Electronic Journal of Probability

Electron. J. Probab. **24** (2019), no. 3, 1–43.
ISSN: 1083-6489 https://doi.org/10.1214/18-EJP261

Differentiability of SDEs with drifts of super-linear growth *

Peter Imkeller† Gonçalo dos Reis‡ William Salkeld §

Abstract

We close an unexpected gap in the literature of Stochastic Differential Equations (SDEs) with drifts of super linear growth and with random coefficients, namely, we prove Malliavin and Parametric Differentiability of such SDEs. The former is shown by proving Stochastic Gâteaux Differentiability and Ray Absolute Continuity. This method enables one to take limits in probability rather than mean square or almost surely bypassing the potentially non-integrable error terms from the unbounded drift. This issue is strongly linked with the difficulties of the standard methodology of [13, Lemma 1.2.3] for this setting. Several examples illustrating the range and scope of our results are presented.

We close with parametric differentiability and recover representations linking both derivatives as well as a Bismut-Elworthy-Li formula.

Keywords: Malliavin calculus; parametric differentiability; monotone growth SDE; one-sided Lipschitz; Bismut-Elworthy-Li formula.
AMS MSC 2010: Primary 60H07, Secondary 60H10; 60H30.
Submitted to EJP on August 31, 2018, final version accepted on December 29, 2018.

1 Introduction

In this manuscript we work with the class of Stochastic Differential Equations (SDEs) with drifts satisfying a super-linear growth (locally Lipschitz) and a monotonicity condition (also called one-sided Lipschitz condition); the coefficients are furthermore assumed to be random. This class of SDEs appears ubiquitously in mathematics and engineering, for example, the stochastic Ginzburg-Landau equation in the theory of superconductivity;

*P.I. was supported in part by DFG Research Unit FOR 2402. G.d.R. acknowledges support from the *Fundação para a Ciência e a Tecnologia* (Portuguese Foundation for Science and Technology) through the project UID/MAT/00297/2019 (Centro de Matemática e Aplicações CMA/FCT/UNL). W.S. acknowledges support from the Laura Wisewell travel fund.

†Humboldt Universität zu Berlin. E-mail: imkeller@mathematik.hu-berlin.de

‡University of Edinburgh and Centro de Matemática e Aplicações (CMA), FCT, UNL. E-mail: G.dosReis@ed.ac.uk

§University of Edinburgh. E-mail: w.j.salkeld@sms.ed.ac.uk

Stochastic Verhulst equation; Feller diffusion with logistic growth; Protein Kinetics and others, see [7] and references.

There is a wealth of results on differentiability and properties of SDEs in general. However, it is surprising that the landscape is (to the best of our knowledge) sparse with respect to the superlinear growth setting apart from [17] which we discuss below. Additionally, in [15] the authors discuss stochastic flows in rough path sense for a class related to ours but only up to linear growth; and using analytical tools, [1, Chapter 1] and [19] require ellipticity and deterministic maps to obtain some results in the same vein as ours. Our arguments are fully probabilistic.

Malliavin differentiability. To establish Malliavin differentiability for an SDE with solution X and with monotone drifts, the most natural path to follow is to try to apply [13, Lemma 1.2.3] by employing a truncation procedure. This yields a sequence X^n of SDEs with Lipschitz coefficients converging to X. Under said Lipschitz conditions the family X^n is Malliavin differentiable under suitable differentiability assumptions, with derivative DX^n, and one is able to appeal to [13, Lemma 1.2.3] to conclude the Malliavin differentiability of X if one is able to show that $\sup_n \mathbb{E}\big[\|DX^n\|_H\big] < \infty$. The truncation procedure, even smoothed out, destroys the monotonicity and, in the multi-dimensional case, it is notoriously difficult to establish the mentioned uniform bound.

To the best of our knowledge this question was studied only in [17]. The authors employ a truncation procedure in order to use [13, Lemma 1.2.3]. Unfortunately their [17, Lemma 4.1] is incorrect. The constant M_l presented in their equation (4.1) depends on the truncation level n in a non-uniformly bounded way; the reader is invited to inspect the 2nd line of page 879. This lemma, which we were not able to fix, is used subsequently to establish the main result in [17].

We prove Malliavin Differentiability through a less well-known method developed by Sugita [16] which uses the concepts of *Ray Absolute Continuity* and *Stochastic Gâteaux Differentiability* see also the posterior developments by [11, 8]. This approach is detailed in Section 3.2 below. The merit of this method is that the limit for the Stochastic Gâteaux derivative is a convergence in probability statement rather than a convergence in mean square statement. Put simply, this allows us to avoid cases such as the “Witches Hat” function where errors are non-integrable but converge to zero almost surely.

We study the case where the coefficients of the SDE are random. We follow the ideas of [6] and present two different sets of conditions which allow for Malliavin Differentiability. One set of conditions is sharp but somewhat difficult to use in practice. The other is much easier to verify but not sharp. We also provide examples discussing the scope and limitations of our approach.

Parametric differentiability. The second contribution of this work is parametric differentiability for SDEs of this type and in particular its implications for the classical case of deterministic coefficients. The methodology takes inspiration from the Malliavin differentiability section and we prove Gâteaux and Fréchet differentiability with respect to the SDEs parameters.

Representations, Absolute continuity of the law and Bismut-Elworthy-Li formulae. We bridge both differentiability results by recovering (a) representation formulae linking the Malliavin derivative and the parametric one; (b) establishing absolute continuity of the solution’s Law; and (c) a Bismut-Elworthy-Li formula.

Technical results. In this setting the drift term is not bounded and, conditional on the coefficients’ integrability, the solution may not be sufficiently integrable - see Remark 2.3 and the examples in Section 3.3. This means that the error terms appearing in the proofs of differentiability will not be assumed to be sufficiently integrable. We negotiate this obstacle by proving everything in convergence in probability and ensuring that adequate conditions are met so that results can be lifted to the relevant setting of mean square and almost sure convergence. Proposition 2.6 contains a Grönwall type inequality for the topology of Convergence in Probability that is of independent interest and is key to the methods used in this paper.

This paper is organized as follows. In Section 2, we lay out the notation and setting for this paper and recall a few baseline results from the literature. In Section 3 we prove Malliavin differentiability of SDEs of the form (2.1). There are two main results: Theorem 3.2 which provides a sharp method and Theorem 3.7 which has easier to verify Assumptions but is not sharp. There is a collection of examples which explain the merits and limitations of the results we present. In Section 4, we use similar methods to describe the Jacobian of the SDE. Finally, Section 5 bridges Section 3 and Section 4 and contains the so-called representations formulae and existence and smoothness results for densities.

Acknowledgments. The authors thank C. Geiss (U. of Jyväskylä), A. Steinicke (U. of Graz) and A. Réveillac (U. of Toulouse) for their helpful comments. In particular to the two referees whose reviews led to nontrivial improvements of the initial manuscript.

2 Preliminaries

2.1 Notation and spaces

We denote by $\mathbb{N} = \{1, 2, \cdots\}$ the set of natural numbers and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$; $\mathbb{R}$ denotes the set of real numbers respectively; $\mathbb{R}^+ = [0, \infty)$. By $a \lesssim b$ we denote the relation $a \leq C\, b$ where $C > 0$ is a generic constant independent of the relevant parameters and may take different values at each occurrence. By $\lfloor x \rfloor$ we denote the largest integer less than or equal to x. Let A be a $d \times m$ matrix, we denote the Transpose of A by A^T. When A is a matrix, we denote $|A|$ by $\mathrm{Tr}(A \cdot A^T)^{1/2}$.

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a differentiable function. Then we denote ∇f to be the gradient operator and $H[f]$ to be the Hessian operator. ∂_{x_i} is the 1st partial derivative wrt i-th position. $\mathbb{1}_A$ denotes the usual indicator function over some set A

We use standard big O and little o notation to mean that for $f_n, f > 0$

$$f_n = O(f) \iff \limsup_{n\to\infty} \frac{f_n}{f} \leq C \qquad \text{and} \qquad f_n = o(f) \iff \lim_{n\to\infty} \frac{f_n}{f} = 0.$$

where C is a constant independent of the limiting variable.

Let $p \in [1, \infty)$. We introduce the following spaces and when there is no ambiguity about the underlying spaces or measures, we omit their arguments. Let $T > 0$, so that the set $[0, T]$ is a compact subset of the real line.

- Let $C([0, T])$ denote the space of continuous functions $f : [0, T] \to \mathbb{R}$ endowed with the uniform norm $\|f\|_\infty = \sup_{s\in[0,T]} |f(s)|$ and $\|f\|_{\infty,t} = \sup_{s\in[0,t]} |f(s)|$; $C_0([0, T])$ be the subspace of continuous functions that start at 0; $C_b^k(\mathbb{R}^m)$ the set of k-times differentiable real valued maps defined on $\mathbb{R}^m$ with bounded partial derivatives up to order k, and $C_b^\infty(\mathbb{R}^m) = \cap_{k\geq 1} C_b^k(\mathbb{R}^m)$; C_b^0 its subspace of continuous bounded functions;
- Let $L^p([0, T])$ denote the space of functions $f : [0, T] \to \mathbb{R}$ satisfying $\|f\|_p = \big(\int_0^T |f(r)|^p dr\big)^{1/p} < \infty$.
- Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{Q})$ be a probability space. Let $L^p(\mathcal{F}_t; \mathbb{R}^d; \mathbb{Q})$, $t \in [0, T]$, is the space of $\mathbb{R}^d$-valued $\mathcal{F}_t$-measurable random variables X with norm $\|X\|_{L^p} = \mathbb{E}^{\mathbb{Q}}[\,|X|^p]^{1/p} < \infty$; L^∞ refers to the subset of bounded random variables with norm $\|X\|_{L^\infty} = \operatorname{ess\,sup}_{\omega\in\Omega} |X(\omega)|$;

Let $L^0(\mathcal{F}_t;\mathbb{R}^d)$ be the space of $\mathbb{R}^d$-valued $\mathcal{F}_t$-measurable, adapted random variables with the topology of convergence in probability.

- $\mathcal{S}^p([0,T],\mathbb{R}^m,\mathbb{Q})$ is the space of $\mathbb{R}^d$-valued processes $(Y_t)_{t\in[0,T]}$ that are $\mathcal{F}$-adapted and satisfying $\|Y\|_{\mathcal{S}^p}=\mathbb{E}^{\mathbb{Q}}[\|Y\|_\infty^p]^{1/p}=\mathbb{E}^{\mathbb{Q}}[\sup_{t\in[0,T]}|Y(t)|^p]^{1/p}<\infty$; additionally $\mathcal{S}^\infty([0,T],\mathbb{R}^m,\mathbb{Q})$ refers to the intersection of $\mathcal{S}^p([0,T],\mathbb{R}^m,\mathbb{Q})$ for every $p\geq 1$.
- Let H be the usual Cameron-Martin Hilbert space for Brownian Motion

$$H=\Big\{h(t)=\int_0^t \dot{h}(s)ds,\ t\in[0,T];\ h(0)=0,\ \dot{h}\in L^2([0,T])\Big\}.$$

2.2 Malliavin calculus

Here we briefly outline the ideas of Malliavin calculus. For more details, see [13]. We denote by $\mathbb{E}$ and $\mathbb{E}[\cdot|\mathcal{F}_t]$ the usual expectation and conditional expectation operator (wrt to $\mathbb{P}$) respectively. For a random variable X we denote its probability distribution (or Law) by $\mathcal{L}^X$; the law of a process $(Y(t))_{t\in[0,T]}$ at time t is denoted by $\mathcal{L}^Y_t$.

The probability space

Let $\tilde{\Omega}=C_0([0,T];\mathbb{R}^m)$ be the canonical m-dimensional Wiener space and let W be the Wiener process with law $\tilde{\mathbb{P}}$. Let $(\tilde{\mathcal{F}}_t)_{t\in[0,T]}$ be the standard augmentation of the filtration generated by the Brownian motion. Then we have the probability space $(\tilde{\Omega},\tilde{\mathcal{F}},(\tilde{\mathcal{F}}_t)_{t\in[0,T]},\tilde{\mathbb{P}})$. Additionally, let $([0,1],\mathcal{B}([0,1]),\overline{\mathbb{P}})$ be a probability space with the Lebesgue measure $\overline{\mathbb{P}}$. Our probability space is structured as follows:

1. The sample space will be $\Omega=[0,1]\times\tilde{\Omega}$
2. The σ-algebra over this space will be $\mathcal{F}=\sigma(\mathcal{B}([0,1])\times\tilde{\mathcal{F}})$ with filtration $\mathcal{F}_t=\sigma(\mathcal{B}([0,1])\times\tilde{\mathcal{F}}_t)$.
3. The probability measure will be the product measure $\mathbb{P}=\overline{\mathbb{P}}\times\tilde{\mathbb{P}}$.

Thus for a random variable θ that is $\mathcal{F}_0=\sigma\big(\mathcal{B}([0,1])\times\tilde{\mathcal{F}}_0\big)$ measurable, the Malliavin Derivative of θ is trivially 0. Hence when for $\omega\in\Omega$ we write $\omega+h$ in Section 3, we think of this perturbation as only being in the canonical Wiener sense. In Section 4, we perturb on the space $L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})$.

Malliavin derivatives via cylindrical functions

Let $\mathcal{H}$ be a Hilbert space and let E, B be separable Banach spaces. Let $\mathfrak{i}:\mathcal{H}\to E$ be an injective, continuous linear map with dense image. The triple $(\mathcal{H},E,\mathfrak{i})$ are called an Abstract Wiener space. The map $\mathfrak{i}$ radonifies the canonical Gaussian cylinder set measure over $\mathcal{H}$ and the induced measure over E is called the Abstract Wiener measure.

Let $W:\mathcal{H}\to L^2(\mathcal{F};\mathcal{R};\mathbb{P})$ a Gaussian random variable. The space $W(\mathcal{H})$ endowed with an inner product $\langle W(h_1),W(h_2)\rangle=\mathbb{E}[W(h_1)W(h_2)]$ is a Gaussian Hilbert space.

A mapping $f:E\to\mathbb{R}$ is called a Polynomial if $\exists n\in\mathbb{N}$, $\exists g_1,...,g_n\in E^*$ and $\exists\tilde{f}:\mathbb{R}^n\to\mathbb{R}$ such that

$$f(x)=\tilde{f}\big(g_1(x),...,g_n(x)\big).$$

The set of all polynomials is denoted $\mathbf{P}$. Similarly, a mapping $F:E\to B$ is said to be an B-valued polynomial if $\exists m\in\mathbb{N}$, $\exists f_1,...,f_m\in\mathbf{P}$ and $\exists b_1,...,b_m\in B$ such that

$$F(x)=\sum_{j=1}^m f_j(x)b_j.$$

The set of all B-valued polynomials is denoted by $\mathbf{P}[B]$. It is well documented that $\mathbf{P}[B]$ is dense in the space $L^p(\mathcal{F};E;\mathbb{P})$ where $\mathbb{P}$ is the induced measure over E. See for instance [16]. For the canonical space, $E=C_0([0,T];\mathbb{R}^m)$ is the space of all continuous $\mathbb{R}^m$ valued paths starting at 0 with uniform norm.

For $F\in\mathbf{P}[B]$, we define the derivative of F to be the $B\otimes\mathcal{H}$ (tensor product of B and H) valued

$$DF(x)=\sum_{k=1}^{n}\sum_{j=1}^{m}\partial_{x_k}f_j\big(g_1(x),...,g_n(x)\big)e_j\otimes(g_k\circ i).$$

Recall that $(g_k\circ i)$ can be thought of as an element of $\mathcal{H}^*$, which is isometrically isomorphic to $\mathcal{H}$. The operator D is closable, and we define the Sobolev space $\mathbb{D}^{1,p}(B)$ to be the closure of $\mathbf{P}[B]$ with respect to the norm

$$\|F\|_{1,p,B}=\left[\int_E\Big\|F(x)\Big\|_B^p d\mathbb{P}(x)+\int_E\Big\|\|DF(x)\|_{\mathcal{H}}\Big\|_B^p d\mathbb{P}(x)\right]^{\frac{1}{p}}.$$

2.3 Existence and uniqueness of SDE with local Lipschitz coefficients

We present the class of SDEs that we will be working with.

Lipschitz and locally Lipschitz coefficients

Let $(t,\omega,\theta)\in[0,T]\times\Omega\times L^0(\mathcal{F}_0;\mathbb{P};\mathbb{R}^d)$.

In this paper, we prove differentiability properties of the SDE

$$X_\theta(t)(\omega)=\theta+\int_0^t b\Big(s,\omega,X(s)(\omega)\Big)ds+\int_0^t\sigma\Big(s,\omega,X(s)(\omega)\Big)dW(s), \tag{2.1}$$

driven by a m-dimensional Brownian motion W.

Assumption 2.1. *Let $p\geq 2$. Let $\theta:\Omega\to\mathbb{R}^d$, $b:[0,T]\times\Omega\times\mathbb{R}^d\to\mathbb{R}^d$ and $\sigma:[0,T]\times\Omega\times\mathbb{R}^d\to\mathbb{R}^{d\times m}$ be progressively measurable maps and $L>0$ such that:*

- $\theta\in L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})$.
- *b and σ are integrable in the sense that*

$$\mathbb{E}\Big[\Big(\int_0^T|b(t,\omega,0)|dt\Big)^p\Big],\ \mathbb{E}\Big[\Big(\int_0^T|\sigma(t,\omega,0)|^2dt\Big)^{\frac{p}{2}}\Big]<\infty. \tag{2.2}$$

- *$\exists L$ such that for almost all $(s,\omega)\in[0,T]\times\Omega$ and $\forall x,y\in\mathbb{R}^d$ we have*

$$\Big\langle x-y,b(s,\omega,x)-b(s,\omega,y)\Big\rangle_{\mathbb{R}^d}\leq L|x-y|^2\quad\text{and}\quad|\sigma(s,\omega,x)-\sigma(s,\omega,y)|\leq L|x-y|.$$

- *For $x,y\in\mathbb{R}^d$ such that $|x|,|y|<N$, $\exists L_N>0$ such that*

$$|b(s,\omega,x)-b(s,\omega,y)|\leq L_N|x-y|,$$

for almost all $(s,\omega)\in[0,T]\times\Omega$.

The next result extends results found in the literature to the case of random coefficients. Existence and uniqueness of a solution follow the methods of [10, Theorem 2.3.6]; the case of random coefficients is not addressed there but the general methodology is applicable in the same way with only more care being taken when proving integrability.

Theorem 2.2. *Let $p \geq 2$. Suppose Assumption 2.1 is satisfied. Then there exists a unique solution $(X(t))_{t\in[0,T]}$ to the SDE (2.1) in $\mathcal{S}^p$ and*

$$\mathbb{E}\Big[\|X\|_\infty^p\Big] \lesssim \Big(\mathbb{E}\Big[|\theta|^p\Big] + \mathbb{E}\Big[\Big(\int_0^T |b(s,\omega,0)|ds\Big)^p\Big] + \mathbb{E}\Big[\Big(\int_0^T \Big|\sigma(s,\omega,0)\Big|^2 ds\Big)^{\frac{p}{2}}\Big]\Big).$$

Moreover, the map $t \mapsto X(t)(\omega)$ is $\mathbb{P}$-a.s. continuous.

Finally, the solution of the SDE is Stochastically Stable in the sense that for $\forall \xi, \theta \in L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})$, (with a constant depending on the other parameters but not on θ or ξ)

$$\mathbb{E}\Big[\|X_\xi - X_\theta\|_\infty^p\Big] \lesssim \mathbb{E}\Big[|\theta - \xi|^p\Big].$$

Proof. This proof can be found in Appendix A.1. □

Remark 2.3 (Issues with integrability and Fubini - Sharp conditions)**.** The integrability conditions of Assumption 2.1 are designed to be sharp. However, they yield processes which can have some problematic properties.

It is very important to note that we cannot (in general) swap the order of integration at this point! This is a key point in our manuscript. We are not able to assume that the drift term is sufficiently integrable (given (2.2)) and hence the error terms appearing in the proofs of differentiability below will not be assumed to be integrable.

To emphasize our point consider the following monotone drift function $b(t,\omega,x) = x - x^5$ and $\sigma(t,\omega,x)$ is chosen so that for some $t' \in [0,T)$

$$\mathbb{E}\Big[\Big|\int_0^T |\sigma(t,\omega,0)|^2 dt\Big|^2\Big] < \infty, \quad \mathbb{E}\Big[\Big|\int_0^{t'} |\sigma(t,\omega,0)|^2 dt\Big|^{\frac{5}{2}}\Big] = \infty.$$

These satisfy the conditions of Assumption 2.1 for $p = 4$ but not for $p = 5$. We can then argue as follows: for $t \in [t', T]$

$$\mathbb{E}[\,|X(t)|^4] < \infty, \quad \mathbb{E}[\,|X(t)|^5] = \infty \quad \text{and in particular} \quad \mathbb{E}\Big[\int_{t'}^t |X(s)|^5 ds\Big] = \infty.$$

The existence of finite fourth moments ensures we have finite first moments and hence for $t > t'$

$$\mathbb{E}\Big[\int_0^t \Big(X(s) - X(s)^5\Big)ds\Big] < \infty \quad \text{which implies that} \quad \mathbb{E}\Big[\int_{t'}^t X(s)^5 ds\Big] < \infty.$$

On SDEs with linear coefficients

Let $(t,\omega,\theta) \in [0,T] \times \Omega \times L^0(\mathcal{F}_0;\mathbb{P};\mathbb{R}^d)$ and take an SDE of the form

$$\begin{aligned} X_\theta(t)(\omega) = \theta &+ \int_0^t \Big[B\big(s,\omega\big)X_\theta(s)(\omega) + b(s,\omega)\Big]ds \\ &+ \int_0^t \Big[\Sigma\big(s,\omega\big)X(s)(\omega) + \sigma(s,\omega)\Big]dW(s), \end{aligned} \tag{2.3}$$

driven by a m-dimensional Brownian motion W. The derivatives of SDEs of the form (2.1) will satisfy linear SDEs of the form (2.3).

