
d\tilde X_t = \tilde b(\tilde X_t) dt + \tilde\sigma(\tilde X_t) dW_t, \ \tilde X_s = \begin{bmatrix} x_0 \\ s \end{bmatrix}
$$

up to $\tau_{x_0,s} > 0$ for all $\begin{bmatrix} x \\ s \end{bmatrix} \in D_T$, where time is treated as x_{p+1}, $\tilde X_t \in \mathbb{R}^{p+1}$, $\tilde\sigma = \begin{bmatrix} \sigma \\ 0 \end{bmatrix}$ for $t \ge 0$, and $\tilde h, \tilde b = \begin{bmatrix} h \\ 1 \end{bmatrix}, \begin{bmatrix} b \\ 1 \end{bmatrix}$ for $t \ge 0$. Moreover, $\tilde\sigma, \tilde b, \tilde h$ satisfy C1, ∂_r, H_r. Therefore, using the Theorem 2 in the time-independent setting as well as Proposition 2 in the appendix, one can conclude that there exists a $(\tilde\Lambda, \kappa, B, \tilde\theta) \in \mathcal{P} = \mathcal{P}_{r,p+1,\tilde\sigma}$ such that $\tilde\Lambda = \begin{bmatrix} \Lambda_t \\ t \end{bmatrix}$ satisfies $\{(\nabla \tilde\Lambda)\tilde\sigma\} \circ (\tilde\Lambda)^{-1} = \begin{pmatrix} I_r & \kappa \\ 0 & 0 \end{pmatrix}$ on some relatively open neighbourhood O of $(x_0, 0)$, and $\tilde h = \tilde h_{(\Lambda,\kappa,B,\tilde\theta)}$, and $\tilde\phi = \tilde\phi_{(\Lambda,\kappa,B,\tilde\theta)}$, where $\tilde{c} = \begin{bmatrix} c \\ c_3 \end{bmatrix}$, $\tilde{G}$, $\tilde{R}$, $\tilde{Q}$ are defined according to (2.9). It remains to prove the general expression for h, ϕ and to show $\tilde{\theta} = \begin{bmatrix} \theta \\ 1 \end{bmatrix}$.

Since $\nabla\tilde{\Lambda} = \begin{pmatrix} \nabla_x\Lambda & \partial_t\Lambda \\ 0 & 1 \end{pmatrix}$, it follows that $[\nabla\tilde{\Lambda}]^{-1} = \begin{pmatrix} [\nabla_x\Lambda]^{-1} & -[\nabla_x\Lambda]^{-1}\partial_t\Lambda \\ 0 & 1 \end{pmatrix}$. Thus $\tilde{h}$ can be written in the form

$$
\begin{aligned}
\tilde{h}(x,t) &= \begin{bmatrix} h_t \\ h_t' \end{bmatrix} \\
&= [\nabla\tilde{\Lambda}]^{-1} \times \left\{ \tilde{\theta}_t \circ \Lambda^{(2)} - \begin{bmatrix} \begin{bmatrix} (B_t \circ \Lambda^{(2)})\Lambda^{(1)} \\ 0 \end{bmatrix} \\ 0 \end{bmatrix} \right\} \\
&= \begin{bmatrix} [\nabla_x\Lambda]^{-1} \times \left[\{\theta - (\partial_t\Lambda)\theta_3\} \circ (\Lambda^{(2)}, t) \right] \\ \theta_3 \circ (\Lambda^{(2)}, t) \end{bmatrix} \\
&\qquad - [\nabla_x\Lambda]^{-1} \times \begin{bmatrix} \begin{bmatrix} (B_t \circ \Lambda^{(2)})\Lambda^{(1)} \\ 0 \end{bmatrix} \\ 0 \end{bmatrix},
\end{aligned}
$$

where $\tilde{\theta} = \begin{bmatrix} \theta \\ \theta_3 \end{bmatrix} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} \in \mathbb{R}^{p+1}$, $\theta_3 \in \mathbb{R}$. Hence, solutions of the form $\tilde{h} = \begin{bmatrix} h \\ 1 \end{bmatrix}$ are only possible when $\theta_3 = 1$. In this case, one has $c_3(t) = t$, so $\dot{c}_2 = \theta_2(c_2, t)$, $c_2(0) = \Lambda^{(2)}\left(\begin{bmatrix} x_0 \\ 0 \end{bmatrix} \right)$ and, consequently h and ϕ clearly have representation

$$h(x,t) = [\nabla_x\Lambda]^{-1} \times \left\{ \theta_t \circ \Lambda^{(2)} - \partial_t\Lambda - \begin{bmatrix} (B_t \circ \Lambda^{(2)})\Lambda^{(1)} \\ 0 \end{bmatrix} \right\},$$