Assumption 2.4. *Let $p \geq 1$. Let $B : [0,T] \times \Omega \to \mathbb{R}^{d\times d}$, $\Sigma : [0,T] \times \Omega \to \mathbb{R}^{(d\times m)\times d}$, $b : [0,T] \times \Omega \to \mathbb{R}^d$ and $\sigma : [0,T] \times \Omega \to \mathbb{R}^{d\times m}$ be progressively measurable maps such that:*

- $\theta \in L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})$.

- *B, b, Σ and σ are integrable in the sense that $\exists L\geq 0$ such that $\forall x\in\mathbb{R}^d$*

$$x^T B(t,\omega)x < L|x|^2 \quad \mathbb{P}\text{-a.s.}, \qquad \int_0^T \|\Sigma(t,\cdot)\|_{L^\infty}^2 dt < \infty,$$
$$\mathbb{E}\Big[\Big(\int_0^T |b(t,\omega)|dt\Big)^p\Big], \quad \mathbb{E}\Big[\Big(\int_0^T |\sigma(t,\omega)|^2 dt\Big)^{\frac{p}{2}}\Big] < \infty.$$

One advantage of SDEs of the form (2.3) is that they have an explicit solution unlike SDEs of the form (2.1) where a solution exists but cannot be explicitly stated. Linear SDEs do have Lipschitz coefficients, but their Lipschitz constants are not uniform over $(t,\omega)\in[0,T]\times\Omega$. Therefore, we cannot apply Theorem 2.2.

Notice that for Assumption 2.4, we do not make any requirement on B being positive definite operator. In fact, we may be interested in cases where $\exists x\in\mathbb{R}^d$ such that $x^T(\int_0^T B(t,\omega)dt)x = -\infty$ with positive probability.

Theorem 2.5. *Let $p\geq 1$. Suppose Assumption 2.4 is satisfied. Then there exists a unique solution $(X(t))_{t\in[0,T]}$ to the SDE (2.3) in $\mathcal{S}^p$ with explicit form*

$$X_\theta(t) = \Psi(t)\Bigg(\theta + \int_0^t \Psi(s)^{-1}\Big[b(s,\omega) - \Big\langle \Sigma(s,\omega), \sigma(s,\omega)\Big\rangle_{\mathbb{R}^m}\Big]ds + \int_0^t \Psi(s)^{-1}\sigma(s,\omega)dW(s)\Bigg),$$

where $\Psi:[0,T]\times\Omega\to\mathbb{R}^{d\times d}$ can be written as

$$\Psi(t) = I_d \exp\Bigg(\int_0^t \Big[B(s,\omega) - \frac{\big\langle\Sigma(s,\omega),\Sigma(s,\omega)\big\rangle_{\mathbb{R}^m}}{2}\Big]ds + \int_0^t \Sigma(s,\omega)dW(s)\Bigg), \tag{2.4}$$

and

$$\mathbb{E}\Big[\|X_\theta\|_\infty\Big] \lesssim \Big(\mathbb{E}\Big[|\theta|^p\Big] + \mathbb{E}\Big[\Big(\int_0^T |b(s,\omega)|ds\Big)^p\Big] + \mathbb{E}\Big[\Big(\int_0^T |\sigma(s,\omega)|^2 ds\Big)^{\frac{p}{2}}\Big]\Big).$$

Moreover, the map $t\mapsto X(t)(\omega)$ is $\mathbb{P}$-a.s. continuous.

Finally, the solution X_θ of the equation is Stochastically stable in the sense that $\forall \xi,\theta\in L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})$

$$\mathbb{E}\Big[\|X_\xi - X_\theta\|_\infty^p\Big] \lesssim \mathbb{E}\Big[|\xi-\theta|^p\Big].$$

Proof. An existence and uniqueness proof is found in [10, Theorem 3.3.1]. Moment calculations are proved in Appendix A.1. Stochastic stability is proved in the same fashion as in Theorem 2.2. □

2.4 A Grönwall inequality

To the best of our knowledge the next result is new and of independent interest. While unsurprising, this is key to the methods of this paper.

Proposition 2.6 (Grönwall Inequality for the Topology of Convergence in Probability)**.** *Let $n\in\mathbb{N}$, $A_n:[0,T]\times\Omega\to\mathbb{R}$ be a sequence of adapted stochastic processes such that $\|A_n\|_\infty \xrightarrow{\mathbb{P}} 0$ as $n\to\infty$ (that is conv. in probability: $\forall\varepsilon>0 \lim_{n\to\infty}\mathbb{P}[\|U_n\|_\infty>\varepsilon]=0$).*

Let U_n be the solution of the SDE

$$U_n(t) = A_n(t) + \int_0^t f\big(U_n(s)\big)ds + \int_0^t g\big(U_n(s)\big)dW(s), \qquad t\in[0,T]$$

where $f,g:\mathbb{R}\to\mathbb{R}$ are Monotone growth and Lipschitz respectively (see 3rd bullet point of Assumption 2.1) and $f(0)=g(0)=0$.

Then $\|U_n\|_\infty \xrightarrow{\mathbb{P}} 0$ as $n\to\infty$.

Notice that since we do not have finite second moments of $\|A_n\|_\infty$, the result cannot be proved using a mean square type argument.

Proof. Fix $\delta > 0$ and let $n \in \mathbb{N}$. We have for any choice of $\eta > 0$ that

$$\mathbb{P}\Big[\|U_n\|_\infty > \delta\Big] \leq \mathbb{P}\Big[\|U_n\|_\infty > \delta, \|A_n\|_\infty \leq \eta\Big] + \mathbb{P}\Big[\|A_n\|_\infty > \eta\Big].$$

We already have that $\lim_{n\to\infty} \mathbb{P}\big[\|A_n\|_\infty > \eta\big] = 0$ for any choice of $\eta > 0$ by assumption. Define the sequence of stopping times $\tau_n = \inf\{t' > 0 : |A_n(t')| > \eta\}$, $n \in \mathbb{N}$.

Firstly, we show that $\lim_{n\to\infty} \tau_n \geq T$ almost surely. Suppose this was not the case. Then $\exists \Omega' \subset \Omega$ with $\mathbb{P}(\Omega') > 0$ and $\forall \omega \in \Omega'$ $\exists n_k(\omega)$ an increasing subsequence of integers such that $\tau_{n_k}(\omega) < T$ for all $k \in \mathbb{N}$. Then $\forall \omega \in \Omega'$, $\|A_{n_k}\|_\infty(\omega) > \eta$ for all $k \in \mathbb{N}$. But that implies that for any $k \in \mathbb{N}$ we have

$$\Omega' \subset \{\omega \in \Omega; \|A_{n_k}\|_\infty(\omega) > \eta\} \quad \text{and hence that} \quad \mathbb{P}\big[\|A_{n_k}\|_\infty > \eta\big] > \mathbb{P}[\Omega'].$$

The latter contradicts the assumption that $\|A_{n_k}\|_\infty$ converges to 0 in probability. So any such set Ω' must have measure 0 and we conclude $\lim_{n\to\infty} \tau_n > T$ almost surely.

The SDE for $U_n(t)$ is well defined for $t \in [0, \tau_n]$. Outside of this interval, A_n may not be integrable so we may not be able construct a solution. However $\forall \omega \in \Omega$ such that $\|A_n\|_\infty(\omega) \leq \eta$ we have that $\tau_n(\omega) > T$. Therefore

$$\mathbb{P}\Big[\|U_n(\cdot)\|_\infty > \delta, \|A_n\|_\infty \leq \eta\Big] = \mathbb{P}\Big[\|U_n(\cdot \wedge \tau_n)\|_\infty > \delta, \|A_n\|_\infty \leq \eta\Big],$$

because the process $U_n(\cdot)$ and the stopped process $U_n(\cdot \wedge \tau_n)$ are $\mathbb{P}$-almost surely equal when one restricts to the event where $\|A_n\|_\infty \leq \eta$.

As we know that the solution $U_n(t \wedge \tau_n)$ will exist and make sense, it serves to introduce this stopping time. Thus we get

$$\begin{aligned}\mathbb{P}\Big[\|U_n\|_\infty > \delta\Big] &\leq \mathbb{P}\Big[\|U_n(\cdot \wedge \tau_n)\|_\infty > \delta, \|A_n\|_\infty \leq \eta\Big] + \mathbb{P}\Big[\|A_n\|_\infty > \eta\Big] \\ &\leq \mathbb{P}\Big[\|U_n(\cdot \wedge \tau_n)\|_\infty > \delta\Big] + \mathbb{P}\Big[\|A_n\|_\infty > \eta\Big].\end{aligned}$$

Now we consider the SDE for $U_n(t \wedge \tau_n)$. The stopping time prevents the term $A_n(t \wedge \tau_n)$ from getting any larger that η and ensures that the stochastic integral is a local martingale. Appealing to Theorem 2.2 yields existence/uniqueness of the solution and moment bounds.

$$\mathbb{E}\Big[\|U_n(\cdot \wedge \tau_n)\|_\infty^2\Big] < \eta^2 e^C \quad \text{and therefore} \quad \mathbb{P}\Big[\|U_n\|_\infty > \delta\Big] \leq \frac{\eta^2 e^C}{\delta^2} + \mathbb{P}\Big[\|A_n\|_\infty > \eta\Big].$$

Choose η such that $\eta^2 e^C/\delta^2 < \varepsilon'/2$. Then find $N \in \mathbb{N}$ such that $\forall n \geq N$ $\mathbb{P}\big[\|A_n\|_\infty > \eta\big] < \varepsilon'/2$. This concludes the proof. □

3 Malliavin Differentiability of SDEs with monotone coefficients

In this section we prove two Malliavin differentiability result for SDEs in the class given by Assumption 2.1. We use a less known method using the concepts of *Ray absolute continuity* and *Stochastic Gâteaux Differentiability* initiated by [16] and later developed by [11, 8].

For SDEs of the form (2.1), the proof of existence and uniqueness of a solution involves a sequence of random variables which converge almost surely to the solution rather than in mean square. Indeed this sequence of random variables does not converge in mean square, unlike in the proof of Existence and Uniqueness for SDEs with Lipschitz coefficients. This means that the classical method from [13, Lemma 1.2.3] cannot be applied; recall further our observation on the role that Proposition 2.6 will play here.

3.1 Main results and their assumptions

We state the main assumptions and results with the proofs postponed for later sections.

Assumption 3.1. *Let $b:[0,T]\times\Omega\times\mathbb{R}^d\to\mathbb{R}^d$ and $\sigma:[0,T]\times\Omega\times\mathbb{R}^d\to\mathbb{R}^{d\times m}$ satisfy Assumption 2.1 for some $p>2$. Further, suppose*

(i) For almost all $(t,\omega)\in[0,T]\times\Omega$ the functions $\sigma(t,\omega,\cdot)$ and $b(t,\omega,\cdot)$ have spatial partial derivatives in all directions.

(ii) For all $h\in H$ and $(\varepsilon,x)\in\mathbb{R}^+\times\mathbb{R}^d$, we have that the maps $\mathbb{R}^+\times\mathbb{R}^d\to L^0(\Omega)$

$$(\varepsilon,x)\mapsto\int_0^T\Big|\nabla_x\sigma(t,\omega+\varepsilon h,x)\Big|^2dt \qquad \text{and} \qquad (\varepsilon,x)\mapsto\int_0^T\Big|\nabla_x b(t,\omega+\varepsilon h,x)\Big|^2dt,$$

are jointly continuous (where convergence in L^0 means convergence in probability).

(iii) $\exists U:[0,T]^2\times\Omega\to\mathbb{R}^{d\times m}$ and $V:[0,T]^2\times\Omega\to\mathbb{R}^{(d\times m)\times m}$ which satisfy that for $s>r$ $U(s,r,\omega)=V(s,r,\omega)=0$ and

$$\mathbb{E}\Big[\Big(\int_0^T\Big(\int_0^T\Big|U(s,r,\omega)\Big|^2ds\Big)^{\frac12}dr\Big)^p\Big]<\infty \quad \text{and}$$
$$\mathbb{E}\Big[\Big(\int_0^T\int_0^T\Big|V(s,r,\omega)\Big|^2dsdr\Big)^{\frac{p}{2}}\Big]<\infty.$$

(iv) b and σ satisfy, as $\varepsilon\to0$, that $\forall h\in H$

$$\mathbb{E}\Big[\Big(\int_0^T\Big|\frac{b(r,\omega+\varepsilon h,X(r))-b(r,\omega,X(r))}{\varepsilon}-\int_0^r U(s,r,\omega)\dot h(s)ds\Big|dr\Big)^2\Big]\to0,$$
$$\mathbb{E}\Big[\int_0^T\Big|\frac{\sigma(r,\omega+\varepsilon h,X(r))-\sigma(r,\omega,X(r))}{\varepsilon}-\int_0^r V(s,r,\omega)\dot h(s)ds\Big|^2dr\Big]\to0.$$

In the above condition neither b or σ are assumed to be in $\mathbb{D}^{1,2}$, they are only assumed to be Malliavin differentiable over the sub-manifold on which X (solution to (2.1)) takes values on. After our main results we give examples of SDE illustrating the scope of our assumptions.

Theorem 3.2 (Malliavin Differentiability of Monotone SDEs)**.** *Take $p>2$. Let Assumption 3.1 hold and denote by X the unique solution of the SDE* (2.1) *in $\mathcal{S}^p$.*

Then X is Malliavin differentiable, i.e. $X\in\mathbb{D}^{1,p}(\mathcal{S}^p)$ and there exist adapted processes U and V such that the Malliavin derivative satisfies for $0\le s\le t\le T$

$$\begin{aligned}D_sX(t)(\omega)=&\sigma(s,\omega,X(s)(\omega))+\int_s^t U(s,r,\omega)dr+\int_s^t V(s,r,\omega)dW(r)\\&+\int_s^t\nabla_x b(r,\omega,X(r)(\omega))D_sX(r)(\omega)dr\\&+\int_s^t\nabla_x\sigma(r,\omega,X(r)(\omega))D_sX(r)(\omega)dW(r),\end{aligned}\tag{3.1}$$

and otherwise $D_sX(t)=0$ for $s>t$.

The proof of Theorem 3.2 can be found in Section 3.4.

Remark 3.3 (Notation)**.** At the simplest level, we have X is $\mathbb{R}^d$-valued and W is $\mathbb{R}^m$-valued. Therefore b, σ are $\mathbb{R}^d$- and $\mathbb{R}^{d\times m}$-valued respectively. Hence we have the collection of one-dimensional SDEs

$$X^{(i)}(t)(\omega) = \theta^{(i)} + \int_0^t b^{(i)}(s,\omega,X(s)(\omega))ds + \sum_{j=1}^m \int_0^t \sigma^{(i,j)}(s,\omega,X(s)(\omega))dW^{(j)}(s),$$

where i is an integer between 1 and d.

The Malliavin Derivative DX is therefore a $\mathbb{R}^{d\times m}$ valued process and we get the system of equations

$$\begin{aligned}D_s^{(k)} X^{(i)}(t)(\omega) =& \sigma^{(i,k)}(s,\omega,X(s)(\omega))ds \\ &+ \int_s^t U^{(i,k)}(s,r,\omega)dr + \sum_{j=1}^m \int_s^t V^{(i,j,k)}(s,r,\omega)dW^{(j)}(r) \\ &+ \int_s^t \Big\langle (\nabla_x b^{(i)})(r,\omega,X(r)(\omega)), D_s^{(k)} X(t)(\omega) \Big\rangle_{\mathbb{R}^d} dr \\ &+ \sum_{j=1}^m \int_s^t \Big\langle (\nabla_x \sigma^{(i,j)})(r,\omega,X(r)(\omega)), D_s^{(k)} X(t)(\omega) \Big\rangle_{\mathbb{R}^d} dW^{(j)}(r),\end{aligned}$$

for i an integer between 1 and d and k an integer between 1 and m.

Remark 3.4 (Mollification and non-differentiability of b and σ)**.** Using classic mollification arguments the assumptions of Theorem 3.2 concerning the behaviour of $x \mapsto b(\cdot,\cdot,x)$ and $x \mapsto \sigma(\cdot,\cdot,x)$ can be further weakened. Namely, σ can be assumed to be uniformly Lipschitz as opposed to continuously differentiable and b can be assumed to have left- and right-derivatives not necessarily equal to each other at every point.

Under these conditions, a canonical mollification argument allows to re-obtain Theorem 3.2 where in (3.1) one replaces $\nabla_x b$ and $\nabla_x \sigma$ by two processes corresponding to their generalized derivatives.

If b and σ are assumed deterministic then one immediately obtains the familiar result.

Corollary 3.5 (Deterministic coefficients case)**.** *Suppose that $b : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0,T]\times\mathbb{R}^d \to \mathbb{R}^{d\times m}$ satisfy Assumption 2.1. Further, suppose that $x \mapsto b(\cdot,x)$ and $x \mapsto \sigma(\cdot,x)$ are continuously differentiable in their spatial variables (uniformly in t).*

Then X is Malliavin differentiable and $D_sX(t) = 0$ for $T \geq s > t \geq 0$ while for $0 \leq s \leq t \leq T$

$$\begin{aligned}D_sX(t)(\omega) = \sigma(s,X(s)(\omega)) + \int_s^t \nabla_x b(r,X(r)(\omega))D_sX(r)(\omega)dr \\ + \int_s^t \nabla_x\sigma(r,X(r)(\omega))D_sX(r)(\omega)dW(r).\end{aligned}$$

Assumption 3.1 is sharp for our construction, nonetheless, it can be slightly strengthened to Assumption 3.6 which is much easier to verify.

Assumption 3.6. *Let $b : [0,T]\times\Omega\times\mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0,T]\times\Omega\times\mathbb{R}^d \to \mathbb{R}^{d\times m}$ satisfy Assumption 2.1 for $p > 2$. Further, suppose Assumption 3.1 (i) and (ii) hold and*

(iii') *b and σ are Malliavin differentiable in the sense that*

$$\forall x \in \mathbb{R}^d,\ b(\cdot,\cdot,x) \in \mathbb{D}^{1,p}\big(L^1([0,T];\mathbb{R}^d)\big) \quad \textit{and} \quad \sigma(\cdot,\cdot,x) \in \mathbb{D}^{1,p}\big(L^2([0,T];\mathbb{R}^{d\times m})\big).$$

(iv') *The Malliavin derivatives of b and σ are progressively measurable and Lipschitz in their spacial variables i.e. $\exists L > 0$ constant such that $\forall (s,t) \in [0,T]^2$ and $x, y \in \mathbb{R}^d$,*

$\mathbb{P}$*-almost surely*

$$|D_s b(t,\omega,x) - D_s b(t,\omega,y)| \le L|x-y|,$$
$$|D_s \sigma(t,\omega,x) - D_s \sigma(t,\omega,y)| \le L|x-y|.$$

The second main result of the section is the following theorem.

Theorem 3.7. *Let $p > 2$. Let Assumption 2.1 hold and denote by X the unique solution of the SDE (2.1) in $\mathcal{S}^p$. Let b and σ satisfy Assumption 3.6. Then the conclusion of Theorem 3.2 still holds: $X \in \mathbb{D}^{1,p}(\mathcal{S}^p)$ and DX satisfies $D_s X(t) = 0$ for $T \ge s > t \ge 0$ while for $0 \le s \le t \le T$*

$$\begin{aligned} &D_s X(t)(\omega) \\ &= \sigma\big(s,\omega,X(s)(\omega)\big) + \int_s^t (D_s b)\big(r,\omega,X(r)(\omega)\big)dr + \int_s^t (D_s\sigma)\big(r,\omega,X(r)(\omega)\big)dW(r) \\ &\quad + \int_s^t \nabla_x b\big(r,\omega,X(r)(\omega)\big)D_s X(r)(\omega)dr + \int_s^t \nabla_x\sigma\big(r,\omega,X(r)(\omega)\big)D_s X(r)(\omega)dW(r). \end{aligned} \tag{3.2}$$

The proof can be found in Section 3.5. We point out that the mollification Remark 3.4 applies to this result as well.

It is a well documented fact, see [13, Theorem 2.2.1], that if one has a SDE with deterministic and Lipschitz drift and diffusion coefficients then the Malliavin derivative is the solution of a homogeneous linear SDE. Both the SDE and the Malliavin Derivative have finite moments of all orders. Therefore the solution of the SDE exists in $\mathbb{D}^{1,\infty}$.

We study the case where the coefficients are random. SDEs of this kind do not always have finite moments of all orders, and the same will apply for the Malliavin derivative. In fact, the integrability of the derivative comes directly from the integrability of the Malliavin derivatives of b and σ.

3.2 Overview of the methodology and results on Wiener spaces

It is important to note that the solution of an SDE is not continuous with respect to $\omega \in \Omega$. As the SDE exists in a probability space with the filtration generated by an m-dimensional Brownian motion, ω can be interpreted to mean the path of an individual Brownian motion plus any extra information about what happens when $t = 0$. However, it will be shown that the random variables are continuous, and indeed differentiable, when perturbed with respect to a path out of the Cameron Martin space. Hence for this section we take $h \in H^{\otimes m}$, an m-dimensional Cameron Martin path and $\dot{h}$ to be its derivative unless stated otherwise. We will not emphasize the difference between H and $H^{\otimes m}$ in this paper.

We start by introducing the concepts of *Ray absolute continuity* and *Stochastic Gâteaux Differentiability* and the results yielding Malliavin differentiability under those properties.

Let E be a separable Banach space. Let $L(H,E)$ be the space of all bounded linear operators $V : H \to E$.

Definition 3.8 (Ray Absolutely Continuous map)**.** *A measurable map $f : \Omega \to E$ is said to be* Ray Absolutely Continuous *if $\forall h \in H$, $\exists$ a measurable mapping $\tilde{f}_h : \Omega \to E$ such that*

$$\tilde{f}_h(\omega) = f(\omega) \quad \mathbb{P}\text{-a.e.}$$

and that $\forall \omega \in \Omega$,

$$t \mapsto \tilde{f}_h(\omega + th) \quad \text{is absolutely continuous on any compact subset of } \mathbb{R}.$$

Definition 3.9 (Stochastically Gâteaux differentiable)**.** *A measurable mapping* $f : \Omega \to E$ *is said to be* Stochastically Gâteaux differentiable *if there exists a measurable mapping* $F : \Omega \to L(H, E)$ *such that* $\forall h \in H$*,*

$$\frac{f(\omega + \varepsilon h) - f(\omega)}{\varepsilon} \xrightarrow{\mathbb{P}} F(\omega)[h] \quad \textit{as } \varepsilon \to 0.$$

Malliavin differentiability follows from [16, Theorem 3.1] which was later improved upon by [11, Theorem 4.1]. We recall both results next.

Theorem 3.10 ([16])**.** *Let* $p > 1$*. The space* $\mathbb{D}^{1,p}(E)$ *is equivalent to the space of all random variables* $f : \Omega \to E$ *such that* $f \in L^p(\Omega; E)$ *is Ray Absolutely Continuous, Stochastically Gâteaux differentiable and the Stochastic Gâteaux derivative* $F : \Omega \to L(H, E)$ *is* $F \in L^p(\Omega; L(H, E))$*.*

Remark 3.11. We know from standard references such as [18] that the map $t \mapsto \tilde{f}_h(\omega + th)$ is continuous as a map from $[0, 1] \to L^0(\Omega)$. The point of proving the stronger absolute continuity is to find a representation of the form

$$\tilde{f}_h(\omega + \varepsilon h) - \tilde{f}_h(\omega) = \int_0^\varepsilon F(\omega + rh)[h] dr,$$

where the object $F(\omega)$ is a candidate for the Malliavin Derivative. Proving Stochastic Gâteaux Differentiability is then verifying that this object is a bounded linear operator and allows one to extend from Gâteaux to Fréchet. Thus a random variable which is Ray Absolutely Continuous but not Stochastic Gâteaux Differentiable has a Malliavin Directional Derivative in all directions, but there is a sequence of elements $h_n \in H$ such that $F(\omega)[h_n] \to \infty$.

By contrast, if one has Stochastic Gâteaux Differentiability but not Ray Absolute Continuity, then one can prove existence of the Malliavin Derivative but which is not in $L^1(\Omega)$ e.g. $\mathbb{E}[\|F(\omega)\|_{L(H,E)}] = \infty$.

Definition 3.12 (Strong Stochastically Gâteaux differentiable)**.** *Let* $p > 1$*. A random variable* $f \in L^p(\Omega; E)$ *is said to be* Strong Stochastically Gâteaux differentiable *if there exists a measurable mapping* $F : \Omega \to L(H, E)$ *such that* $\forall h \in H$

$$\lim_{\varepsilon \to 0} \mathbb{E}\Big[\Big\| \frac{f(\omega + \varepsilon h) - f(\omega)}{\varepsilon} - F(\omega)[h] \Big\|\Big] \to 0. \tag{3.3}$$

Theorem 3.13 ([11])**.** *Let* $p > 1$*. The space* $\mathbb{D}^{1,p}(E)$ *is equivalent to the space of all random variables* $f \in L^p(\Omega; E)$ *that are Strong Stochastically Gâteaux differentiable and have measurable mapping* $F \in L^p(\Omega; L(H, E))$*.*

The merit of [16] is that it allows one to prove Malliavin differentiability by first establishing existence of a Gâteaux derivative and then extending to the full Frechét derivative. The convergence of the Gâteaux derivative in probability is a very weak condition that is much easier to prove than full Malliavin differentiability. [11] extends this result to the stronger Strong Stochastic Gâteaux Differentiability condition and removed the Ray Absolute Continuity condition.

Both of these methods have their merits. While studying different examples of processes with monotone growth, we became interested in the particular example where the drift term has polynomial growth of order q but only finite moments up to $p < q - 2$. In this case, one cannot in general find a dominating function for the error terms coming from the drift of the SDE while trying to prove Stochastic Gâteaux Differentiability. It therefore became necessary to prove only a convergence in probability statement.

Corollary 3.14. *Suppose a measurable map $f:\Omega\to E$ is Stochastically Gâteaux Differentiable and additionally that for $\delta>0$*

$$\sup_{\varepsilon\leq 1}\mathbb{E}\left[\left|\frac{f(\omega+\varepsilon h)-f(\omega)}{\varepsilon}\right|^{1+\delta}\right]<\infty. \tag{3.4}$$

Then f is Malliavin Differentiable (and so f is Ray Absolutely Continuous).

Proof. Condition (3.4) yields the collection of random variables $\big((f(\omega+\varepsilon h)-f(\omega))/\varepsilon\big)_{\varepsilon\leq 1}$ to be uniformly integrable. Stochastic Gâteaux Differentiability means that this collection of random variables converges in probability to a limit. Since $\delta>0$, we conclude that the sequence of random variables converges in mean, or equivalently we have Strong Stochastic Gâteaux differentiability. Theorem 3.13 shows this is equivalent to Malliavin Differentiability and Theorem 3.10 implies we must have Ray Absolute Continuity. □

The convergence conditions on U and V in Assumption 3.1(iii) and (iv) could equivalently been stated in terms of a *Ray Absolute Continuity* and *Stochastic Gâteaux Differentiability* criterion instead of *Strong Stochastic Gâteaux Differentiability*.

Classical results on the Cameron Martin transforms

We recall two useful results from [18]. First we introduce the notation for a Doléans-Dade exponential over $[0,T]$ of some sufficient integrable $\mathbb{R}^m$-valued process, $(M(t))_{t\in[0,T]}$, namely, we define for $t\in[0,T]$ and an m-dimensional Brownian motion W,

$$\mathcal{E}(M)(t)=\exp\Big(\int_0^t M(s)dW(s)-\frac{1}{2}\int_0^t|M(s)|^2ds\Big). \tag{3.5}$$

Proposition 3.15 (The Cameron-Martin Formula – [18])**.** *Let F be an $\mathcal{F}_T$-measurable random variable. For $h\in H$ let $\mathcal{E}(\dot h)(\cdot)$ be the associated Doléans-Dade exponential.*

Then, when both sides are well defined,

$$\mathbb{E}\big[\,F(\omega+h)\,\big]=\mathbb{E}\Big[F\exp\Big(\int_0^T\dot h(s)dW(s)-\frac{1}{2}\int_0^T|\dot h(s)|^2ds\Big)\Big]=\mathbb{E}\big[\,F(\omega)\mathcal{E}(\dot h)(T)\,\big].$$

Moreover, $\forall h\in H$ and $\forall p\geq 1$ that $\mathcal{E}(\dot h)(\cdot)\in\mathcal{S}^p([0,T])$.

Proposition 3.16 (Continuity of the Cameron Martin Transform – [18])**.** *The map $\tau_h:[0,1]\to L^0(\Omega)$ defined by $t\mapsto f(\omega+th)$ is continuous map from a compact interval of the real line to a measurable function with respect to the topology of convergence in probability.*

3.3 Examples

In this section, we discuss some interesting examples which emphasize the scope and sharpness of the assumptions made.

Example 3.17 (Concerning the continuity of $s\mapsto D_sX(\cdot)$)**.** Previous works on Malliavin calculus, see for example [13], treat the solution of this SDE as being continuous in s. While this is true for those examples studied, it is not true in the general case that we study here. We only have that it is square integrability; this example shows that it is not necessary for the derivative to be continuous in s. Take $g\in L^2([0,T])$ be a deterministic discontinuous function (a step function would be adequate) and assume the one dimensional setting. Consider σ of the form

$$\sigma(t,\omega,x)=x+\int_0^t g(s)dW(s)\quad\text{and}\quad b(t,\omega,x)=0.$$

Hence $X(t)$ satisfies $X(t) = 1 + \int_0^t \left[X(s) + \int_0^s g(r)dW(r)\right]dW(s)$. It can be shown that the explicit solution of this equation is

$$X(t) = \exp\Big(W(t) - \frac{t}{2}\Big)\Big[1 - \int_0^t \int_0^r \exp\big(\frac{r}{2} - W(r)\big)g(u)dudr \\ + \int_0^t \int_0^r \exp\big(\frac{r}{2} - W(r)\big)g(u)dW(u)dr\Big].$$

Note that, as expected, X is a continuous process.

The process V, which represents the Malliavin derivative of σ, is

$$V(s,t,\omega) = D_s\sigma(t,\omega,X(t)(\omega)) = g(s)\mathbb{1}_{(0,t)}(s) \\ \Rightarrow \quad \int_s^t V(s,r,\omega)dW(r) = g(s)[W(t) - W(s)].$$

Clearly, the latter map is not continuous in s. The Malliavin derivative of X solves

$$D_sX(t) = X(s) + \int_0^s g(r)dW(r) + g(s)[W(t) - W(s)] + \int_s^t D_sX(r)dW(r).$$

Define $J_s(t) = \exp\Big([W(t) - W(s)] - \frac{t-s}{2}\Big)$. Then the Malliavin derivative has the explicit solution

$$D_sX(t) = J_s(t)\Big[X(s) + \int_0^s g(r)dW(r) + g(s)\Big(\int_s^t J_s(r)^{-1}dW(r) - \int_s^t J_s(r)^{-1}dr\Big)\Big].$$

Since g is assumed not to be continuous, this will also not be continuous in s.

We present a case where the coefficients are not Malliavin differentiable in general but are only differentiable on the set where the solution X takes its values. In other words, Assumption 3.1 is satisfied but Assumption 3.6 is not.

Example 3.18 (Malliavin Differentiable on the right manifold)**.** Let $d = m = 1$ for simplicity. Let $b(t,\omega,x) = -x$ and

$$\sigma(t,\omega(t),x) = \begin{cases} (x-1)^2(x+1)^2 & , x \in [-1,1] \\ \phi(x)\cdot f(\omega(t)) & , |x| > 1 \end{cases},$$

where $\phi \in C^\infty$, $\phi(x) = 0$ for $|x| \le 1$ and $\phi(x) = 1$ for $|x| \ge 2$. The function f is any function $f : \mathbb{R} \to \mathbb{R}$ which is bounded, continuous but not differentiable and ω is the path of the Brownian motion.

An example of such a function f could be

$$f(x) = \begin{cases} W'(x) & , x \in [-1,1] \\ -2 & , |x| > 1 \end{cases},$$

where $W'(x)$ is the Weierstrass function. The Weierstrass function is continuous but not differentiable anywhere and satisfies $W'(-1) = W'(1) = -2$. The latter implies that f is continuous. Hence $f(\omega(t))$ will not be Malliavin differentiable but $\varepsilon \mapsto f(\omega(t) + \varepsilon h(t))$ will be continuous.

The derivative of σ will satisfy

$$\partial_x\sigma(t,\omega,x) = \begin{cases} 4x(x-1)(x+1) & , x \in [-1,1] \\ \phi'(x)\cdot f(\omega(t)) & , 1 < x < 2 \\ 0 & , |x| > 2 \end{cases},$$

so since f is bounded, we conclude that σ is Lipschitz $\forall\omega \in \Omega$ and differentiable.

When the initial conditions determine that the process starts inside the interval $[-1, 1]$, this is a so-called Wright-Fisher process (see [12]) and the solution will remain within the interval $[-1, 1]$ with probability 1. This is important because the non-Malliavin Differentiability only affects the system when the process exits the $[-1, 1]$ interval. The conditions of Assumption 3.1 are satisfied but $\sigma(\cdot, x)$ is not Malliavin differentiable for all $x \in \mathbb{R}^d$.

Remark 3.19 (The square-integrability case)**.** In [11], it is proved that one does not require the Ray Absolute Continuity condition if one can prove a *Strong Stochastic Gâteaux Differentiability* condition, see Theorem 3.13 and Equation (3.3). However, in [8], the authors provide a random variable $Z \in \mathbb{D}^{1,2}$ which is not Strong Stochastic Gâteaux differentiable in the sense that

$$\mathbb{E}\Big[\Big|\frac{Z(\omega+\varepsilon h)-Z(\omega)}{\varepsilon} - D^h Z\Big|^2\Big] \nrightarrow 0, \quad \text{as } \varepsilon \to 0.$$

It is however true that for all values $q \in [1, 2)$

$$\mathbb{E}\Big[\Big|\frac{Z(\omega+\varepsilon h)-Z(\omega)}{\varepsilon} - D^h Z\Big|^q\Big] \to 0, \quad \text{as } \varepsilon \to 0.$$

In our framework, it is necessary to study the square of increments of the process due to the nature of the monotonicity property. Therefore we require that our SDE has finite moment of order p for some $p > 2$. However, in light of the example provided in [8], we believe (but do not show) that that there exists a case where the solution to an SDE of the form (2.1) which has finite moments of order up to $p = 2$ which is Malliavin Differentiable. Stochastic Gâteaux Differentiability would follow as before, but it was unclear to us how one would prove Ray Absolute Continuity of such a process.

Remark 3.20 (The spatial Lipschitz condition for the Malliavin Derivatives of b and σ)**.** In Assumption 3.6 (iv') we assume that Db and $D\sigma$ are Lipschitz in the spacial variable. We chose this condition because it is easy to verify and strong enough to ensure that $\forall x \in \mathbb{R}^d$

$$\mathbb{E}\Big[\Big(\int_0^T \Big(\int_0^t |D_s b(t,\omega,X(t))|^2 ds\Big)^{\frac{1}{2}} dt\Big)^p\Big] < \infty,$$
$$\mathbb{E}\Big[\Big(\int_0^T \int_0^t |D_s \sigma(t,\omega,X(t))|^2 ds dt\Big)^{\frac{p}{2}}\Big] < \infty.$$

However, this condition is by no means necessary. One could consider the case where Db is locally Lipschitz in space and satisfies a linear growth condition and equivalently prove Theorem 3.7. However, the proof is more involved as it involves a careful interplay using Hölder's inequality between the maximal integrability of X, Db, $D\sigma$ and several other stochastic terms.

3.4 Proofs of the 1st main result - Theorem 3.2

In what follows, the choice of θ (the initial condition in (2.1)) does not affect the Malliavin derivative because θ is $\mathcal{F}_0$-measurable. If Y is $\mathcal{F}_t$-measurable then $D_s Y = 0$ for any $t < s$, see [13, Corollary 1.2.1].

Existence and uniqueness of the Malliavin derivative $D_s X(t)$

We start by establishing that (3.1) has a unique solution where X solves (2.1). At this point, nothing is said about the solution of (3.1) being the Malliavin derivative to X solution of (2.1), showing it is the subsequent step.

Theorem 3.21. *Let $p>2$. For $(s,t)\in[0,T]^2$, let X be the solution to the SDE* (2.1) *under Assumption 3.1. Let $(M_s(t))$ be defined by the matrix of $L^2([0,T])$-valued SDEs*

$$\begin{aligned} M_s(t)(\omega) =&\sigma(s,\omega,X(s)(\omega)) + \int_s^t U(s,r,\omega)dr + \int_s^t V(s,r,\omega)dW(r) \\ &+ \int_s^t \nabla_x b(r,\omega,X(r)(\omega))M_s(r)(\omega)dr + \int_s^t \nabla_x \sigma(r,\omega,X(r)(\omega))M_s(r)(\omega)dW(r), \end{aligned} \tag{3.6}$$

for $s<t$ and $M_s(t)=0$ for $s>t$.

Then a unique solution exists in $\mathcal{S}^p([0,T];L^2([0,T]))$ for (3.6) *and the process M has finite p^{th} moment, namely*

$$\mathbb{E}\Big[\Big(\sup_{t\in[0,T]}\int_0^T |M_s(t)|^2 ds\Big)^{\frac{p}{2}}\Big]<\infty.$$

Observe that Equation (3.6) is linear in M, so the sharpness of the integrability is determined by the integrability of U, V and σ (given the assumed behavior of $\nabla_x b$ and $\nabla_x\sigma$). In the trivial case where $U=V=0$ and $\sigma=1$ then M has finite moments of all orders.

Proof of Theorem 3.21. For brevity, $t\in[0,T]$ and we omit the explicit ω dependency throughout.

Equation (3.6) is an infinite dimensional SDE. We see this when we think of the Malliavin Derivative as being an $L^2([0,T])$ valued stochastic process. Therefore, we need to extend results from Section 2 to infinite dimensional spaces. Let e_n be an orthonormal basis of the space $L^2([0,T];\mathbb{R}^m)$. This is a separable Hilbert space, so without loss of generality we can say the orthonormal basis is countably infinite. Let V_n be the linear span of the set $\{e_1,...,e_n\}$. Let $P_n: L^2([0,T];\mathbb{R}^m)\to V_n$ be the canonical projection operators

$$P_n[f](t)=\sum_{k=1}^n \langle f,e_k\rangle_{L^2([0,T];\mathbb{R}^m)}e_k(t).$$

Then it is clear that $\lim_{n\to\infty}\|P_n[f]-f\|_{L^2([0,T];\mathbb{R}^m)}=0$. For $k\in\mathbb{N}$, consider the sequence of 1-dimensional Linear Stochastic Differential Equations

$$\begin{aligned} M_k(t)=&\int_0^t \sigma(u,X(u))e_k(u)du+\int_0^t\Big(\int_0^r U(u,r)e_k(u)du\Big)dr\\ &+\int_0^t\Big(\int_0^r V(u,r)e_k(u)du\Big)dW(r)+\int_0^t \nabla_x b(r,X(r))M_k(r)dr\\ &+\int_0^t \nabla_x\sigma(r,X(r))M_k(r)dW(r). \end{aligned}$$

These equations are of the same form as (2.3), hence a unique solution exists for each k by Theorem 2.5. Also, observe that the fundamental matrix Ψ will be the same for each choice of $k\in\mathbb{N}$. Ψ will have the explicit solution

$$\begin{aligned}\Psi(t)=&\exp\Big(\int_0^t \nabla b(r,X(r))dr-\frac{1}{2}\int_0^t\Big\langle\nabla\sigma\big(r,X(r)\big),\nabla\sigma\big(r,X(r)\big)\Big\rangle_{\mathbb{R}^m}dr\\ &+\int_0^t\nabla\sigma(r,X(r))dW(r)\Big)\end{aligned}$$

and M_k has explicit solution

$$
\begin{aligned}
M_k(t) =&\Psi(t)\Bigg(\int_0^t \sigma\big(u,X(u)\big)e_k(u)du + \int_0^t \Psi(r)^{-1}\Big[\int_0^r U\big(u,r\big)e_k(u)du\\
&-\Big\langle \nabla\sigma\big(r,X(r)\big), \int_0^r V\big(u,r\big)e_k(u)du\Big\rangle_{\mathbb{R}^m}\Big]dr + \int_0^t \Psi(r)^{-1}\int_0^r V\big(u,r\big)e_k(u)dudW(r)\Bigg).
\end{aligned}
$$

Next define for $0\leq s,t\leq T$, $n\in\mathbb{N}$ the process $M_{(n),s}(t)=\sum_{k=1}^n M_k(t)\otimes e_k(s)\mathbb{1}_{[0,t)}(s)$. This process makes sense as the projection space is finite dimensional so we can rewrite it in a finite dimensional vector form. The solution exists in the space $\mathcal{S}^p(L^2([0,T];\mathbb{R}^{d\times m}))$ and has the explicit solution

$$
\begin{aligned}
M_{(n),s}(t) =&\sum_{k=1}^n \Psi(t)\Bigg(\int_0^t \sigma\big(u,X(u)\big)e_k(u)du + \int_0^t \Psi(r)^{-1}\Big[\int_0^r U\big(u,r\big)e_k(u)du\\
&-\Big\langle \nabla\sigma\big(r,X(r)\big), \int_0^r V\big(u,r\big)e_k(u)du\Big\rangle_{\mathbb{R}^m}\Big]dr\\
&+\int_0^t \Psi(r)^{-1}\Big[\int_0^r V\big(u,r\big)e_k(u)du\Big]dW(r)\Bigg)\otimes e_k(s)\mathbb{1}_{[0,t)}(s),\\
=&\Psi(t)\Bigg(P_n\Big[\sigma(\cdot,X(\cdot))\Big](s) + \int_s^t \Psi(r)^{-1}\Big(P_n\Big[U(\cdot,r)\Big](s)\\
&-\Big\langle \nabla\sigma\big(r,X(r)\big), P_n\Big[V\big(\cdot,r\big)\Big](s)\Big\rangle_{\mathbb{R}^m}\Big)dr + \int_s^t \Psi(r)^{-1}P_n\Big[V\big(\cdot,r\big)\Big](s)dW(r)\Bigg).
\end{aligned}
$$

This process satisfies the SDE

$$
\begin{aligned}
M_{(n),s}(t) =&P_n\Big[\sigma(\cdot,X(\cdot))\Big](s) + \int_s^t P_n\Big[U(\cdot,r)\Big](s)dr + \int_s^t P_n\Big[V(\cdot,r)\Big](s)dW(r)\\
&+\int_s^t \nabla_x b(r,X(r))M_{(n),s}(r)dr + \int_s^t \nabla_x\sigma(r,X(r))M_{(n),s}(r)dW(r).
\end{aligned}
$$

For $a^{(i,j)}\in L^2([0,T])$ and for $A(u)=(a^{(i,j)}(u))_{i\in\{1,\dots,d\},j\in\{1,\dots,m\}}$, define the norm

$$
\|A.\| = \Big(\sum_{i=1}^d\sum_{j=1}^m \int_0^T |a^{(i,j)}(u)|^2du\Big)^{1/2} = \Big(\int_0^T |A(u)|^2du\Big)^{1/2}. \tag{3.7}
$$