and

$$\phi(t,y) = \phi_{(\Lambda,\kappa,B,\theta)}(t,y) = \Lambda_t^{-1} \left(c(t) + \begin{bmatrix} Q^{-1}(t)G(0)y \\ 0 \end{bmatrix} \right),$$

as stated in Theorem 2. □

5. Appendix: Local Diffeomorphisms

We fix $r \in \{1, 2, ...\}$, set $\bar{r} = r + 1$, take $q = p + 1$ if σ or b depend on t or $q = p$ otherwise and assume in this appendix that $\sigma \in \mathcal{C}^{\bar{r}}(D_T; \mathbb{R}^{q\times d})$ satisfies H_r and $\sigma_{p+1} = 0$ if $q > p$. Next, we let $D_T^2 = \begin{cases} D \times (-T, T) & \text{if } \sigma \text{ or b depend on } t \\ D & \text{otherwise} \end{cases}$, fix $\bar{x} \in D_T$, set $\partial_t\sigma(x,t) = \partial_t\sigma(x,0)$, $\partial_{x_i}\sigma(x,t) = \partial_{x_i}\sigma(x,0)$ for $t < 0$, $i = 1, 2, ..., q$ and use exactness of the corresponding 1-form to extend σ uniquely to D_T^2 such that $\sigma \in \mathcal{C}^{\bar{r}}(D_T^2; \mathbb{R}^{q\times d})$. By making $T > 0$ smaller if necessary, we can assume that the first r columns of σ are linearly independent on D_T^2.

The following lemma can be proven by elementary calculations.

Lemma 1. *Suppose that the mappings α_1, α_2 and α_3 from a domain $O \subset \mathbb{R}^q$ to $\mathbb{R}^q$ are differentiable and satisfy $(\nabla\alpha_1)\alpha_2 - (\nabla\alpha_2)\alpha_1 = \alpha_3$. Let Λ be a $\mathcal{C}^2$-diffeomorphism on O and set $\tilde{\alpha}_i = \{(\nabla\Lambda)\alpha_i\} \circ \Lambda^{-1}$, $i = 1, 2, 3$. Then $(\nabla\tilde{\alpha}_1)\tilde{\alpha}_2 - (\nabla\tilde{\alpha}_2)\tilde{\alpha}_1 = \tilde{\alpha}_3$ on the domain $\Lambda(O)$.*

The following two results are motivated from differential geometry, e.g. Brickell and Clark (1970), Propositions 8.3.2 and 11.5.2. Their full proofs are included because they illustrate how to construct the diffeomorphisms needed in applications. Below the phrase maximal solution means the unique local solution that lasts until the solution leaves Δ or time infinity.

Proposition 1. *Suppose $\Delta \subset \mathbb{R}^p$ is open and σ-compact, and $\alpha \in \mathcal{C}^{\bar{r}}(\Delta; \mathbb{R}^q \setminus \{0\})$. Then, for any $\delta \in \{1, ..., q\}$ and $\bar{x} \in \Delta$ such that $\alpha_\delta(\bar{x}) \neq 0$, there exists a $\bar{x}$-local diffeomorphism $(O^{\bar{x}}, \Lambda_\delta)$ such that $\Lambda_\delta \in \mathcal{C}^{\bar{r}}(O^{\bar{x}}; \mathbb{R}^q)$ and $\{\nabla \Lambda_\delta\}\alpha = e_\delta$ on $O^{\bar{x}}$, where $e_\delta^\top = (0, \ldots, 0, 1, 0, \ldots, 0)$. If $q > 1$ and $\alpha_q \equiv 0$, then $\Lambda_\delta(x) = \begin{bmatrix} \hat{\Lambda}(x) \\ x_q \end{bmatrix}$, where $(x_1, \ldots, x_{q-1})^T \to \hat{\Lambda}(x)$ is a local diffeomorphism for each x_q. Moreover, when $D \subset \Delta$, $\bar{x} \in D$, and $\alpha\big|_D$ only depends on $(x_\delta, ..., x_p)$ we have that*

$$\Lambda_\delta(x) = \sum_{i=1}^{\delta-1} x_i e_i + \begin{bmatrix} M(x_\delta, ..., x_p) \\ L(x_\delta, ..., x_p) \end{bmatrix}, \qquad M(x_\delta, ..., x_p) \in \mathbb{R}^{\delta-1},$$

for the $x \in D$ such that $\dot{y} = \alpha(y)$, $y(0) = (x_1, ..., x_{\delta-1}, \bar{x}_\delta, x_{\delta+1}, ..., x_q)$ stays in D for t between 0 and $x_\delta - \bar{x}_\delta$.