By Itô's formula, we have

$$
\begin{aligned}
&\Big|M_{(n),s}(t)-M_{(m),s}(t)\Big|^2 = \sum_{i=1}^d\sum_{k=1}^m \Big|M_{(n),s}^{(i,k)}(t)-M_{(m),s}^{(i,k)}(t)\Big|^2,\\
&=\sum_{i,k}\Big|(P_n-P_m)\Big[\sigma(\cdot,X(\cdot))\Big]^{(i,k)}(s)\Big|^2\\
&+2\sum_{i,k}\int_s^t \Big(M_{(n),s}^{(i,k)}(r)-M_{(m),s}^{(i,k)}(r)\Big)\cdot(P_n-P_m)\Big[U(\cdot,r)\Big]^{(i,k)}(s)dr\\
&+2\sum_{i,j,k}\int_s^t \Big(M_{(n),s}^{(i,k)}(r)-M_{(m),s}^{(i,k)}(r)\Big)\cdot(P_n-P_m)\Big[V(\cdot,r)\Big]^{(i,j,k)}(s)dW^{(j)}(r)\\
&+2\sum_{i,k}\int_s^t \Big(M_{(n),s}^{(i,k)}(r)-M_{(m),s}^{(i,k)}(r)\Big)\cdot\Big\langle \nabla_x b^{(i)}(r,X(r)), M_{(n),s}^{(\cdot,k)}(r)-M_{(m),s}^{(\cdot,k)}(r)\Big\rangle dr
\end{aligned}
$$

$$+2\sum_{i,j,k}\int_s^t \Big(M^{(i,k)}_{(n),s}(r)-M^{(i,k)}_{(m),s}(r)\Big)\cdot\Big\langle \nabla_x\sigma^{(i,j)}(r,X(r)),M^{(\cdot,k)}_{(n),s}(r)-M^{(\cdot,k)}_{(m),s}(r)\Big\rangle dW^{(j)}(r)$$

$$+\sum_{i,j,k}\int_s^t \Big|(P_n-P_m)\Big[V(\cdot,r)\Big]^{(i,j,k)}(s)+\Big\langle \nabla_x\sigma^{(i,j)}(r,X(r)),M^{(\cdot,k)}_{(n),s}(r)-M^{(\cdot,k)}_{(m),s}(r)\Big|^2 dr.$$

Denote $N_s(t)=M_{(n),s}(t)-M_{(m),s}(t)$ and $(P_n-P_m)=Q$ for brevity. Integrating over s and since every term is positive, we can change the order of integration to obtain

$$\begin{aligned}\int_0^t \Big|N^{(i,k)}_s(t)\Big|^2 ds &= \int_0^t \Big|Q\big[\sigma(\cdot,\omega,X(\cdot))\big]^{(i,k)}(s)\Big|^2 ds\\
&+2\int_0^t\int_0^r N^{(i,k)}_s(r)\cdot\Big[Q\big[U(\cdot,r)\big]^{(i,k)}(s)+\Big\langle \nabla_x b^{(i)}(r,X(r)),N^{(\cdot,k)}_s(r)\Big\rangle\Big]dsdr\\
&+2\sum_j\int_0^t\int_0^r N^{(i,k)}_s(r)\cdot\Big[Q\big[V(\cdot,r)\big]^{(i,j,k)}(s)+\Big\langle \nabla_x \sigma^{(i,j)}(r,X(r)),N^{(\cdot,k)}_s(r)\Big\rangle\Big]dsdW^{(j)}(r)\\
&+\sum_j\int_0^t\int_0^r \Big|Q\big[V(\cdot,r)\big]^{(i,j,k)}(s)+\Big\langle \nabla_x \sigma^{(i,j)}(r,X(r)),N^{(\cdot,k)}_s(r)\Big\rangle\Big|^2 dsdr.\end{aligned}$$

Next, we use Itô's formula for the function $g(x)=\big(\sum_i x^{(i)}\big)^{p/2}$ to get

$$\begin{aligned}\|N_\cdot(t)\|^p = \Big(\sum_{i,k}\int_0^t |N^{(i,k)}_s(t)|^2 ds\Big)^{\frac{p}{2}} &= \Big(\int_0^t \Big|Q\big[\sigma(\cdot,\omega,X(\cdot))\big](s)\Big|^2 ds\Big)^{\frac{p}{2}}\\
&+p\int_0^t \|N_\cdot(r)\|^{p-2}\Big(\sum_{i,k}\int_0^r N^{(i,k)}_s(r)\Big[Q\big[U(\cdot,r)\big]^{(i,k)}(s)\\
&\qquad+\Big\langle \nabla_x b^{(i)}\big(r,X(r)\big),N^{(\cdot,k)}_s(r)\Big\rangle ds\Big)dr \qquad (3.8)\\
+\frac{p}{2}\int_0^t \|N_\cdot(r)\|^{p-2}\sum_{i,j,k}\int_0^r \Big|Q\big[V(\cdot,r)\big]^{i,j,k}(s)&\\
&\qquad+\Big\langle \nabla_x \sigma^{(i,j)}(r,X(r)),N^{(\cdot,k)}_s(r)\Big\rangle\Big|^2 dsdr \qquad (3.9)\\
+p\int_0^t \|N_\cdot(r)\|^{p-2}\sum_{i,j,k}\int_0^r N^{(i,k)}_s(r)\Big(Q\big[V(\cdot,r)\big]^{(i,j,k)}(s)&\\
&\qquad+\Big\langle \nabla_x \sigma^{(i,j)}\big(r,X(r)\big),N^{(\cdot,k)}_s(r)\Big)dsdW(r) \qquad (3.10)\\
+p(p-2)\int_0^t \|N_\cdot(r)\|^{p-4}\sum_{i,j,k}\Big(\int_0^r N^{(i,k)}_s(r)\Big(Q\big[V(\cdot,r)\big]^{(i,j,k)}(s)&\\
&\qquad+\Big\langle \nabla_x \sigma^{(i,j)}\big(r,X(r)\big),N^{(\cdot,k)}_s(r)\Big)ds\Big)^2 dr. \qquad (3.11)\end{aligned}$$

We take a supremum over $t\in[0,T]$ then expectations to show that $\mathbb{E}[\|N\|^2_\infty]$ can be made arbitrarily small for $n,m\in\mathbb{N}$ large enough. Let $a\in\mathbb{N}$ be an integer which we will choose later.

Firstly,

$$(3.8)\le p\mathbb{E}\Big[\int_0^T \|N_\cdot(r)\|^{p-2}\Big(\sum_{i,k}\int_0^r N^{(i,k)}_s(r)Q\big[U(\cdot,r)\big]^{(i,k)}(s)ds\Big)dr\Big] \qquad (3.12)$$

$$+p\mathbb{E}\Big[\int_0^T \|N_\cdot(r)\|^{p-2}\Big(\sum_{i,k}\int_0^r N^{(i,k)}_s(r)\Big\langle \nabla_x b^{(i)}\big(r,X(r)\big),N^{(\cdot,k)}_s(r)\Big\rangle ds\Big)dr\Big]. \qquad (3.13)$$

Now we deal with (3.12) using Hölder inequality, the norm (3.7), then dominate via the supremum norm and move the term outside the integral to merge it with the outer integrand term

$$\begin{aligned}(3.12) \leq & p\mathbb{E}\Big[\|N_\cdot\|_\infty^{p-1}\int_0^T\Big(\sum_{i,k}\int_0^r|Q[U(\cdot,r)]^{(i,k)}(s)|^2ds\Big)^{\frac{1}{2}}dr\Big]\\ \leq & \frac{\mathbb{E}\Big[\|N_\cdot\|_\infty^p\Big]}{a}+[a(p-1)]^{p-1}\mathbb{E}\Big[\Big(\int_0^T\|Q[U(\cdot,r)](s)\|_2ds\Big)^p\Big],\end{aligned}$$

and

$$(3.13) \leq pL\int_0^T\mathbb{E}\Big[\|N_\cdot\|_{\infty,r}^p\Big]dr.$$

using the Monotonicity property of b. Secondly,

$$\begin{aligned}(3.9) \leq & p\mathbb{E}\Big[\int_0^T\|N_\cdot(r)\|^{p-2}\Big(\sum_{i,j,k}\int_0^r|Q[V(\cdot,r)]^{(i,j,k)}|^2ds\Big)dr\Big]\\ & +p\mathbb{E}\Big[\int_0^T\|N_\cdot(r)\|^{p-2}\Big(\sum_{i,j,k}\int_0^r\Big\langle\nabla_x\sigma^{(i,j)}(r,X(r)),N_s^{(\cdot,k)}(r)\Big\rangle^2ds\Big)dr\Big]\\ \leq & \frac{\mathbb{E}\Big[\|N_\cdot\|_\infty^p\Big]}{a}+2[a(p-2)]^{\frac{p-2}{2}}\mathbb{E}\Big[\Big(\int_0^T\|Q[V(\cdot,r)](\cdot)\|_2^2dr\Big)^{\frac{p}{2}}\Big]+pL^2\int_0^t\mathbb{E}\Big[\|N_\cdot\|_{\infty,r}^p\Big]dr,\end{aligned}$$

using the boundedness of $\nabla\sigma$. Thirdly, using the Burkholder-Davis-Gundy Inequality

$$\begin{aligned}(3.10) \leq & pC_1\mathbb{E}\Big[\Big(\int_0^T\|N_\cdot(r)\|^{2p-4}\sum_j\Big(\sum_{i,k}\int_0^r N_s^{(i,k)}(r)\Big[Q[V(\cdot,r)]^{(i,j,k)}(s)\\ & +\Big\langle\nabla_x\sigma^{(i,j)}(r,X(r)),N_s^{(\cdot,k)}(r)\Big\rangle\Big]ds\Big)^2dr\Big)^{\frac{1}{2}}\Big]\\ \leq & \sqrt{2}pC_1\mathbb{E}\Big[\|N_\cdot\|_\infty^{p-2}\Big(\int_0^T\Big[\sum_{i,j,k}\int_0^r|N_s^{(i,k)}(r)|\cdot|Q[V(\cdot,r)]^{(i,j,k)}(s)|ds\Big]^2dr\Big)^{\frac{1}{2}}\Big] \quad (3.14)\\ & +\sqrt{2}pC_1\mathbb{E}\Big[\|N_\cdot\|_\infty^{p-2}\Big(\int_0^T\Big[\sum_{i,j,k}\int_0^r|N_s^{(i,k)}(r)|\\ & \qquad\cdot\Big|\Big\langle\nabla_x\sigma^{(i,j)}(r,X(r)),N_s^{(\cdot,k)}(r)\Big\rangle\Big|ds\Big]^2dr\Big)^{\frac{1}{2}}\Big]. \quad (3.15)\end{aligned}$$

As before, we have

$$\begin{aligned}(3.14) \leq & \sqrt{2}pC_1\mathbb{E}\Big[\|N_\cdot\|_\infty^{p-1}\Big(\int_0^T\|Q[V(\cdot,r)](\cdot)\|_2^2dr\Big)^{\frac{1}{2}}\Big]\\ \leq & \frac{\mathbb{E}\Big[\|N_\cdot\|_\infty^p\Big]}{a}+(\sqrt{2}C_1)^p[a(p-1)]^{p-1}\mathbb{E}\Big[\Big(\int_0^T\|Q[V(\cdot,r)](\cdot)\|_2^2dr\Big)^{\frac{p}{2}}\Big],\end{aligned}$$

and put together

$$\begin{aligned}(3.15) \leq & \sqrt{2}pC_1L\mathbb{E}\Big[\Big(\int_0^T\|N_\cdot(r)\|^{2p}dr\Big)^{\frac{1}{2}}\Big]\leq\sqrt{2}pC_1L\mathbb{E}\Big[\|N_\cdot\|_\infty^{\frac{p}{2}}\Big(\int_0^T\|N_\cdot(r)\|^pdr\Big)^{\frac{1}{2}}\Big]\\ \leq & \frac{\mathbb{E}\Big[\|N_\cdot\|_\infty^p\Big]}{a}+\frac{(pC_1L)^2a}{2}\int_0^T\mathbb{E}\Big[\|N_\cdot\|_{\infty,r}^p\Big]dr.\end{aligned}$$

Finally, for

$$(3.11) \leq 2p(p-2)\mathbb{E}\Big[\int_0^T \|N.(r)\|^{p-4}\Big(\sum_{i,j,k}\int_0^r |N_s^{(i,k)}(r)|\cdot|Q[V(\cdot,r)]^{(i,j,k)}(s)|ds\Big)^2 dr\Big] \tag{3.16}$$

$$+ 2p(p-2)\mathbb{E}\Big[\int_0^T \|N.(r)\|^{p-4}\Big(\sum_{i,j,k}\int_0^r |N_s^{(i,k)}(r)| \cdot\Big|\Big\langle \nabla_x\sigma^{(i,j)}(r,X(r)), N_s^{(\cdot,k)}(r)\Big\rangle\Big|ds\Big)^2 dr\Big]. \tag{3.17}$$

Repeating the same ideas as before, we get

$$\begin{aligned}(3.16) \leq& 2p(p-2)\mathbb{E}\Big[\|N.\|_\infty^{p-2}\cdot\int_0^T \|Q[V(\cdot,r)](\cdot)\|_2^2 dr\Big]\\ \leq& \frac{\mathbb{E}\Big[\|N.\|_\infty^p\Big]}{a} + 2^{\frac{p+2}{2}}\cdot a^{\frac{p-2}{2}}\cdot(p-2)^{p-1}\mathbb{E}\Big[\Big(\int_0^T \|Q[V(\cdot,r)](\cdot)\|_2^2 dr\Big)^{\frac{p}{2}}\Big].\end{aligned}$$

and $(3.17) \leq 2p(p-2)L^2\mathbb{E}\big[\int_0^T \|N.\|_{\infty,r}^p dr\big]$. Therefore, choosing $a=6$ we conclude

$$\begin{aligned}\frac{\mathbb{E}\Big[\|N.\|_\infty^p\Big]}{6} \leq& \Bigg(\mathbb{E}\Big[\|Q[\sigma(\cdot,X(\cdot))](\cdot)\|_2^p\Big] + \widetilde{C}_1\mathbb{E}\Big[\Big(\int_0^T \|Q[U(\cdot,r)](\cdot)\|_2 dr\Big)^p\Big]\\ &+ \widetilde{C}_2\mathbb{E}\Big[\Big(\int_0^T \|Q[V(\cdot,r)](\cdot)\|_2^2 dr\Big)^{\frac{p}{2}}\Big]\Bigg) + \widetilde{C}_3\int_0^T \mathbb{E}\Big[\|N.\|_{\infty,r}^p\Big]dr.\end{aligned}$$

By applying Grönwall's inequality we conclude that

$$\begin{aligned}\mathbb{E}\Big[\sup_{t\in[0,T]}\|M_{(n),\cdot}-M_{(m),\cdot}\|^p\Big] \lesssim& \Bigg(\mathbb{E}\Big[\|(P_n-P_m)\big[\sigma(\cdot,X(\cdot))\big]\|_2^p\Big]\\ &+\mathbb{E}\Big[\Big(\int_0^T \|(P_n-P_m)[U(\cdot,r)]\|_2 dr\Big)^p\Big] + \mathbb{E}\Big[\Big(\int_0^T \|(P_n-P_m)[V(\cdot,r)]\|_2^2 dr\Big)^{\frac{p}{2}}\Big]\Bigg).\end{aligned}$$

Given that by assumption we already have

$$\mathbb{E}\Big[\|\sigma(\cdot,X(\cdot))\|_2^p\Big],\quad \mathbb{E}\Big[\Big(\int_0^T \|U(\cdot,r)\|_2 dr\Big)^p\Big],\quad \mathbb{E}\Big[\Big(\int_0^T \|V(\cdot,r)\|_2^2 dr\Big)^{\frac{p}{2}}\Big] < \infty,$$

we are able to apply the Dominated Convergence Theorem to swap the order of limits and integrals. Taking a limit as m,n go to infinity lets us conclude that the sequence $M_{(n)}$ is Cauchy in $\mathcal{S}^p(L^2([0,T];\mathbb{R}^{d\times m}))$. This is a Banach space, so a limit must exist which we denote by M',

$$\begin{aligned}M'_s(t) = \lim_{n\to\infty}\Psi(t)\Bigg(&P_n\Big[\sigma(\cdot,X(\cdot))\Big](s) + \int_s^t \Psi(r)^{-1}\Big(P_n\Big[U(\cdot,r)\Big](s)\\ &- \Big\langle\nabla\sigma\big(r,X(r)\big), P_n\Big[V\big(\cdot,r\big)\Big](s)\Big\rangle_{\mathbb{R}^m}\Big)dr + \int_s^t \Psi(r)^{-1}P_n\Big[V\big(\cdot,r\big)\Big](s)dW(r)\Bigg).\end{aligned}$$

Now let $g\in L^2([0,T];\mathbb{R}^m)$ be chosen arbitrarily. Then we define $M^{g\prime}(\cdot)$ as

$$
\begin{aligned}
M^{g\prime}(t) =& \int_0^t M_s'(t) g(s) ds \\
=& \Psi(t)\Bigg(\int_0^t \sigma(s, X(s)) g(s) ds + \int_s^t \Psi(r)^{-1} \Big(\int_0^t U(s,r) g(s) ds \\
&- \Big\langle \nabla\sigma\big(r, X(r)\big), \int_0^t V\big(s,r\big) g(s) ds \Big\rangle_{\mathbb{R}^m} \Big) dr + \int_s^t \Psi(r)^{-1} \int_0^t V\big(s,r\big) g(s) ds dW(r) \Bigg).
\end{aligned}
$$

In order to move the limit inside the different integrals, we use the Dominated Convergence Theorem again.

Given an explicit solution, we know $M^{g\prime}$ will satisfy the SDE

$$
\begin{aligned}
M^{g\prime}(t) =& \int_0^t \sigma(s, X(s)) g(s) ds + \int_0^t \Big(\int_0^r U(s,r) g(s) ds \Big) dr + \int_0^t \Big(\int_0^r V(s,r) g(s) ds \Big) dW(r) \\
&+ \int_0^t \nabla_x b(r, X(r)) M^{g\prime}(r) dr + \int_0^t \nabla_x \sigma(r, X(r)) M^{g\prime}(r) dW(r).
\end{aligned}
$$

Therefore by a duality argument

$$
\begin{aligned}
M_s'(t) =& \sigma(s, X(s)) + \int_0^t U(s,r) dr + \int_0^t V(s,r) dW(r) \\
&+ \int_0^t \nabla_x b(r, X(r)) M_s'(r) dr + \int_0^t \nabla_x \sigma(r, X(r)) M_s'(r) dW(r),
\end{aligned}
$$

which is the same SDE as (3.6).

Next we prove uniqueness. Suppose that there are two solutions to the SDE (3.6), M and M'. Denote $M - M' = \tilde{N}$. Then $\tilde{N}$ will satisfy the linear SDE

$$
d\tilde{N}_s(t) = \nabla_x b(t, X(t)) \tilde{N}_s(t) dt + \nabla_x \sigma(t, X(t)) \tilde{N}_s(t) dW(t), \quad \tilde{N}_s(s) = 0.
$$

Let $g \in L^2([0,T]; \mathbb{R}^m)$ be chosen arbitrarily. Define $\tilde{N}^g(t) = \int_0^t \tilde{N}_s(t) g(s) ds$. Clearly, this linear SDE will almost surely be equal to 0 independently of the choice of g. Hence $\tilde{N}$ must also be equal to 0. So $M = M'$ and we have proved uniqueness. ☐

Ray absolute continuity of X

We show that the expectation of $\|\big(X(\cdot)(\omega + \varepsilon h) - X(\cdot)(\omega)\big)/\varepsilon\|_\infty^2$ has a bound uniform in ε. This relies on having finite p^{th} moments of the random variable $\|X\|_\infty$ for $p > 2$.

The case $p = 2$ is problematic. It is not the case that $Z \in \mathbb{D}^{1,2}$ implies that $\big(Z(\omega + \varepsilon h) - Z(\omega)\big)/\varepsilon$ converges in mean square as $\varepsilon \searrow 0$, see Remark 3.19 and [8] for in-depth discussion. If we were dealing with the sharp case where the solution of the SDE exists in $\mathcal{S}^2$, it would be unreasonable to expect the Malliavin Derivatives of b and σ to satisfy Assumption 3.1(iv), which is necessary for the following Proposition. The power p must be greater that 2, as opposed to 1, because the monotonicity condition lends itself to studying the moments of the SDE for moments of greater than or equal to 2 but is a hindrance for the moments of order less than 2 (computations may involve local times).

Proposition 3.22. *Let X be solution to the SDE* (2.1) *under Assumption 3.1. We have*

$$
\mathbb{E}\Big[\Big\| \frac{X(\cdot)(\omega + \varepsilon h) - X(\cdot)(\omega)}{\varepsilon} \Big\|_\infty^2 \Big] = O(1) \quad \text{as } \varepsilon \to 0. \tag{3.18}
$$

After we have proved Stochastic Gâteaux Differentiability (see Theorem 3.23), Corollary 3.14 and Equation (3.18) will imply Ray Absolute Continuity.

Proof. Let $t \in [0,T]$. Using Assumption 3.1, we have

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\int_0^T \Big|\frac{b(t,\omega+\varepsilon h, X(t)(\omega)) - b(t,\omega,X(t)(\omega))}{\varepsilon}\Big|dt\Big)^2\Big] \\
&\quad \le 2\mathbb{E}\Bigg[\|h\|_2^2\Big(\int_0^T \Big(\int_0^t |U(s,t,\omega)|^2 ds\Big)^{\frac{1}{2}} dt\Big)^2 \\
&\qquad + \Big(\int_0^T \Big|\frac{b(t,\omega+\varepsilon h, X(t)) - b(t,\omega,X(t))}{\varepsilon} - \int_0^t U(s,t,\omega)\dot{h}(s)ds\Big|dt\Big)^2\Bigg] \le O(1),
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(t,\omega+\varepsilon h, X(t)(\omega)) - \sigma(t,\omega,X(t)(\omega))}{\varepsilon}\Big|^2 ds\Big] \\
&\quad \le 2\mathbb{E}\Bigg[\|h\|_2^2 \int_0^T \int_0^t |V(s,t,\omega)|^2 ds dt \\
&\qquad + \int_0^T \Big|\frac{\sigma(t,\omega+\varepsilon h, X(t)) - \sigma(t,\omega,X(t))}{\varepsilon} - \int_0^t V(s,t,\omega)\dot{h}(s)ds\Big|^2 dt\Bigg] \le O(1).
\end{aligned}$$

For notational compactness let us introduce $P_\varepsilon(t)(\omega) = \big(X(t)(\omega+\varepsilon h) - X(t)(\omega)\big)/\varepsilon$. We have

$$\begin{aligned}
P_\varepsilon(t)(\omega) = &\int_0^t \sigma(s,\omega,X(s)(\omega))\dot{h}(s)ds \\
&+ \int_0^t \Big(\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega,X(s)(\omega))\Big)\dot{h}(s)ds \\
&+ \frac{1}{\varepsilon}\int_0^t \Big(b(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - b(s,\omega,X(s)(\omega))\Big)ds \\
&+ \frac{1}{\varepsilon}\int_0^t \Big(\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega,X(s)(\omega))\Big)dW(s).
\end{aligned}$$

Using Itô's formula for $f(x) = x^2$ we have

$$\Big|P_\varepsilon(t)(\omega)\Big|^2 = 2\int_0^t \Big\langle P_\varepsilon(s)(\omega), \sigma(s,\omega,X(s)(\omega))\dot{h}(s)\Big\rangle ds \tag{3.19}$$

$$+ 2\int_0^t \Big\langle P_\varepsilon(s)(\omega), \Big(\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega+\varepsilon h, X(s)(\omega))\Big)\dot{h}(s)\Big\rangle ds \tag{3.20}$$

$$+ 2\int_0^t \Big\langle P_\varepsilon(s)(\omega), \Big(\sigma(s,\omega+\varepsilon h, X(s)(\omega)) - \sigma(s,\omega,X(s)(\omega))\Big)\dot{h}(s)\Big\rangle ds \tag{3.21}$$

$$+ 2\int_0^t \Big\langle P_\varepsilon(s)(\omega), \frac{b(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - b(s,\omega+\varepsilon h, X(s)(\omega))}{\varepsilon}\Big\rangle ds \tag{3.22}$$

$$+ 2\int_0^t \Big\langle P_\varepsilon(s)(\omega), \frac{b(s,\omega+\varepsilon h, X(s)(\omega)) - b(s,\omega,X(s)(\omega))}{\varepsilon}\Big\rangle ds \tag{3.23}$$

$$+ 2\int_0^t \Big\langle P_\varepsilon(s)(\omega), \frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega+\varepsilon h, X(s)(\omega))}{\varepsilon}dW(s)\Big\rangle \tag{3.24}$$

$$+ 2\int_0^t \Big\langle P_\varepsilon(s)(\omega), \frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega)) - \sigma(s,\omega,X(s)(\omega))}{\varepsilon}dW(s)\Big\rangle \tag{3.25}$$

$$+ \int_0^t \Big|\frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega,X(s)(\omega))}{\varepsilon}\Big|^2 ds. \tag{3.26}$$

We take a supremum over t then expectations. Let n be an integer that we will choose later. By using a combination of Young's Inequality, Cauchy-Schwartz Inequality, Burkholder-Davis-Gundy Inequality and the continuity properties from Assumption 3.1 we find the following upper bounds:

$$\begin{aligned}\text{For (3.19)} \Rightarrow &\mathbb{E}\Big[2\int_0^T \Big|\Big\langle P_\varepsilon(s)(\omega), \sigma(s,\omega,X(s)(\omega))\dot{h}(s)\Big\rangle\Big|ds\Big]\\ \leq&\frac{\mathbb{E}[\|P_\varepsilon\|_\infty^2]}{n} + n\|\dot{h}\|_2^2\mathbb{E}\Big[\int_0^T \Big|\sigma(s,\omega,X(s)(\omega))\Big|^2 ds\Big]\\ \leq&\frac{\mathbb{E}[\|P_\varepsilon\|_\infty^2]}{n} + 2n\|\dot{h}\|_2^2\Big(L^2\mathbb{E}\Big[\|X\|_\infty^2\Big] + \mathbb{E}\Big[\int_0^T \Big|\sigma(s,\omega,0)\Big|^2 ds\Big]\Big),\end{aligned}$$

For (3.20)

$$\begin{aligned}&\Rightarrow \mathbb{E}\Big[2\int_0^T \Big|\Big\langle P_\varepsilon(s)(\omega), \Big(\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega+\varepsilon h, X(s)(\omega))\Big)\dot{h}(s)\Big\rangle\Big|ds\Big]\\ &\leq 2L\varepsilon\int_0^T \mathbb{E}\Big[\|P_\varepsilon\|_{\infty,s}^2\Big]\cdot|\dot{h}(s)|ds,\end{aligned}$$

$$\begin{aligned}\text{For (3.21)} \Rightarrow &\mathbb{E}\Big[2\int_0^T \Big|\Big\langle P_\varepsilon(s)(\omega), \Big(\sigma(s,\omega+\varepsilon h, X(s)(\omega)) - \sigma(s,\omega,X(s)(\omega))\Big)\dot{h}(s)\Big\rangle\Big|ds\Big]\\ \leq&\frac{\mathbb{E}[\|P_\varepsilon\|_\infty^2]}{n} + n\|\varepsilon\dot{h}\|_2^2\mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega)) - \sigma(s,\omega,X(s)(\omega))}{\varepsilon}\Big|^2 ds\Big],\end{aligned}$$

$$\begin{aligned}\text{For (3.22)} \Rightarrow &\mathbb{E}\Big[2\int_0^T \Big|\Big\langle P_\varepsilon(s)(\omega), \frac{b(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - b(s,\omega+\varepsilon h, X(s)(\omega))}{\varepsilon}\Big\rangle\Big|ds\Big]\\ \leq&2L\int_0^T \mathbb{E}\Big[\|P_\varepsilon\|_{\infty,s}^2\Big]ds,\end{aligned}$$

$$\begin{aligned}\text{For (3.23)} \Rightarrow &\mathbb{E}\Big[2\int_0^T \Big|\Big\langle P_\varepsilon(s)(\omega), \frac{b(s,\omega+\varepsilon h, X(s)(\omega)) - b(s,\omega,X(s)(\omega))}{\varepsilon}\Big\rangle\Big|ds\Big]\\ \leq&\frac{\mathbb{E}[\|P_\varepsilon\|_\infty^2]}{n} + n\mathbb{E}\Big[\Big(\int_0^T \Big|\frac{b(s,\omega+\varepsilon h, X(s)(\omega)) - b(s,\omega,X(s)(\omega))}{\varepsilon}\Big|ds\Big)^2\Big],\end{aligned}$$

For (3.24)

$$\begin{aligned}&\Rightarrow \mathbb{E}\Big[\sup_{t\in[0,T]} 2\int_0^t \Big\langle P_\varepsilon(s), \frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega+\varepsilon h, X(s)(\omega))}{\varepsilon}dW(s)\Big\rangle\Big]\\ &\leq 2C_1\mathbb{E}\Big[\|P_\varepsilon\|_\infty\Big(\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega+\varepsilon h, X(s)(\omega))}{\varepsilon}\Big|^2 ds\Big)^{1/2}\Big]\\ &\leq \frac{\mathbb{E}[\|P_\varepsilon\|_\infty^2]}{n} + nC_1^2\mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega+\varepsilon h, X(s)(\omega))}{\varepsilon}\Big|^2 ds\Big]\\ &\leq \frac{\mathbb{E}[\|P_\varepsilon\|_\infty^2]}{n} + nC_1^2\int_0^T \mathbb{E}\Big[\|P_\varepsilon\|_{\infty,s}^2\Big]ds,\end{aligned}$$

$$\begin{aligned}\text{For (3.25)} \Rightarrow &\mathbb{E}\Big[\sup_{t\in[0,T]} 2\int_0^t \Big\langle P_\varepsilon(s), \frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega)) - \sigma(s,\omega,X(s)(\omega))}{\varepsilon}dW(s)\Big\rangle\Big]\\ \leq&\frac{\mathbb{E}[\|P_\varepsilon\|_\infty^2]}{n} + nC_1\mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega)) - \sigma(s,\omega,X(s)(\omega))}{\varepsilon}\Big|^2 ds\Big],\end{aligned}$$

$$
\begin{aligned}
\text{For (3.26)} \Rightarrow &\mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega,X(s)(\omega))}{\varepsilon}\Big|^2 ds\Big] \\
&\leq 2\mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega+\varepsilon h,X(s)(\omega))}{\varepsilon}\Big|^2 ds\Big] \qquad (3.27)\\
&\quad + 2\mathbb{E}\Big[\int_0^T \Big|\frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega)) - \sigma(s,\omega,X(s)(\omega))}{\varepsilon}\Big|^2 ds\Big],
\end{aligned}
$$

and finally that (3.27) $\leq 2L^2 \int_0^T \mathbb{E}\Big[\|P_\varepsilon\|^2_{\infty,s}\Big] ds$.

Combining all these inequalities and choosing $n = 6$, we have

$$
\frac{1}{6}\mathbb{E}\Big[\|P_\varepsilon\|^2_\infty\Big] \leq \mathbb{E}\Big[\|A_\varepsilon\|^2_\infty\Big] + \bar{C}_1 \int_0^T \mathbb{E}\Big[\|P_\varepsilon\|^2_{\infty,s}\Big] ds,
$$

where $\mathbb{E}\big[\|A_\varepsilon\|^2_\infty\big] = O(1)$ as $\varepsilon \to 0$. Grönwall's inequality yields that $\mathbb{E}[\,\|P_\varepsilon\|^2_\infty] = O(1)$ as $\varepsilon \to 0$. □

Stochastic Gateaux differentiability of X

Next we prove the convergence in probability statement of Definition 3.9.

Theorem 3.23. *Let X be solution to the SDE* (2.1) *under Assumption 3.1 and let $h \in H$. Then we have as $\varepsilon \to 0$*

$$
\Big\|\frac{X(\cdot)(\omega+\varepsilon h) - X(\cdot)(\omega)}{\varepsilon} - \int_0^{\cdot} M_s(\cdot)(\omega)\dot{h}(s)ds\Big\|_\infty \xrightarrow{\mathbb{P}} 0.
$$

Hence X satisfies Definition 3.9, i.e. is Stochastically Gâteaux differentiable.

Proof. Let $t \in [0,T]$. To make the proof more readable we introduce several shorthand notations M^h, P_ε and Y_ε, to denote increments and its differences, namely, define

$$
M^h(t)(\omega) := \int_0^t M_s(t)(\omega)\dot{h}(s)ds, \quad P_\varepsilon(t)(\omega) := \tfrac{X(t)(\omega+\varepsilon h) - X(t)(\omega)}{\varepsilon},
$$

and $Y_\varepsilon(t)(\omega) := P_\varepsilon(t)(\omega) - M^h(t)(\omega)$. The proof's goal is to show that $\|Y_\varepsilon(\cdot)(\omega)\|_\infty \xrightarrow{\mathbb{P}} 0$ as $\varepsilon \searrow 0$.

Methodologically, we write out the SDE for $Y_\varepsilon(t)(\omega) = P_\varepsilon(t)(\omega) - M^h(t)(\omega)$ which we then break into a sequence of terms that are manipulated individually to yield an final inequality amenable to our Grönwall type result for Convergence in Probability of Proposition 2.6.

Firstly, we have

$$
\begin{aligned}
P_\varepsilon(t)(\omega) = &\int_0^t \sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h))\dot{h}(s)ds \\
&+ \int_0^t \Big[b(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - b(s,\omega,X(s)(\omega))\Big]ds \\
&+ \int_0^t \Big[\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega,X(s)(\omega))\Big]dW(s).
\end{aligned}
$$

This would mean we can decompose the SDE for $Y_\varepsilon = P_\varepsilon - M^h$ as

$$
\begin{aligned}
Y_\varepsilon(t)(\omega) &= P_\varepsilon(t)(\omega) - M^h(t)(\omega) = \frac{X(t)(\omega+\varepsilon h) - X(t)(\omega)}{\varepsilon} - M^h(t)(\omega) \\
&= \int_0^t \Big[\sigma(s,\omega+\varepsilon h, X(s)(\omega+\varepsilon h)) - \sigma(s,\omega,X(s)(\omega))\Big]\dot{h}(s)ds \qquad (3.28)
\end{aligned}
$$

$$+\int_0^t \Big[\frac{b(s,\omega+\varepsilon h, X(s)(\omega))-b(s,\omega,X(s)(\omega))}{\varepsilon}-\int_0^s U(r,s,\omega)\dot h(r)dr\Big]ds \tag{3.29}$$
$$+\int_0^t \Big[\frac{\sigma(s,\omega+\varepsilon h, X(s)(\omega))-\sigma(s,\omega,X(s)(\omega))}{\varepsilon}-\int_0^s V(r,s,\omega)\dot h(r)dr\Big]dW(s) \tag{3.30}$$
$$+\int_0^t \Big[\int_0^1 \nabla_x b(s,\omega+\varepsilon h,\Xi(s))d\xi-\nabla_x b(s,\omega,X(s)(\omega))\Big]P_\varepsilon(s)(\omega)ds \tag{3.31}$$
$$+\int_0^t \Big[\int_0^1 \nabla_x \sigma(s,\omega+\varepsilon h,\Xi(s))d\xi-\nabla_x \sigma(s,\omega,X(s)(\omega))\Big]P_\varepsilon(s)(\omega)dW(s) \tag{3.32}$$
$$+\int_0^t \nabla_x b(s,\omega,X(s)(\omega))Y_\varepsilon(s)(\omega)ds+\int_0^t \nabla_x \sigma(s,\omega,X(s)(\omega))Y_\varepsilon(s)(\omega)dW(s),$$

where $\Xi(\cdot)=X(\cdot)(\omega)+\xi[X(\cdot)(\omega+\varepsilon h)-X(\cdot)(\omega)]$.

Then we take $\sup$ over $t\in[0,T]$. Notice that we will not use an Itô type formula on the SDE, but proving convergence for each of the individual terms.

Firstly we consider the mean convergence of (3.28),

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\int_0^T\Big|\Big[\sigma(s,\omega+\varepsilon h,X(s)(\omega+\varepsilon h))-\sigma(s,\omega,X(s)(\omega))\Big]\dot h(s)\Big|ds\Big)^2\Big]\\
&\quad\le\|\dot h\|_2^2\cdot\mathbb{E}\Big[\int_0^T\Big|\sigma(s,\omega+\varepsilon h,X(s)(\omega+\varepsilon h))-\sigma(s,\omega,X(s)(\omega))\Big|^2ds\Big]\\
&\quad\le 2\|\dot h\|_2^2\Big(\mathbb{E}\Big[\int_0^T\Big|\sigma(s,\omega+\varepsilon h,X(s)(\omega))-\sigma(s,\omega,X(s)(\omega))\Big|^2ds\Big]\\
&\qquad+2L^2T\mathbb{E}\Big[\|X(\omega+\varepsilon h)-X(\omega)\|_\infty^2\Big]\Big)\\
&\quad\le O\Big(\varepsilon^2\Big)+O\Big(\varepsilon^2\Big),
\end{aligned}$$

hence this random variable converges to zero in mean square as $\varepsilon\to0$.

The term (3.29) converges in mean from Assumption 3.1 since as $\varepsilon\to0$

$$\mathbb{E}\Big[\int_0^T\Big|\frac{b(s,\omega+\varepsilon h,X(s)(\omega))-b(s,\omega,X(s)(\omega))}{\varepsilon}-\int_0^s U(r,s,\omega)\dot h(r)dr\Big|ds\Big]\to0.$$

The term (3.30) converges in mean from Assumption 3.1, namely as $\varepsilon\to0$

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,T]}\Big|\int_0^t\Big[\frac{\sigma(s,\omega+\varepsilon h,X(s)(\omega))-\sigma(s,\omega,X(s)(\omega))}{\varepsilon}-\int_0^s V(r,s,\omega)\dot h(r)dr\Big]dW(s)\Big|\Big]\\
&\quad\le C_1\mathbb{E}\Big[\Big(\int_0^T\Big|\frac{\sigma(s,\omega+\varepsilon h,X(s)(\omega))-\sigma(s,\omega,X(s)(\omega))}{\varepsilon}-\int_0^s V(r,s,\omega)\dot h(r)dr\Big|^2ds\Big)^{\frac12}\Big]\to0.
\end{aligned}$$

For equation (3.31), we are not able to use mean convergence arguments because the terms $\nabla_x b(s,\omega,x)$ have polynomial growth in x and we will not necessarily have enough finite moments to ensure that this term can be dominated. We already have $\lim_{\varepsilon\to0}\mathbb{E}[\|X(\omega+\varepsilon h)-X(\omega)\|_\infty]=0$, so clearly we also have convergence in probability. Also by Proposition 3.16, we have

$$\int_0^T\Big|\nabla_x b(s,\omega+\varepsilon h,X(s)(\omega+\varepsilon h))-\nabla_x b(s,\omega,X(s)(\omega))\Big|ds\xrightarrow{\mathbb{P}}0.$$

for any choice of $x\in\mathbb{R}^d$. Therefore, by continuity of $\nabla_x b$ from Assumption 3.1, we get

$$\int_0^T\Big|\int_0^1\nabla_x b(s,\omega+\varepsilon h,\Xi(s))d\xi-\nabla_x b(s,\omega,X(s)(\omega))\Big|ds\xrightarrow{\mathbb{P}}0,\quad\text{as }\varepsilon\to0.$$

Since we also have finite moments of $\|X(\omega+\varepsilon h)-X(\omega)\|_\infty/\varepsilon$ by Proposition 3.22, we can conclude that (3.31) converges to zero in probability.

For (3.32) we know that σ is Lipschitz so we have $\nabla_x\sigma$ is bounded. Hence, we won't have the same integrability issues as with (3.31). Therefore, we use convergence in mean. By the Burkholder-Davis-Gundy Inequality and recalling Proposition 3.22 we get

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{t'\in[0,T]}\Big|\int_0^{t'}\Big(\int_0^1\nabla_x\sigma(s,\omega+\varepsilon h,\Xi(s))d\xi-\nabla_x\sigma(s,\omega,X(s)(\omega))\Big)\cdot P_\varepsilon(s)(\omega)dW(s)\Big|\Big]\\
\leq& C_1\mathbb{E}\Big[\Big(\int_0^T\Big|\int_0^1\nabla_x\sigma(s,\omega+\varepsilon h,\Xi(s))d\xi-\nabla_x\sigma(s,\omega,X(s)(\omega))\Big|^2\cdot\Big|P_\varepsilon(s)(\omega)\Big|^2ds\Big)^{\frac{1}{2}}\Big]\\
\leq& C_1\mathbb{E}\Big[\Big(\int_0^T\Big|\int_0^1\nabla_x\sigma(s,\omega+\varepsilon h,\Xi(s))d\xi-\nabla_x\sigma(s,\omega,X(s)(\omega))\Big|^2ds\Big)^{\frac{1}{2}}\cdot\|P_\varepsilon(\omega)\|_\infty\Big]\\
\leq& C_1\mathbb{E}\Big[\|P_\varepsilon(\omega)\|_\infty^2\Big]^{\frac{1}{2}}\cdot\mathbb{E}\Big[\int_0^T\Big|\int_0^1\nabla_x\sigma(s,\omega+\varepsilon h,\Xi(s))d\xi-\nabla_x\sigma(s,\omega,X(s)(\omega))\Big|^2ds\Big]^{\frac{1}{2}}.
\end{aligned}$$

In the same way as earlier, by continuity of $\nabla_x\sigma$ from Assumption 3.1 and Proposition 3.16 we get

$$\int_0^T\Big|\int_0^1\nabla_x\sigma(s,\omega+\varepsilon h,\Xi(s))d\xi-\nabla_x\sigma(s,\omega,X(s)(\omega))\Big|^2ds\xrightarrow{\mathbb{P}}0.$$

Also, by boundedness of $\nabla_x\sigma$, we have the immediate domination

$$\int_0^T\Big|\int_0^1\nabla_x\sigma(s,\omega+\varepsilon h,\Xi(s))d\xi-\nabla_x\sigma(s,\omega,X(s)(\omega))\Big|^2ds\leq 4L^2T,$$

so we clearly have uniform integrability of all orders. Hence

$$\lim_{\varepsilon\to0}\mathbb{E}\Big[\int_0^T\Big|\int_0^1\nabla_x\sigma(s,\omega+\varepsilon h,\Xi(s))d\xi-\nabla_x\sigma(s,\omega,X(s)(\omega))\Big|^2ds\Big]^{\frac{1}{2}}=0.$$

Finally, the SDE for the process $Y_\varepsilon(t)(\omega)$ can be written in the convenient form

$$\begin{aligned}
Y_\varepsilon(t)(\omega)=&A_\varepsilon(\omega)+\int_0^t\nabla_x b(s,\omega,X(s)(\omega))Y_\varepsilon(s)(\omega)ds\\
&+\int_0^t\nabla_x\sigma(s,\omega,X(s)(\omega))Y_\varepsilon(s)(\omega)dW(s),
\end{aligned}$$

where the sequence A_ε is a sequence of random variables which converge to zero in probability. By Proposition 2.6 the random variable $\|Y_\varepsilon\|_\infty$ converges in probability to zero as $\varepsilon\to0$. $\square$

Strong stochastic Gâteaux differentiability

Theorem 3.24. *Let X be solution to the SDE* (2.1) *under Assumption 3.1. Then for any $h\in H$*

$$\lim_{\varepsilon\to0}\mathbb{E}\left[\left\|\frac{X(\omega+\varepsilon h)-X(\omega)}{\varepsilon}-M^h(\omega)\right\|_\infty\right]=0.$$

Hence X satisfies Equation (3.3)*, i.e. is Strong Stochastically Gâteaux differentiable.*

Proof. By Theorem 3.23, we have convergence in Probability. Combining this with Proposition 3.22 and Theorem 3.21, we have

$$\mathbb{E}\Big[\Big\|\frac{X(\omega+\varepsilon h)-X(\omega)}{\varepsilon}\Big\|_\infty^2\Big], \quad \mathbb{E}\Big[\|M^h(\omega)\|_\infty^2\Big] < \infty.$$

Apply Corollary 3.14 to conclude. □

Remark 3.25. Although convergence in probability may seem to be rather a weak result relative to the much stronger Almost sure convergence or convergence in mean square, it is actually the case that we now have both. After all, we proved that the sequence of random variables $\big(X(\cdot)(\omega+\varepsilon h)-X(\cdot)(\omega)\big)/\varepsilon$ have uniform finite p moments over ε and the limit $D^hX(\cdot)$ has finite p moments. Therefore, by standard probability theory we have mean square convergence.