Proof. We let $\theta(t, x)$ be the maximal solution of $\dot{y} = \alpha(y)$, $y(0) = x$ and set $\psi(x) = \theta(x_\delta - \bar{x}_\delta, x_1, \ldots, x_{\delta-1}, \bar{x}_\delta, x_{\delta+1}, \ldots, x_p)$ for the x such that it is well defined so $\partial_{x_\delta}\psi = \alpha(\psi)$. Next,

$$\psi(x_1, \ldots, x_{\delta-1}, \bar{x}_\delta, x_{\delta+1}, \ldots, x_p) = (0, x_1, \ldots, x_{\delta-1}, \bar{x}_\delta, x_{\delta+1}, \ldots, x_p)^\top,$$

so $\nabla\psi(\bar{x})$ has determinant $\alpha_\delta(\bar{x}) \neq 0$. Therefore, applying the Inverse Function Theorem, one obtains that ψ has a inverse $\Lambda_\delta \in \mathcal{C}^{\bar{r}}(O^{\bar{x}}, \mathbb{R}^q)$, where $O^{\bar{x}}$ is a neighborhood of $\bar{x}$. Therefore $\nabla\Lambda_\delta = [\nabla\psi]^{-1}(\Lambda_\delta)$ on $O^{\bar{x}}$. Hence, $(\nabla\Lambda)\alpha = e_\delta$ on $O^{\bar{x}}$ since $\{\partial_{x_\delta}\psi\}(\Lambda) = \alpha(\psi \circ \Lambda) = \alpha$ so clearly $\nabla\Lambda\alpha(x) = e_\delta$ when $x_\delta = \bar{x}_\delta$ and it is easy to show that $\partial_{x_\delta}(\nabla\Lambda\alpha) = 0$ on $O^{\bar{x}}$. The final claims follow easily from the fact ψ will have the form $\psi(x) = (\hat{\psi}(x_1, ..., x_{q-1}, x_q), x_q)$ when $\alpha_q = 0$ and the form

$$\psi(x) = \begin{bmatrix} (x_1 \ldots x_{\delta-1})^\top + f(x_\delta, x_{\delta+1}, ..., x_q) \\ g(x_\delta, x_{\delta+1}, ..., x_q) \end{bmatrix}$$

when α does not depend on $x_1, ..., x_{\delta-1}$. □

In the previous proposition the σ-compact condition was for convenience. It can always be satisfied by making open O smaller if necessary. In the following proposition, the diffeomorphism domains are open subsets of D_T^2. These domains are restricted by intersecting them with D_T in the proof of Theorem 2.

Proposition 2. *Suppose $(\nabla\sigma_j)\sigma_k - (\nabla\sigma_k)\sigma_j = 0$ on D_T, for $1 \leq j, k \leq d$. Then, there exists a $\bar{x}$-local diffeomorphism $(O^{\bar{x}}, \Lambda)$ such that*

$$\{(\nabla\Lambda)\sigma\} \circ \Lambda^{-1} = \begin{pmatrix} I_r & \kappa \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{q \times d} \text{ on } \Lambda(O^{\bar{x}} \cap D_T),$$

where $\kappa \in \mathcal{C}^1(\Lambda(O^{\bar{x}} \cap D_T); \mathbb{R}^{r \times (d-r)})$ does not depend on $x_1, \ldots, x_r$ and $\Lambda \in \mathcal{C}^2(O^{\bar{x}}; \mathbb{R}^q)$. In particular, κ is constant if $r = q$. If $q > r$ and the q^{th} row of

σ is zero, then $\Lambda = \begin{bmatrix} \hat\Lambda(x) \\ x_q \end{bmatrix}$, *where* $(x_1, \ldots, x_{q-1})^T \to \hat\Lambda(x)$ *is a diffeomorphism for each* x_q.

Proof. Suppose there is a $\bar x$-local diffeomorphism $\Lambda_1 \in \mathcal{C}^{r+3-\delta}(O_1; \mathbb{R}^q)$ such that $\{(\nabla\Lambda_1)\sigma\}\circ\Lambda_1^{-1} = \left(e_1\middle| \cdots \middle| e_{\delta-1}\middle| s^\delta\right)$ on $\Lambda_1(O_1 \cap D_T)$ for some $s^\delta \in \mathcal{C}^{r+2-\delta}(\Lambda_1(O_1))$. This is true for $\delta = 2$ by Proposition 1 and we proceed by induction. By Lemma 1,

$$(\nabla s_j^\delta)e_i = (\nabla s_j^\delta)e_i - (\nabla e_i)s_j^\delta = (\nabla\sigma_j)\sigma_i - (\nabla\sigma_i)\sigma_j = 0$$