Proof of the Malliavin differentiability result, Theorem 3.2

Proof of Theorem 3.2. The proof is straightforward and follows from Theorem 3.24 and Theorem 3.13. Further, the Malliavin Derivative satisfies the SDE (3.1) which has a unique solution as proved in Theorem 3.21. □

3.5 Proofs of the 2nd main result - Theorem 3.7

In order to prove the Malliavin differentiability (Theorem 3.2) under the weakest possible conditions, we only assumed enough properties to ensure convergence of the Stochastic Gâteaux Derivatives. However, the Stochastic Gâteaux differentiability conditions for b and σ do not require that b and σ are Malliavin differentiable. These conditions need to be checked by the user on a case-by-case basis. Under slightly stronger conditions, but much easier to verify, we present an argument to establish integrability and convergence of b and σ to prove Theorem 3.2.

In [6], there is a discussion about how much continuity is required for the spacial variable in the Malliavin Derivatives of b and σ in order to prove Malliavin Differentiability of the solution X. The authors prove results similar to those in this paper using much weaker continuity condition, but in doing so assume the integrability of the terms $D_sb(t,\omega,X(t))$ and $D_s\sigma(t,\omega,X(t))$. In our manuscript, we were unable to ensure integrability of b and σ evaluated at X without the Lipschitz (or otherwise tractable assumptions). Weaker continuity conditions would have allowed for examples where $b(t,\omega,X(t)(\omega)$ and $\sigma(t,\omega,X(t)(\omega))$ were not adequately integrable. Therefore, for easy to check conditions, we work under Assumption 3.6 (iii') and (iv') (see Remark 3.20).

For simplicity, we introduce Assumption 3.26 which contains all of the relevant properties of Assumption 3.6 that we require for this section. The function f represents b or σ depending on the choice of m.

Assumption 3.26. *Let $m\in\{1,2\}$. Suppose that $f:[0,T]\times\Omega\times\mathbb{R}^d\to\mathbb{R}^d$ such that*

(i) $\forall x\in\mathbb{R}^d$ $f(\cdot,\cdot,x)\in\mathbb{D}^{1,p}(L^m([0,T];\mathbb{R}^d))$.

(ii) f is Locally Lipschitz in the spacial variable i.e $\exists L_N>0$ such that $\forall x,y\in\mathbb{R}^d$ such that $|x|,|y|\le N$ and $\forall t\in[0,T]$,

$$|f(t,\omega,x)-f(t,\omega,y)|\le L_N|x-y| \quad \mathbb{P}\text{-almost surely.}$$

(iii) Df are Lipschitz in their spatial variables i.e. $\exists L>0$ constant such that $\forall(s,t)\in[0,T]^2$ and $\forall x,y\in\mathbb{R}^d$,

$$|D_sf(t,\omega,x)-D_sf(t,\omega,y)|\le L|x-y| \quad \mathbb{P}\text{-almost surely.}$$

Integrability and indistinguishability of the Malliavin Derivative

Lemma 3.27. *Let $m \in \{1,2\}$ and $p > 2$. Let X be solution to the SDE* (2.1) *under Assumption 2.1 and let f satisfy Assumption 3.26. Then*

$$\mathbb{E}\Big[\Big(\int_0^T \Big(\int_0^t |D_s f(t,\omega,X(t)(\omega))|^2 ds\Big)^{\frac{m}{2}} dt\Big)^{\frac{p}{m}}\Big] < \infty.$$

Proof. By the definition of $\mathbb{D}^{1,p}(L^m([0,T];\mathbb{R}^d))$ we have for any $t \in [0,T]$

$$\mathbb{E}\Big[\Big(\int_0^T \Big(\int_0^t |D_s f(t,\omega,0)|^2 ds\Big)^{\frac{m}{2}} dt\Big)^{\frac{p}{m}}\Big] < \infty.$$

Therefore for some constant C (depending on p, m, T, L) we have

$$\begin{aligned}\mathbb{E}\Big[\Big(\int_0^T &\Big(\int_0^t \Big|D_s f(t,\omega,X(t)(\omega))\Big|^2 ds\Big)^{\frac{m}{2}} dt\Big)^{\frac{p}{m}}\Big] \\ &\leq 2^{\frac{p-m}{m}} C\left(\mathbb{E}\Big[\Big(\int_0^T \Big(\int_0^t |D_s f(t,\omega,0)|^2 ds\Big)^{\frac{m}{2}} dt\Big)^{\frac{p}{m}}\Big] + \mathbb{E}\Big[\|X\|_\infty^p\Big]\right) < \infty. \quad \square\end{aligned}$$

We have by Assumption 3.26 that for every $x \in \mathbb{R}^d$ the random field $f(\cdot,\cdot,x)$ is a Malliavin differentiable process. However, it is not immediate that we have the same for $f(\cdot,\cdot,X(\cdot)(\cdot))$. We first prove an indistinguishability property for when we replace x by $X(\cdot)(\omega)$.

Lemma 3.28. *Let $m \in \{1,2\}$ and $p > 2$. Let X be solution to the SDE* (2.1) *under Assumption 2.1. Let f satisfy Assumption 3.26 and recall the directional derivative notation introduced previously, $D^h F = \langle DF, h\rangle$ for any choice of $h \in H$.*

Then, for $h \in H$ we have, (t,ω)-almost surely that

$$f\Big(t,\omega+\varepsilon h, X(t)(\omega)\Big) - f\Big(t,\omega,X(t)(\omega)\Big) = \int_0^\varepsilon D^h f(t,\omega+rh,X(t)(\omega))dr.$$

Proof. We have that $\forall x \in \mathbb{R}^d$ that $\exists C_x \subset [0,T]\times\Omega$ with $\mathbb{E}[\int_0^T \mathbb{1}_{C_x}(t,\omega)dt] = 0$, dependent on the choice of x, for which $\forall (t,\omega) \in [0,T]\times\Omega\backslash C_x$ that

$$f(t,\omega+\varepsilon h,x) - f(t,\omega,x) = \int_0^\varepsilon D^h f(t,\omega+rh,x)dr. \tag{3.33}$$

We wish to prove that we can choose a null set C which is independent of x outside of which the equality holds. To do this, it suffices to prove almost sure continuity with respect to x of both the left and right hand side of (3.33).

Almost sure continuity of the left hand side is immediate since f is locally Lipschitz. For the right hand side, we use the Lipschitz properties of the Malliavin derivative. Let r_i be an enumeration of the rationals $\mathbb{Q}^d$. Then we have $\bigcup_i C_{r_i}$ is also a null set since it is the countable union of null sets. Then for $(t,\omega) \in [0,T]\times\Omega\backslash\Big(\bigcup_i C_{r_i}\Big)$ and $\forall x \in \mathbb{Q}^d$ equation (3.33) holds. Then by the continuity of f and its Malliavin derivative we conclude that this also holds $\forall x \in \mathbb{R}^d$. $\square$

Strong stochastic Gâteaux differentiability

Lemma 3.29. *Let $m \in \{1,2\}$ and $p > 2$. Let X be solution to the SDE* (2.1) *under Assumption 2.1. Let f satisfy Assumption 3.26. Then*

$$\mathbb{E}\left[\Big(\int_0^T \Big|\frac{f(t,\omega+\varepsilon h,X(t)(\omega)) - f(t,\omega,X(t)(\omega))}{\varepsilon}\Big|^m dt\Big)^{\frac{2}{m}}\right] = O(1), \quad \text{as } \varepsilon \searrow 0.$$

Proof. Fix $\varepsilon > 0$. By Lemma 3.28, for almost all $\omega \in \Omega$ we have that

$$\int_0^T \big|f(t,\omega+\varepsilon h, X(t)(\omega)) - f(t,\omega,X(t)(\omega))\big|^m dt = \int_0^T \Big|\int_0^\varepsilon D^h f(t,\omega+rh,X(t)(\omega))dr\Big|^m dt.$$

Arguing from this, we have with the help of the directional derivative D^h, Jensen and reverse Jensen inequality,

$$\begin{aligned}
&\Big(\int_0^T \big|f(t,\omega+\varepsilon h, X(t)(\omega)) - f(t,\omega,X(t)(\omega))\big|^m dt\Big)^{\frac{2}{m}} \\
&\quad= \Big(\int_0^T \Big|\int_0^\varepsilon D^h f(t,\omega+rh,X(t)(\omega))dr\Big|^m dt\Big)^{\frac{2}{m}} \\
&\quad\le \varepsilon \int_0^\varepsilon \Big(\int_0^T |D^h f(t,\omega+rh,X(t)(\omega))|^m dt\Big)^{\frac{2}{m}} dr \\
&\quad\le \varepsilon \|\dot h\|_2^2 \int_0^\varepsilon \Big(\int_0^T \Big(\int_0^t |D_s f(t,\omega+rh,X(t)(\omega))|^2 ds\Big)^{\frac{m}{2}} dt\Big)^{\frac{2}{m}} dr \\
&\quad\le 2^{\frac{2}{m}} \varepsilon \|\dot h\|_2^2 \Bigg(\int_0^\varepsilon \Big(\int_0^T \Big(\int_0^t |D_s f(t,\omega+rh,0)|^2 ds\Big)^{\frac{m}{2}} dt\Big)^{\frac{2}{m}} dr + \varepsilon \|X(\omega)\|_\infty^2 \cdot T^{\frac{2}{m}+1}\Bigg).
\end{aligned}$$

Therefore

$$\begin{aligned}
&\mathbb{E}\Big[\frac{1}{\varepsilon^2}\Big(\int_0^T \big|f(t,\omega+\varepsilon h, X(t)(\omega)) - f(t,\omega,X(t)(\omega))\big|^m dt\Big)^{\frac{2}{m}}\Big] \\
&\qquad\le 2^{\frac{2}{m}} \|\dot h\|_2^2 \mathbb{E}\Big[\frac{1}{\varepsilon}\int_0^\varepsilon \Big(\int_0^T \Big(\int_0^t |D_s f(t,\omega+rh,0)|^2 ds\Big)^{\frac{m}{2}} dt\Big)^{\frac{2}{m}} dr\Big] \qquad (3.34)\\
&\qquad\quad + 2^{\frac{2}{m}} \|\dot h\|_2^2 T^{\frac{2}{m}+1} \mathbb{E}\Big[\|X(\omega)\|_\infty^2\Big].
\end{aligned}$$

We estimate term (3.34) as follows and with the help of Proposition 3.15

$$\begin{aligned}
(3.34) \le\,& 2^{\frac{2}{m}} \|\dot h\|_2^2 \frac{1}{\varepsilon} \int_0^\varepsilon \mathbb{E}\Big[\int_0^T \Big(\int_0^t |D_s f(t,\omega,0)|^2 ds\Big)^{\frac{m}{2}} \cdot \mathcal{E}(r\dot h)(t) dt\Big]^{\frac{2}{m}} dr \\
\le\,& 2^{\frac{2}{m}} \|\dot h\|_2^2 \mathbb{E}\Big[\Big(\int_0^T \Big(\int_0^t |D_s f(t,\omega,0)|^2 ds\Big)^{\frac{m}{2}} dt\Big)^{\frac{p}{m}}\Big]^{\frac{2}{p}} \frac{1}{\varepsilon}\int_0^\varepsilon \mathbb{E}\Big[\|\mathcal{E}(r\dot h)(\cdot)\|_\infty^{\frac{p}{p-m}}\Big]^{\frac{2(p-m)}{pm}} dr \\
<\,& O(1),
\end{aligned}$$

with $\mathcal{E}(r\dot h)$ denoting the stochastic exponential of $r\dot h$ as introduced in (3.5). □

Lemma 3.30. *Let $m \in \{1,2\}$ and $p > 2$. Let X be solution to the SDE* (2.1) *under Assumption 2.1. Let f satisfy Assumption 3.26. Then for $h \in H$ and any $\delta > 0$*

$$\lim_{\varepsilon\to 0} \mathbb{P}\Big[\Big(\int_0^T \Big|\frac{1}{\varepsilon}\int_0^\varepsilon D^h f\big(t,\omega+rh,X(t)(\omega)\big)dr - D^h f\big(t,\omega,X(t)(\omega)\big)\Big|^m dt > \delta\Big] = 0. \quad (3.35)$$

Proof. By Proposition 3.16, we know that for any $\delta > 0$ that

$$\lim_{\varepsilon\to 0} \mathbb{P}\Big[\int_0^T \Big|D^h f(t,\omega+\varepsilon h, X(t)(\omega+\varepsilon h)) - D^h f(t,\omega,X(t)(\omega))\Big|^m dt > \delta\Big] = 0. \qquad (3.36)$$

Similarly

$$\lim_{\varepsilon\to 0} \mathbb{P}\Big[\|X(\omega+\varepsilon h) - X(\omega)\|_\infty > \delta\Big] = 0,$$

so by Lipschitz continuity of Df we also have

$$\lim_{\varepsilon\to 0}\mathbb{P}\Big[\int_0^T \Big|D^h f(t,\omega+\varepsilon h, X(t)(\omega+\varepsilon h)) - D^h f(t,\omega+\varepsilon h, X(t)(\omega))\Big|^m dt > \delta\Big] = 0. \quad (3.37)$$

Combining Equations (3.36) and (3.37), we conclude

$$\lim_{\varepsilon\to 0}\mathbb{P}\Big[\int_0^T \Big|D^h f(t,\omega+\varepsilon h, X(t)(\omega)) - D^h f(t,\omega, X(t)(\omega))\Big|^m dt > \delta\Big] = 0.$$

Next, using the Fundamental Theorem of Calculus, we also have

$$\lim_{\varepsilon\to 0}\mathbb{P}\Big[\int_0^T \Big|\frac{1}{\varepsilon}\int_0^\varepsilon D^h f\big(t,\omega+rh, X(t)(\omega)\big)dr - D^h f\big(t,\omega, X(t)(\omega)\big)\Big|^m dt > \delta\Big] = 0. \qquad \square$$

The next result establishes the Strong Stochastic Gâteaux differentiability, see Definition 3.12.

Lemma 3.31. *Let $m \in \{1,2\}$ and $p > 2$. Let X be solution to the SDE* (2.1) *under Assumption 2.1. Let f satisfy Assumption 3.26. Then for $h \in H$*

$$\lim_{\varepsilon\to 0}\mathbb{E}\Big[\Big(\int_0^T \Big|\frac{f(t,\omega+\varepsilon h, X(t)(\omega)) - f(t,\omega,X(t)(\omega))}{\varepsilon} - D^h f(t,\omega,X(t)(\omega))\Big|^m dt\Big)^{\frac{1}{m}}\Big] = 0.$$

Proof. First, using Lemma 3.28, we have $\mathbb{P}$-almost surely that

$$\begin{aligned}\int_0^T &\Big|\frac{f(t,\omega+\varepsilon h, X(t)(\omega)) - f(t,\omega,X(t)(\omega))}{\varepsilon} - D^h f(t,\omega,X(t)(\omega))\Big|^m dt\\ &= \int_0^T \Big|\frac{1}{\varepsilon}\int_0^\varepsilon D^h f\big(t,\omega+rh, X(t)(\omega)\big)dr - D^h f\big(t,\omega, X(t)(\omega)\big)\Big|^m dt.\end{aligned}$$

By Lemma 3.30, both sides converge to 0 in probability (as $\varepsilon \to 0$).

Next, by Lemma 3.27 and Lemma 3.29, we have uniform L^1 integrability of this collection of random variables since they are bounded in L^2. Convergence in probability and Uniform Integrability imply convergence in mean. $\square$

Proof of Theorem 3.7

Proof of Theorem 3.7. The difference between Assumptions 3.1 and Assumptions 3.6 is (iii') and (iv'). Here we verify that b and σ satisfying Assumption 3.6 implies Assumptions 3.1.

Lemma 3.27 implies Assumptions 3.1 (iii) is satisfied. Lemma 3.31 implies Assumptions 3.1 (iv) is satisfied. In this case, the identification U, V with Db and $D\sigma$ respectively is straightforward.This also means that the Existence proof in Theorem 3.21 holds so a solution to the SDE (3.2) must exist. $\square$

4 Parametric differentiability

In this section, we study the differentiability properties of solutions of SDEs with respect to the initial condition. For a detailed exploration of the subject of Stochastic flows, see [9]. The main contribution of this section is to prove similar results for SDEs with only locally Lipschitz and monotone coefficients as opposed to previous results which rely on a Lipschitz condition. Similar problems have been studied in [15], [1, Chapter 1] and [19].

4.1 Gâteaux and Frechét differentiability of monotone SDEs

We start by recalling the concept of Gâteaux and Frechét Differentiability for abstract Banach Spaces.

Definition 4.1 (Gâteaux and Frechét Differentiability). *Let V and W be Banach spaces and let U be an open subset of V. Let $f : U \to W$. The map f is* Gâteaux differentiable *at $x \in U$ in direction $h \in V$ if the limit*

$$\lim_{\varepsilon \to 0} \frac{f(x+\varepsilon h) - f(x)}{\varepsilon} = \frac{d}{d\varepsilon} f(x+\varepsilon h),$$

exists. The limit is called the Gâteaux derivative in direction h.

The map f is said to be Frechét differentiable at $x \in U$ if there exists a bounded linear operator $A : U \to W$ such that

$$\lim_{\|h\|_V \to 0} \frac{\|f(x+h) - f(x) - Ah\|_W}{\|h\|_V} = 0.$$

The linear operator A is called the Frechét derivative of f at x

Let X_θ be the solution of SDE (2.1). We next show that the map $\theta \in L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P}) \mapsto X_\theta(\cdot) \in \mathcal{S}^p([0,T])$ is Frechét differentiable. As we will be differentiating with respect to θ for this section, we emphasize the dependency on θ.

Assumption 4.2. *Let $b : [0,T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0,T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ satisfy Assumption 2.1 for some $p \geq 2$. Further, suppose*

(i) For almost all $(t,\omega) \in [0,T] \times \Omega$ we have the functions $\sigma(t,\omega,\cdot)$ and $b(t,\omega,\cdot)$ have partial derivatives in all directions.

(ii) For all $x \in \mathbb{R}^d$, we have $\mathbb{P}$-almost sure continuity of the maps

$$x \mapsto \int_0^T \Big|\nabla_x \sigma(t,\omega,x)\Big|^2 dt \quad \text{and} \quad x \mapsto \int_0^T \Big|\nabla_x b(t,\omega,x)\Big|^2 dt.$$

Theorem 4.3. *Let $p \geq 2$ and let $1 \leq q < p$. Let X_θ be the solution of SDE* (2.1) *under Assumption 4.2 in $\mathcal{S}^q$. Then the map $\theta \to X_\theta$ is Gâteaux Differentiable in direction h and the derivative is equal to $F[h]$ the solution of the SDE* (4.1)

Further, the operator $F : L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P}) \to S^q([0,T])$ is the Frechét derivative.

Remark 4.4. It is important to note that we were unable to prove Gâteaux Differentiability in the Banach space $\mathcal{S}^p$. Convergence in $\mathcal{S}^p$ would be equivalent to uniform integrability of the random variable

$$\Big\|\frac{X_{\theta+h} - X_\theta - F[h]}{\|h\|_{L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})}}\Big\|_\infty^p,$$

over all possible choices of $h \in L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P})$. Unlike in the case where the coefficients are Lipschitz, see [2], this is not true.

The proof is given after several intermediary results. The first results relates to Gâteaux differentiability and its properties, we address the Frechét differentiability afterwards. For the proof once one has established Gâteaux differentiability, extending to Frechét differentiability is remarkably easy. Gâteaux differentiability is the weaker condition and is usually considered the easier property to prove.

Existence and uniqueness for the candidate process

Theorem 4.5. *Let $p \geq 2$ and suppose Assumption 4.2 holds. Let X_θ be the solution to* (2.1)*. Let $h \in L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P})$. Then the SDE*

$$F(t)[h] = h + \int_0^t \nabla_x b\Big(s, \omega, X_\theta(s)(\omega)\Big) F(s)[h] ds + \int_0^t \nabla_x \sigma\Big(s, \omega, X_\theta(s)(\omega)\Big) F(s)[h] dW(s), \tag{4.1}$$

has a unique solution in $\mathcal{S}^p([0,T]; \mathbb{R}^d)$.