for all $1 \le i < \delta \le j$ on $\Lambda_1(O_1 \cap D_T)$, and each $s_j^\delta\big|_{D_T}$ depends only on $x_\delta, \ldots, x_q$. However, if y^δ is a local solution to $\dot y = \sigma_\delta(y)$, $y(0) = x$ in O_1, then $z^\delta = \Lambda_1(y^\delta)$ is a local solution to $\dot z^\delta(t) = s_\delta^\delta(z^\delta(t))$, $z^\delta(0) = \Lambda_1(x)$ in $\Lambda_1(O_1)$ that stays in $\Lambda_1(O_1 \cap D_T)$ if started there. Hence, by Proposition 1, we find a $(r+2-\delta)$-times continuously differentiable $\Lambda_1(\bar x)$-local diffeomorphism $(O_\delta, \Lambda_\delta)$ that takes the form $\Lambda_\delta(x) = \sum_{i=1}^{\delta-1} x_i e_i + \begin{bmatrix} M(x_\delta, \ldots, x_q) \\ L(x_\delta, \ldots, x_q) \end{bmatrix}$ on $\Lambda_1(O_1 \cap D_T) \cap O_\delta$ and satisfies $(\nabla\Lambda_\delta)s_\delta^\delta = e_\delta$ on O_δ. Hence, $(O = O_1 \cap \Lambda_1^{-1}(O_\delta), \Lambda = \Lambda_\delta \circ \Lambda_1)$ is a $\bar x$-local diffeomorphism such that $(\nabla\Lambda)\sigma \circ \Lambda^{-1} = \left(e_1\middle| \cdots \middle| e_{\delta-1}\middle| s^\delta\right) \circ \Lambda_\delta^{-1} = \left(e_1\middle| \cdots \middle| e_\delta\middle| s^{\delta+1}\right)$ on $\Lambda(O \cap D_T)$ and by a second, identical application of Lemma 1 $s^{\delta+1}\big|_{D_T}$ does not depend on $x_1, \ldots, x_\delta$. The end result of the induction is a $\bar x$-local diffeomorphism (O, Λ) such that

$$\tilde\sigma = \{(\nabla\Lambda)\sigma\} \circ \Lambda^{-1} = \begin{pmatrix} I_r & \kappa \\ 0 & \bar\kappa \end{pmatrix} \in \mathbb{R}^{q\times d} \text{ on } \Lambda(O \cap D_T),$$

where $\kappa \in \mathbb{R}^{r\times(d-r)}$ and $\bar\kappa \in \mathbb{R}^{(q-r)\times(d-r)}$ do not depend on the variables $x_1, \ldots, x_r$. Since $\tilde\sigma$ has also rank r, it follows that $\bar\kappa = 0$. □

References

[1] Brickell, F. and Clark, R. S. (1970). *Differentiable Manifolds.* Van Nostrand Reinhold, London.

[2] Doss, H. (1977). Liens entre équations différentielles stochastiques et ordinaires. *Ann. Inst. H. Poincaré Probab. Statist.* **13**, 99–125.

[3] Ikeda, N. and Watanabe, S. (1981). *Stochastic Differential Equations and Diffusion Processes*, North-Holland, Amsterdam-Oxford-New York.

[4] Karatzas, I. and Shreve, S.E. (1987). *Brownian Motion and Stochastic Calculus*, Springer, New York.

[5] Kouritzin, M. (1998). On exact filters for continuous signals with discrete observations. *IEEE Trans. Automat. Control* **43**, 709–715.

[6] Kouritzin, M. (2000). Exact infinite dimensional filters and explicit solutions. In *Stochastic Models, A Volume in honour of Donald A. Dawson*, L. Gorostiza and G. Ivanoff Eds, Canadian Mathetical Society Conference Proceedings **26**, American Mathematical Society, Providence, RI, 265–282.

[7] Kunita, H. (1984). Stochastic differential equations and stochastic flows of diffeomorphisms. In *École d'été de probabilités de Saint-Flour, XII—1982*, P. L. Hennequin ed., Lecture Notes in Mathematics **1097**, Springer, Berlin, 143–303.

[8] Rémillard, B. (1994). Chung's law of the iterated logarithm for some stochastic integrals, *Ann. Probab.* **22**, 1794–1802.

[9] Sussmann, H. (1978). On the gap between deterministic and stochastic differential equations. *Ann. Probab.* **6**, 19–41.

[10] Yamato, Y. (1979). Stochastic differrential equations and nilpotent Lie algebras. *Z. Wahrsch. Verw. Gebiete* **47**, 213–229.

Current address: Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton (Alberta), Canada T6G 2G1

E-mail address: mkouritz@math.ualberta.ca

URL: http://www.math.ualberta.ca/People/FacultyPages/Kouritzin/Kouritzin.M.html

Current address: Service de l'enseignement des méthodes quantitatives de gestion, HEC Montréal, Montréal (Québec), Canada H3T 2A7

E-mail address: bruno.remillard@hec.ca

URL: http://www.hec.ca/pages/bruno.remillard