Proof. This just follows from Theorem 2.5. We simply verify that Assumption 2.4 holds:

1. $|\nabla_x \sigma| < L$ by the Lipschitz property we have $\mathbb{E}\big[\int_0^T |\nabla_x \sigma(s, \omega, X_\theta(s))|^2 ds\big] < \infty$.
2. From the differentiability and the monotonicity property of b, we have that $\nabla_x b$ is $\mathbb{P}$-almost surely negative semidefinite[1]. Therefore, for $z \in \mathbb{R}^d$

$$z^T \Big(\int_0^T \nabla_x b(s, \omega, X_\theta(s)) ds\Big) z \leq \int_0^T L|z|^2 ds \leq LT|z|^2,$$

Hence, using the moment estimates we conclude that $\mathbb{E}\big[\|F[h]\|_\infty^p\big] \lesssim \|h\|^p_{L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P})}$. $\square$

Unlike with the Malliavin Derivative, the SDE (4.1) is not a general linear stochastic differential equation. As b and σ do not have dependency on θ, we do not have extra terms akin to the Malliavin derivatives Db and $D\sigma$. This means that, unlike the Malliavin Derivative, F has finite moments of all orders provided the initial condition has adequate integrability.

Proposition 4.6. *Let $p \geq 2$. Suppose Assumption 4.2. Let X_θ be the solution to* (2.1)*. The operator $F : L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P}) \to \mathcal{S}^p([0,T])$ defined by $h \mapsto F[h]$ the solution of Equation* (4.1)*, is a bounded linear operator $\|F[h]\|_{\mathcal{S}^p} \lesssim \|h\|_{L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P})}$.*

Proof. Firstly, we show that $F[0](\cdot) = 0_d$ a.s. (0_d is the $\mathbb{R}^d$-vector of zeros). Since $F[0]$ is the solution to the SDE

$$\begin{aligned} F(t)[0] =& \int_0^t \nabla_x b(s, \omega, X_\theta(s)(\omega)) F(s)[0] ds + \int_0^t \nabla_x \sigma(s, \omega, X_\theta(s)(\omega)) F(s)[0] dW(s), \\ F(0)[0] =& 0 \end{aligned}$$

and this SDE has a unique solution, we only need to show that $F[0](\cdot) = 0_d$ is a solution. Clearly we have $\mathbb{P}$-almost surely that

$$\int_0^t \nabla_x b(s, \omega, X_\theta(s)(\omega)) \cdot 0_d ds = 0 \quad \text{and} \quad \int_0^t \nabla_x \sigma(s, \omega, X_\theta(s)(\omega)) \cdot 0_d dW(s) = 0,$$

so this is immediate.

Let $\lambda \in \mathbb{R}$. Next we have

$$\begin{aligned} &F[h_1](t) + \lambda F[h_2](t) \\ &= h_1 + \lambda h_2 + \int_0^t \nabla_x b(s, \omega, X_\theta(s)(\omega)) F[h_1](s) ds + \lambda \int_0^t \nabla_x b(s, \omega, X_\theta(s)(\omega)) F[h_2](s) ds \\ &\quad + \int_0^t \nabla_x \sigma(s, \omega, X_\theta(s)(\omega)) F[h_1](s) dW(s) + \lambda \int_0^t \nabla_x \sigma(s, \omega, X_\theta(s)(\omega)) F[h_2](s) dW(s), \end{aligned}$$

[1] We do not prove this fact; it is straightforward using inner products and the definition of derivative.

$$
\begin{aligned}
&\Big(F[h_1] + \lambda F[h_2]\Big)(t) \\
&= (h_1 + \lambda h_2) + \int_0^t \nabla_x b(s, \omega, X_\theta(s)(\omega))\Big(F[h_1](s) + \lambda F[h_2](s)\Big)(s)ds \\
&\qquad + \int_0^t \nabla_x \sigma(s, \omega, X_\theta(s)(\omega))\Big(F[h_1](s) + \lambda F[h_1](s)\Big)(s)dW(s),
\end{aligned}
$$

which is the same as the SDE for $F[h_1 + \lambda h_2]$. Hence, by existence and uniqueness, the two must be equal up to a null set.

For the boundedness, observe that $\|F[h]\|_{\mathcal{S}^p} \lesssim \|h\|_{L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})}$ from Theorem 4.5. □

Differentiability of $\theta \mapsto X_\theta$

It is immediate to prove the stochastic stability result that $\mathbb{E}\big[\|X_{\theta+h} - X_\theta\|_\infty^p\big]^{1/p} = O\big(\|h\|_{L^p}\big)$ as $\|h\|_{L^p} \to 0$, see Theorem 2.2. Hence we have

$$
\lim_{\|h\|_{L^p}\to 0} \mathbb{E}\big[\|X_{\theta+h} - X_\theta\|_\infty^p\big] \to 0.
$$

Theorem 4.7. *Let $p \geq 2$ and $1 \leq q < p$. Let $h \in L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})$. Suppose we have Assumption 4.2, let X_θ be the solution of the SDE* (2.1) *and let $F(t)[h]$ be the solution to the SDE* (4.1)*. Then we have*

$$
\|X_{\theta+h} - X_\theta - F[h]\|_{\mathcal{S}^q} = o\big(\|h\|_{L^p}\big),
$$

and therefore $F[h]$ is the Gâteaux derivative of X.

Proof. Let $t \in [0, T]$. Define $\Xi(\cdot) = X_\theta(\cdot) + \xi[X_{\theta+h}(\cdot) - X_\theta(\cdot)]$ and consider

$$
\begin{aligned}
\frac{X_{\theta+h}(t) - X_\theta(t) - F[h](t)}{\|h\|_{L^p}} &= \frac{(\theta+h) - \theta - h}{\|h\|_{L^p}} \\
&+ \int_0^t \Big[\int_0^1 \nabla_x b(s,\omega,\Xi(s))d\xi - \nabla_x b(s,\omega,X_\theta(s))\Big] \cdot \Big[\tfrac{X_{\theta+h}(s)-X_\theta(s)}{\|h\|_{L^p}}\Big]ds \qquad (4.2) \\
&+ \int_0^t \Big[\int_0^1 \nabla_x \sigma(s,\omega,\Xi(s))d\xi - \nabla_x \sigma(s,\omega,X_\theta(s))\Big] \cdot \Big[\tfrac{X_{\theta+h}(s)-X_\theta(s)}{\|h\|_{L^p}}\Big]dW(s) \qquad (4.3) \\
&+ \int_0^t \nabla_x b(s,\omega,X_\theta(s))\Big[\tfrac{X_{\theta+h}(s)-X_\theta(s)-F(s)[h](s)}{\|h\|_{L^p}}\Big]ds \\
&+ \int_0^t \nabla_x \sigma(s,\omega,X_\theta(s))\Big[\tfrac{X_{\theta+h}(s)-X_\theta(s)-F(s)[h]}{\|h\|_{L^p}}\Big]dW(s).
\end{aligned}
$$

Arguing the same way as in Theorem 3.23, we show that Equation (4.2) and (4.3) converge to zero in probability as $\|h\|_{L^p} \to 0$. Then we apply Proposition 2.6 to conclude that

$$
\frac{\|X_{\theta+h} - X_\theta - F[h]\|_\infty}{\|h\|_{L^p}} \xrightarrow{\mathbb{P}} 0.
$$

Finally, from Theorem 2.2 and Theorem 4.5 we have that

$$
\frac{\mathbb{E}\Big[\|X_{\theta+h} - X_\theta\|_\infty^p\Big]}{\|h\|_{L^p}^p} = O(1), \quad \frac{\mathbb{E}\Big[\|F[h]\|_\infty^p\Big]}{\|h\|_{L^p}^p} = O(1) \qquad \text{as} \quad \|h\|_{L^p} \to 0.
$$

Therefore, the random variable $\Big\|\frac{X_{\theta+h}(t)-X_\theta(t)-F[h](t)}{\|h\|_{L^p}}\Big\|_\infty^q$ is uniformly integrable and we conclude

$$
\Big\|\frac{X_{\theta+h} - X_\theta - F[h]}{\|h\|_{L^p}}\Big\|_{\mathcal{S}^q} \to 0.
$$

□

Proof of the Frechét differentiability theorem

Proof of Theorem 4.3. In Proposition 4.6 we proved that F is a bounded linear operator and in Theorem 4.7 we proved that it satisfies Definition 4.1. □

4.2 Classical differentiability of SDEs

For this section, we will be studying the specific case where $\theta = x$ (a constant point in $\mathbb{R}^d$) and our perturbations are all in the constant function directions. Fix $(t,\omega) \in [0,T]\times\Omega$ and consider the map $x \in \mathbb{R}^d \mapsto X_x(t,\omega)$. We will be proving that, with probability 1 and for Lebesgue almost all $t \in [0,T]$, it is a diffeomorphism from $\mathbb{R}^d$ to $\mathbb{R}^d$. For this section, $h \in \mathbb{R}^d$ will represent some deterministic vector in Euclidean space. We will be calculating the partial derivatives in direction h.

The Jacobian matrix J

Definition 4.8. *Let $p \geq 2$. Let X_x be solution to the SDE* (2.1) *under Assumption 4.2 and with initial condition $X_x(0) = x \in \mathbb{R}^d$. Let I_d be the d-dimensional identity matrix. For $q \geq 1$ and let $J \in \mathcal{S}^q([0,T];\mathbb{R}^{d\times d})$ be the solution of the matrix valued SDE, $t \in [0,T]$*

$$J(t) = I_d + \int_0^t \nabla_x b\Big(s,\omega,X_x(s)(\omega)\Big)J(s)ds + \int_0^t \nabla_x\sigma\Big(s,\omega,X_x(s)(\omega)\Big)J(s)dW(s). \tag{4.4}$$

Notice that Equation (4.4) is the same SDE as (2.4). This means the Jacobian has an explicit solution which will be useful in Section 5 below.

Theorem 4.9. *Let $p \geq 2$. Let X_x be solution to the SDE* (2.1) *under Assumption 4.2 and with initial condition $x \in \mathbb{R}^d$. Then the SDE* (4.4) *has a unique solution in $\mathcal{S}^p$ and for any choice of $t \in [0T]$ the map $x \mapsto X_x(t)(\omega)$ is differentiable $\mathbb{P}$-almost surely. The derivative is almost surely equal to the solution of the Jacobian Equation, SDE* (4.4)*.*

Differentiability of X_x

In the previous section we proved almost sure continuity of $\|X_{x+\varepsilon h} - X_x\|_\infty/\varepsilon$, we need to show that the limit as $\varepsilon \to 0$ is equal to the solution of the Jacobian SDE.

Assumption 4.10. *Let $b : [0,T]\times\Omega\times\mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : [0,T]\times\Omega\times\mathbb{R}^d \to \mathbb{R}^{d\times m}$ satisfy Assumption 2.1 for some $p \geq 2$. Further, suppose that $\nabla_x b : [0,T]\times\Omega\times\mathbb{R}^d \to \mathbb{R}^{d\times d}$ and $\nabla_x\sigma : [0,T]\times\Omega\times\mathbb{R}^d \to \mathbb{R}^{d\times m\times d}$ are progressively measurable and that*

(i) For almost all $(t,\omega) \in [0,T]\times\Omega$ we have the functions $\sigma(t,\omega,\cdot)$ and $b(t,\omega,\cdot)$ have partial derivatives in all directions.

(ii) For $x \in \mathbb{R}^d$, we have that the maps $\mathbb{R}^d \to L^0(\Omega)$

$$x \mapsto \int_0^T \Big|\nabla_x\sigma(t,\omega,x)\Big|^2 dt \qquad \text{and} \qquad x \mapsto \int_0^T \Big|\nabla_x b(t,\omega,x)\Big|^2 dt,$$

are continuous (where convergence in L^0 means convergence in probability).

(iii) For almost all $(t,\omega) \in [0,T]\times\Omega$ we have

$$|\nabla_x\sigma(t,\omega,x) - \nabla_x\sigma(t,\omega,y)| \leq L|x-y|.$$

(iv) For $x, y \in \mathbb{R}^d$ such that $|x|, |y| < N$ and for almost all $(s,\omega) \in [0,T]\times\Omega$, $\exists L_N > 0$ such that

$$|\nabla_x b(s,\omega,x) - \nabla_x b(s,\omega,y)| \leq L_N|x-y|.$$

Proposition 4.11. *Let $p \geq 2$. Let X_x be solution to the SDE* (2.1) *under Assumption 4.10 and with initial condition $x \in \mathbb{R}^d$. Then we have that the map*

$$\varepsilon \mapsto \Big\| \frac{X_{x+\varepsilon h}(\omega) - X_x(\omega)}{\varepsilon} \Big\|_\infty ,$$

can be extended to when $\varepsilon = 0$ and the extension is almost surely continuous.

Proof. By the Stochastic Stability from Theorem 2.2 we have $\mathbb{E}\big[\|X_x - X_y\|_\infty^p\big] \lesssim |x-y|^p$, hence by the Kolmogorov Continuity Criterion we have that the map $\varepsilon \mapsto X_{x+\varepsilon h}$ is almost surely continuous. In fact, one can show α-Hölder continuity for $\alpha < 1$ but not for when $\alpha = 1$ (which would imply Lipschitz Continuity). Therefore, we additionally need to prove almost sure continuity of the map $\varepsilon \mapsto (X_{x+\varepsilon h} - X_x)/\varepsilon$.

Denote for any $t \in [0,T]$ the auxiliary process $K_\varepsilon(t) = (X_{x+\varepsilon h}(t) - X_x(t))/\varepsilon$. This process satisfies the Linear SDE

$$\begin{aligned} K_\varepsilon(t) = h &+ \int_0^t \Big[\int_0^1 \nabla_x b\big(s, \omega, X_x(s) + \xi[X_{x+\varepsilon h}(s) - X_x(s)]\big) d\xi \Big] K_\varepsilon(s) ds \\ &+ \int_0^t \Big[\int_0^1 \nabla_x \sigma\big(s, \omega, X_x(s) + \xi[X_{x+\varepsilon h}(s) - X_x(s)]\big) d\xi \Big] K_\varepsilon(s) dW(s), \end{aligned}$$

and, introducing the auxiliary process $\Xi_\varepsilon(\cdot) := X_x(\cdot) + \xi[X_{x+\varepsilon h}(\cdot) - X_x(\cdot)]\big)$, we can write the explicit solution of K_ε (as it is the solution a geometric Brownian motion type SDE)

$$K_\varepsilon(t) = h \cdot \exp\left(\int_0^t \Big[\int_0^1 \nabla_x b\Big(s, \omega, \Xi_\varepsilon(s)\Big) d\xi \Big] ds \right) \cdot \mathcal{E}\left(\int_0^1 \nabla_x \sigma\Big(\cdot, \omega, \Xi_\varepsilon(\cdot)\Big) d\xi \right)(t), \quad (4.5)$$

where $\mathcal{E}$ is the Doléan-Dade operator introduced in (3.5), which for shorthand notation we denote $K'_\varepsilon(t) = \mathcal{E}\big(\int_0^1 \nabla_x \sigma\big(\cdot, \omega, \Xi_\varepsilon(\cdot)\big) d\xi\big)(t)$.

We now analyze the behaviour of differences of increments of K'_ε in ε parameter. Take $\delta > 0$, using the properties of the Doléan-Dade exponential, we have

$$\begin{aligned} K'_\varepsilon(t) - K'_\delta(t) = &\int_0^t \Big[\int_0^1 \nabla_x \sigma\Big(s, \omega, \Xi_\varepsilon(s)\Big) - \nabla_x \sigma\Big(s, \omega, \Xi_\delta(s)\Big) d\xi \Big] K'_\varepsilon(s) dW(s) \\ &+ \int_0^t \Big[\int_0^1 \nabla_x \sigma\Big(s, \omega, \Xi_\delta(s)\Big) d\xi \Big] \cdot \Big[K'_\varepsilon(s) - K'_\delta(s) \Big] dW(s). \end{aligned}$$

Applying Itô's formula for $f(x) = |x|^p$ and denoting $L(\cdot) = K'_\varepsilon(\cdot) - K'_\delta(\cdot)$ we get

$$\begin{aligned} |L(t)|^p = p &\int_0^t |L(s)|^{p-2} L(s)^T \cdot \Big[\int_0^1 \nabla_x \sigma\Big(s, \omega, \Xi_\varepsilon(s)\Big) \\ &\qquad - \nabla_x \sigma\Big(s, \omega, \Xi_\delta(s)\Big) d\xi \Big] K'_\varepsilon(s) dW(s) && (4.6) \\ &+ p \int_0^t |L(s)|^{p-2} L(s)^T \cdot \Big[\int_0^1 \nabla_x \sigma\Big(s, \omega, \Xi_\delta(s)\Big) d\xi \Big] \cdot L(s) dW(s) && (4.7) \\ &+ \frac{p}{2} \int_0^t |L(s)|^{p-2} \Big| \Big[\int_0^1 \nabla_x \sigma\Big(s, \omega, \Xi_\varepsilon(s)\Big) - \nabla_x \sigma\Big(s, \omega, \Xi_\delta(s)\Big) d\xi \Big] K'_\varepsilon(s) \Big|^2 ds && (4.8) \\ &+ \frac{p}{2} \int_0^t |L(s)|^{p-2} \Big| \Big[\int_0^1 \nabla_x \sigma\Big(s, \omega, \Xi_\delta(s)\Big) d\xi \Big] \cdot L(s) \Big|^2 ds && (4.9) \\ &+ \frac{p(p-2)}{2} \int_0^t |L(s)|^{p-4} \Big| L(s)^T \Big[\int_0^1 \nabla_x \sigma\Big(s, \omega, \Xi_\varepsilon(s)\Big) \\ &\qquad - \nabla_x \sigma\Big(s, \omega, \Xi_\delta(s)\Big) d\xi \Big] K'_\varepsilon(s) \Big|^2 ds && (4.10) \end{aligned}$$

$$+\frac{p(p-2)}{2}\int_0^t |L(s)|^{p-4}\Big|L(s)^T\big[\int_0^1 \nabla_x\sigma\Big(s,\omega,\Xi_\delta(s)\Big)d\xi\big]\cdot L(s)\Big|^2 ds. \tag{4.11}$$

Next, take a supremum over time then expectations. Using the methods that have already been explored in detail for the proof of Theorem 3.21, we know that the terms from lines (4.7), (4.9) and (4.11) will all yield terms of the form $\lesssim \int_0^T \mathbb{E}\big[\|L\|_{\infty,t}^p\big]dt$ which will be accounted for with the Grönwall inequality.

Firstly, following the same methods for Theorem 3.21 and using the additional Assumption 4.10(iii), for

for (4.6)

$$\begin{aligned}
&\Rightarrow p\mathbb{E}\Big[\sup_{t\in[0,T]}\int_0^t |L(s)|^{p-2}\cdot\Big[\int_0^1 \nabla_x\sigma\Big(s,\omega,\Xi_\varepsilon(s)\Big)-\nabla_x\sigma\Big(s,\omega,\Xi_\delta(s)\Big)d\xi\Big]K'_\varepsilon(s)dW(s)\Big],\\
&\le pC_1\mathbb{E}\Big[\Big(\int_0^T |L(s)|^{2p-4}\cdot\Big|\big[\int_0^1 \nabla_x\sigma\Big(s,\omega,\Xi_\varepsilon(s)\Big)-\nabla_x\sigma\Big(s,\omega,\Xi_\delta(s)\Big)d\xi\big]K'_\varepsilon(s)\Big|^2 ds\Big)^{\frac12}\Big],\\
&\le \frac{\mathbb{E}\Big[\|L\|_\infty^p\Big]}{n}+C_1^p\big[n(p-1)\big]^{p-1}\mathbb{E}\Big[\Big(TL\|X_{x+\varepsilon h}-X_{x+\delta h}\|_\infty^2\cdot\|K'_\varepsilon\|_\infty^2\Big)^{\frac p2}\Big],\\
&\le \frac{\mathbb{E}\Big[\|L\|_\infty^p\Big]}{n}+C_1^p\big[n(p-1)\big]^{p-1}(TL)^{\frac p2}\mathbb{E}\Big[\|X_{x+\varepsilon h}-X_{x+\delta h}\|_\infty^{2p}\Big]^{\frac12}\cdot\mathbb{E}\Big[\|K'_\varepsilon\|_\infty^{2p}\Big]^{\frac12}.
\end{aligned}$$

Secondly,

$$\begin{aligned}
\text{for (4.8)}\Rightarrow &\frac p2\mathbb{E}\Big[\int_0^T |L(s)|^{p-2}\Big|\big[\int_0^1 \nabla_x\sigma\Big(s,\omega,\Xi_\varepsilon(s)\Big)-\nabla_x\sigma\Big(s,\omega,\Xi_\delta(s)\Big)d\xi\big]K'_\varepsilon(s)\Big|^2 ds\Big],\\
&\le \frac{\mathbb{E}\Big[\|L\|_\infty^p\Big]}{n}+\Big[\frac{n(p-2)}{2}\Big]^{\frac{p-2}{2}}(LT)^{\frac p2}\mathbb{E}\Big[\|X_{x+\varepsilon h}-X_{x+\delta h}\|_\infty^{2p}\Big]^{\frac12}\cdot\mathbb{E}\Big[\|K'_\varepsilon\|_\infty^{2p}\Big]^{\frac12}.
\end{aligned}$$

The terms from (4.10) are treated in exactly the same way.

Finally, we use that $\mathbb{E}\big[\|K'_\varepsilon\|_\infty^{2p}\big]^{1/2}<\infty$ and $\mathbb{E}\big[\|X_{x+\varepsilon h}-X_{x+\delta h}\|_\infty^{2p}\big]^{1/2}\lesssim|\delta-\varepsilon|^p|h|^p$ to conclude

$$\mathbb{E}\Big[\|K'_\varepsilon-K'_\delta\|_\infty^p\Big]\lesssim|\varepsilon-\delta|^p|h|^p.$$

Hence by Kolmogorov Continuity Criterion, we have the map $\varepsilon\mapsto K'_\varepsilon(t)(\omega)$ is almost surely continuous for any $t\in[0,T]$ $\mathbb{P}$-almost surely.

Now, we return to Equation (4.5). Using the almost sure continuity of $\varepsilon\mapsto X_{x+\varepsilon h}(t)(\omega)$ and Assumption 4.10 (iv), we have that

$$\varepsilon\mapsto\exp\left(\int_0^t\Big[\int_0^1\nabla_x b\Big(s,\omega,X_x(\cdot)+\xi[X_{x+\varepsilon h}(\cdot)-X_x(\cdot)]\Big)d\xi\Big]ds\right),$$

is almost surely continuous. Hence $\varepsilon\mapsto K_\varepsilon(t)(\omega)$ is also almost surely continuous. □

Theorem 4.12. *Let $p\ge2$. Let X_x be solution to the SDE* (2.1) *under Assumption 4.10 and with initial condition $x\in\mathbb{R}^d$. Then we have that $\forall t\in[0,T]$*

$$\frac{X_{x+\varepsilon h}(t)(\omega)-X_x(t)(\omega)}{\varepsilon}\to h\cdot J(t)(\omega)\quad\mathbb{P}\text{-almost surely as }\varepsilon\to0.$$

Proof. Let $t\in[0,T]$. First, we show convergence in probability of $(X_{x+\varepsilon h}(t)-X_x(t))/\varepsilon$ to $h\cdot J(t)$ using Proposition 2.6. Convergence in probability will imply the existence of a subsequence which converges almost sure. Finally, using Proposition 4.11 we know the limit will be almost surely unique.

Writing out the SDE for the increments' process, we have

$$
\begin{aligned}
&\frac{X_{x+\varepsilon h}(t)-X_x(t)}{\varepsilon} - hJ(t)\\
&\quad= \int_0^t \Big[\int_0^1 \nabla_x b(s,\omega,\Xi_\varepsilon(s))d\xi - \nabla_x b(s,\omega,X_x(s))\Big]\Big[\frac{X_{x+\varepsilon h}(s)-X_x(s)}{\varepsilon}\Big]ds \qquad (4.12)\\
&\qquad + \int_0^t \Big[\int_0^1 \nabla_x \sigma(s,\omega,\Xi_\varepsilon(s))d\xi - \nabla_x \sigma(s,\omega,X_x(s))\Big]\Big[\frac{X_{x+\varepsilon h}(s)-X_x(s)}{\varepsilon}\Big]dW(s) \qquad (4.13)\\
&\qquad + \int_0^t \nabla_x b(s,\omega,X_x(s))\Big[\frac{X_{x+\varepsilon h}(s)-X_x(s)}{\varepsilon} - hJ(s)\Big]ds\\
&\qquad + \int_0^t \nabla_x \sigma(s,\omega,X_x(s))\Big[\frac{X_{x+\varepsilon h}(s)-X_x(s)}{\varepsilon} - hJ(s)\Big]dW(s),
\end{aligned}
$$

where $\Xi_\varepsilon(\cdot) = X_x(\cdot)+\xi[X_{x+\varepsilon h}(\cdot)-X_x(\cdot)]$. As with Theorem 3.23, we argue that the terms (4.12) and (4.13) converge in probability to 0, then use Proposition 2.6 to conclude that

$$\Big\|\frac{X_{x+\varepsilon h}(\omega)(\cdot)-X_x(\omega)(\cdot)}{\varepsilon} - hJ(\omega)(\cdot)\Big\|_\infty \xrightarrow{\mathbb{P}} 0.$$

Thus there exists a sequence ε_n such that $\varepsilon_n \to 0$ as $n\to\infty$ and an event $C_1 \subset \Omega$ with $\mathbb{P}[C_1]=0$ such that $\forall \omega \in \Omega \backslash C_1$

$$\lim_{n\to\infty} \Big\|\frac{X_{x+\varepsilon_n h}(\omega)-X_x(\omega)}{\varepsilon_n} - hJ(\omega)\Big\|_\infty \to 0.$$

Finally, by Proposition 4.11 there exists an event $C_2 \subset \Omega$ with $\mathbb{P}[C_2]=0$ such that $\forall \omega \in \Omega\backslash C_2$ the map

$$\varepsilon \mapsto \Big\|\frac{X_{x+\varepsilon h}(\omega)-X_x(\omega)}{\varepsilon} - hJ(\omega)\Big\|_\infty,$$

is continuous for ε at 0. Then for $\forall \omega \in \Omega\backslash(C_1\cup C_2)$

$$\lim_{\varepsilon\to 0} \Big\|\frac{X_{x+\varepsilon h}(\omega)-X_x(\omega)}{\varepsilon} - hJ(\omega)\Big\|_\infty \to 0,$$

and $\mathbb{P}[C_1\cup C_2]=0$. □

Invertibility of the Jacobian matrix

Next, we wish to show that the Jacobian Matrix $J(t)$ is $\mathbb{P}$-almost surely invertible for any choice of $t\in[0,T]$. Notice that due to the initial condition, we have that this is true for $t=0$ since $J(0)=I_d$.

To prove the Jacobian is invertible, we consider a matrix valued stochastic process and observe that for any choice of $t\in[0,T]$, this process will take value equal to the left inverse of J. This proof follows that of Nualart, [13, Chapter 2.3; Equation 2.8].

We introduce the SDE

$$
\begin{aligned}
K(t) = I_d - \int_0^t K(s)\Big[\nabla_x b(s,\omega,X(s)) - \Big\langle \nabla_x\sigma, \nabla_x \sigma\Big\rangle_{\mathbb{R}^m}\big(s,\omega,X(s)\big)\Big]ds&\\
- \int_0^t K(s)\nabla_x\sigma(s,\omega,X(s))dW(s).& \qquad (4.14)
\end{aligned}
$$

Proposition 4.13. *Let $p\geq 2$. Let X_x be solution to the SDE (2.1) under Assumption 4.2 and with initial condition $x\in\mathbb{R}^d$. Then we have the following identity $K(t)J(t)=I_d$ for all $t\in[0,T]$ $\mathbb{P}$-a.s.*

Proof. We deal here with matrix valued processes which cannot necessarily be assumed commutative, this makes the analysis slightly more involved. Itô's formula for matrices gives that $(KJ)(0) = I_d$ and

$$\begin{aligned} d(KJ)(t) =& K(t)dJ(t) + dK(t)J(t) + d[K,J](t), \\ =& K(t)\nabla_x b(t,\omega,X(t))J(t)dt + K(t)\sigma(t,\omega,X(t))J(t)dW(t) \\ &- K(t)\nabla_x b(t,\omega,X(t))J(t)dt - K(t)\sigma(t,\omega,X(t))J(t)dW(t) \\ &+ K(t)\Big\langle \nabla_x\sigma, \nabla_x\sigma\Big\rangle_{\mathbb{R}^m}\big(s,\omega,X(s)\big)J(t)dt \\ &- K(t)\Big\langle \nabla_x\sigma, \nabla_x\sigma\Big\rangle_{\mathbb{R}^m}\big(s,\omega,X(s)\big)J(t)dt = 0dt + 0dW(t). \end{aligned}$$ □

SDE (4.14) does not necessarily satisfy Assumption 2.4, the issue being that the term $-z^T\nabla_x b(t,\omega,X(t))z$ is not bounded from above by a constant almost surely for any choice of vector $|z| = 1$. However, an explicit solution to the SDE can be written out pathwise, even if it does not have finite moments. This construction has the property that it is the left inverse of J.

Proposition 4.14. *The determinant of the Matrix $J(t)$, denoted $D(t)$, is called the Stochastic Wronskian and satisfies the SDE*

$$\begin{aligned} dD(t) =& \mathit{Tr}\Big(\nabla_x b(t,\omega,X(t))\Big)D(t)dt + \mathit{Tr}\Big(\nabla_x\sigma(t,\omega,X(t))\Big)D(t)dW(t) \\ &+ \Big[\Big\langle \mathit{Tr}\big(\nabla_x\sigma(t,\omega,X(t))\big), \mathit{Tr}\big(\nabla_x\sigma(t,\omega,X(t))\big)\Big\rangle_{\mathbb{R}^m} \\ &\qquad - \mathit{Tr}\Big(\big\langle \nabla_x\sigma(t,\omega,X(t)), \nabla_x\sigma(t,\omega,X(t))\big\rangle_{\mathbb{R}^m}\Big)\Big]D(t)dt, \end{aligned} \tag{4.15}$$

with $D(0) = 1$. $D(t)$ has explicit form

$$\begin{aligned} D(t) = \exp\Bigg(&\int_0^t \mathit{Tr}\Big(\nabla_x b(s,\omega,X(s))\Big) - \frac{1}{2}\mathit{Tr}\Big(\Big\langle \nabla_x\sigma(s,\omega,X(s)), \nabla_x\sigma(s,\omega,X(s))\Big\rangle_{\mathbb{R}^m}\Big)ds \\ &+ \int_0^t \mathit{Tr}\Big(\nabla_x\sigma(s,\omega,X(s))\Big)dW(s)\Bigg). \end{aligned} \tag{4.16}$$

Proof. The proof can be found in [10, Theorem 3.2.2]. The proof involves applying Itô's formula to the determinant of $J(t)$ and establishing that it satisfies Equation (4.15). Then one applies Itô's formula to Equation (4.16) and verifies that this likewise satisfies (4.15). Finally, by Theorem 2.5, the solution is unique. □

The matrix $\nabla_x b$ being lower semidefinite means that $\mathrm{Tr}(\nabla_x b)$ is bounded from above, but not necessarily from below. We can conclude the $D(\cdot)$ is almost surely positive and therefore the process K is $\mathbb{P}$-almost surely the inverse (left or right) of J provided $\mathrm{Tr}(\nabla_x b) \neq -\infty$ with positive probability.

5 Applications

In this section, we recover and discuss some standard applications of Malliavin Differentiation and evaluate some of the problems that occur under our framework.

5.1 Representation formulae

Firstly, we present a way of writing the Malliavin Derivative of X_θ in terms of the Jacobian.

Proposition 5.1 (Representation formulae)**.** *Let X_x be solution to the SDE* (2.1) *under Assumption 3.1 and with initial condition $X_x(0) = x \in \mathbb{R}^d$. Let J satisfy the SDE* (4.4)*. Consider the SDE for the process $J(t)J(s)^{-1}$ for $t > s$.*

$$\begin{aligned}J_s(t) =&J(t)J(s)^{-1}\\ =&J(s)J(s)^{-1} + \int_s^t \nabla_x b(r,\omega,X(r)(\omega))J(r)J(s)^{-1}dr\\ &\qquad\qquad + \int_s^t \nabla_x \sigma(r,\omega,X(r)(\omega))J(r)J(s)^{-1}dW(r)\\ =&I_d + \int_s^t \nabla_x b(r,\omega,X(r)(\omega))J_s(r)dr + \int_s^t \nabla_x \sigma(r,\omega,X(r)(\omega))J_s(r)dW(r). \qquad (5.1)\end{aligned}$$

Equation (5.1) *is the Fundamental Matrix of the Linear Stochastic Differential Equation* (3.1)*. As such, under Assumption 3.1 the Malliavin Derivative of X can be expressed for $t > s$ as*

$$D_sX(t) = J_s(t)A(s,t),$$

where $A(s,t)$ is defined for $t > s$ as

$$\begin{aligned}A(s,t) = \sigma(s,\omega,X(s)(\omega)) + \int_s^t J_s(r)^{-1}\Big(U(s,r,\omega) - \Big\langle \nabla_x\sigma(r,\omega,X(r)(\omega)), V(s,r,\omega)\Big\rangle_{\mathbb{R}^m}\Big)dr\\ + \int_s^t J_s^{-1}(r)V(s,r,\omega)dW(r).\end{aligned}$$

Proof. The proof of this representation formula follows the same ideas as Theorem 3.21. Equation (3.1) is an infinite dimensional SDE, so we project from the infinite dimensional space into a finite dimensional space. We follow the method of [10, Theorem 3.3.1] to solve the solution explicitly in the projection space then use the Dominated Convergence Theorem to ensure the passage to the limit. □

Absolute continuity

In [13, Theorem 2.3.1], it is proved that the solution of a Stochastic Differential Equation with Lipschitz, deterministic coefficients and elliptic diffusion term has a law which is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^d$. This proof can be easily extended to the case where the drift term has monotone growth.

Theorem 5.2. *Let X_x be solution to the SDE* (2.1) *under Assumption 3.1 and with initial condition $X_x(0) = x \in \mathbb{R}^d$. Suppose additionally that $\forall z \in \mathbb{R}^d$ that*

$$z^T A(s,t)A(s,t)^T z > \lambda(s,t)|z|^2 \geq 0, \quad \int_0^t \lambda(s,t)ds > 0 \quad \mathbb{P}\text{-almost surely.}$$

Then the law of $X_x(t)$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$.

Proof. For this proof, recall [13, Corollary 2.1.2] and following that our strategy is to show that the Malliavin matrix is $\mathbb{P}$-almost surely non zero.

The Malliavin Matrix, $Q(t)$ is defined to be

$$Q(t) = \int_0^t D_sX(t)D_sX(t)^T ds = J(t)\left(\int_0^t K(s)A(s,t)A(s,t)^T K(s)^T ds\right)J(t)^T.$$

Therefore, for $z \in \mathbb{R}^d$ we have $z^TQ(t)z \geq \int_0^t \lambda(s,t)|K(s)|^2 ds \cdot |J(t)|^2 \cdot |z|^2$ which is greater than zero because $|J|, |K| > 0$. □

Remark 5.3. Observe that the Ellipticity condition for σ is no longer enough to ensure that the law is absolutely continuous. When b and σ are deterministic, U and V are uniformly 0 and Ellipticity is enough.

5.2 Bismut-Elworthy-Li formula

In [3], the author uses Malliavin Differentiability of an SDE X_x to prove differentiability for functions of the form $u(x) = \mathbb{E}[\phi(X_x(t))]$ where ϕ is assumed to be a continuous function and $t \in [0,T]$. This was later extended in [5] and [4] to cover functions ϕ which are integrable and even measurable (provided u remains finite).

Define for $t \in (0,T]$ the set $\Gamma_t = \big\{a \in L^2([0,T]); \int_0^t a(s)ds = 1\big\}$.

Theorem 5.4 (Bismut-Elworthy-Li formula)**.** *Let $\Phi : \mathbb{R}^d \to \mathbb{R}^d$ be a bounded, measurable function. Let X_x be solution to the SDE* (2.1) *under Assumption 3.1 and with initial condition $X_x(0) = x \in \mathbb{R}^d$. Let $t \in (0,T]$. Suppose additionally that ($\delta(\cdot)$ stands for the usual Skorokhod integral, see [13])*

1. *$\forall s \in [0,t]$ the matrix $A(s,t)$ has a right inverse,*
2. *$\exists a \in \Gamma_t$ such that $a(\cdot)A(\cdot,t)^{-1}J(\cdot) \in$ dom(δ).*

Then

$$\nabla_x \mathbb{E}\Big[\Phi(X_x(t)\Big] = \mathbb{E}\Big[\Phi(X_x(t))\delta\Big(a(s)A(s,t)^{-1}J(s)\Big)\Big].$$

Proof. We give only a sketch of the proof. For a more detailed proof, see [5] and [4]. First suppose that Φ is continuously differentiable with bounded derivatives, then

$$\begin{aligned}\nabla_x \mathbb{E}\Big[\Phi(X_x(t))\Big] =& \mathbb{E}\Big[\nabla_x \Phi(X_x(t))\Big] = \mathbb{E}\Big[\nabla\Phi(X_x(t))J(t)\Big]\\ =& \mathbb{E}\Big[\nabla\Phi(X_x(t))D_sX_x(t)A(s,t)^{-1}J(s)\Big].\end{aligned}$$

Multiplying both sides by $a \in \Gamma_s$, integrating over $[0,t]$ (using $\int_0^t a(s)ds = 1$ on the LHS) and Fubini gives

$$\begin{aligned}\nabla_x \mathbb{E}\Big[\Phi(X_x(t))\Big] &= \mathbb{E}\Big[\int_0^t a(s)\nabla\Phi(X_x(t))D_sX_x(t)A(s,t)^{-1}J(s)ds\Big]\\ =& \mathbb{E}\Big[\int_0^t D_s\Big(\Phi(X_x(t))\Big)a(s)A(s,t)^{-1}J(s)ds\Big] = \mathbb{E}\Big[\Phi\big(X_x(t)\big)\delta\Big(a(s)A(s,t)^{-1}J(s)\Big)\Big],\end{aligned}$$

where in the last line we used integration-by-parts formula.

Secondly, let Φ be bounded and measurable. Then using that C_b^1 is dense in the set of bounded measurable functions, we approximate Φ by a sequence of functions $\Phi_n \in C_b^1$. Finally, using a domination argument it is shown that one can swap the limits and integrals and one reaches the conclusion. □

A Proofs

A.1 The existence and uniqueness theorem plus moment calculations

Proof of Theorem 2.2. As $p \geq 2$, we also have that

$$\mathbb{E}\Big[\int_0^T \Big|\sigma(s,\omega,X_\theta(s))\Big|^2 ds\Big] \leq 2\mathbb{E}\Big[\int_0^T \Big|\sigma(s,\omega,0)\Big|^2 ds\Big] + 2L^2T\mathbb{E}\Big[\|X\|_\infty^2\Big] < \infty.$$

This means we can use [14, Theorem 3.2.5] to get $\mathbb{P}$-almost sure continuity of the stochastic integral. The drift term is a Lebesgue integral so likewise is continuous in time. Hence $\mathbb{P}$-almost sure continuity of $t \mapsto X_\theta(t)$ is immediate.

Finally, let $t \in [0,T]$ and $\xi, \theta \in L^p(\mathcal{F}_0;\mathbb{R}^d;\mathbb{P})$. We have

$$X_\xi(t) - X_\theta(t) = \xi - \theta + \int_0^t \Big[b(s,\omega,X_\xi(s)) - b(s,\omega,X_\theta(s))\Big]ds \\ + \int_0^t \Big[\sigma(s,\omega,X_\xi(s)) - \sigma(s,\omega,X_\theta(s))\Big]dW(s).$$

We write $Q(s) = X_\xi(s) - X_\theta(s)$ and by applying Itô's formula with $f(x) = |x|^p$ we get

$$\begin{aligned}|Q(t)|^p = |\xi-\theta|^p &+ p\int_0^t |Q(s)|^{p-2}\Big\langle Q(s), b(s,\omega,X_\xi(s)) - b(s,\omega,X_\theta(s))\Big\rangle ds \\ &+ p\int_0^t |Q(s)|^{p-2}\Big\langle Q(s), \Big[\sigma(s,\omega,X_\xi(s)) - \sigma(s,\omega,X_\theta(s))\Big]dW(s)\Big\rangle \\ &+ \frac{p}{2}\int_0^t |Q(s)|^{p-2}\cdot\Big|\sigma(s,\omega,X_\xi(s)) - \sigma(s,\omega,X_\theta(s))\Big|^2 ds \\ &+ \frac{p(p-2)}{2}\int_0^t |Q(s)|^{p-4}\Big|Q(s)^T\cdot\big[\sigma(s,\omega,X_\xi(s)) - \sigma(s,\omega,X_\theta(s))\big]\Big|^2 ds.\end{aligned}$$

Taking a supremum over time and then taking expectations, we get

$$\begin{aligned}\mathbb{E}\Big[\|X_\xi - X_\theta\|_\infty^p\Big] = \mathbb{E}\Big[\|Q\|_\infty^p\Big] \le \mathbb{E}\Big[|\xi-\theta|^p\Big] \\ + p\mathbb{E}\Big[\int_0^T |Q(s)|^{p-2}\Big|\Big\langle Q(s), b(s,\omega,X_\xi(s)) - b(s,\omega,X_\theta(s))\Big\rangle\Big|ds\Big] \quad &(A.1) \\ + p\mathbb{E}\Big[\sup_{t\in[0,T]}\int_0^t |Q(s)|^{p-2}\Big\langle Q(s), \Big[\sigma(s,\omega,X_\xi(s)) - \sigma(s,\omega,X_\theta(s))\Big]dW(s)\Big\rangle\Big] \quad &(A.2) \\ + \frac{p}{2}\mathbb{E}\Big[\int_0^T |Q(s)|^{p-2}\Big|\sigma(s,\omega,X_\xi(s)) - \sigma(s,\omega,X_\theta(s))\Big|^2 ds\Big] \quad &(A.3) \\ + \frac{p(p-2)}{2}\mathbb{E}\Big[\int_0^T |Q(s)|^{p-4}|Q(s)^T\Big[\sigma(s,\omega,X_\xi(s)) - \sigma(s,\omega,X_\theta(s))\Big]|^2 ds\Big]. \quad &(A.4)\end{aligned}$$

Firstly by monotonicity of b we have (A.1) $\le pL\int_0^T \mathbb{E}\Big[\|Q\|_{\infty,s}^p\Big]ds$. Secondly, by the Burkholder-Davis-Gundy inequality we have

$$\text{(A.2)} \le pC_1 L\mathbb{E}\Big[\|Q\|_\infty^{\frac{p}{2}}\Big(\int_0^T \|Q\|_{\infty,s}^p ds\Big)^{\frac{1}{2}}\Big] \le \frac{\mathbb{E}\Big[\|Q\|_\infty^p\Big]}{2} + \frac{p^2C_1^2L^2}{2}\int_0^T \mathbb{E}\Big[\|Q\|_{\infty,s}^p\Big]ds.$$

Finally, we have

$$\text{(A.3)} \le \frac{pL}{2}\int_0^T \mathbb{E}\Big[\|Q\|_{\infty,s}^p\Big]ds \qquad \text{and} \qquad \text{(A.4)} \le \frac{p(p-2)L}{2}\int_0^T \mathbb{E}\Big[\|Q\|_{\infty,s}^p\Big]ds.$$

Gathering all the estimates we have finally

$$\frac{1}{2}\mathbb{E}\Big[\|X_\xi - X_\theta\|_\infty^p\Big] \le \mathbb{E}\Big[|\xi-\theta|^p\Big] + \widehat{C}\int_0^T \mathbb{E}\Big[\|X_\xi - X_\theta\|_{\infty,s}^p\Big]ds,$$

where $\widehat{C} = p^2L(LC_1^2+2)/2$. Grönwall's inequality yields $\mathbb{E}\Big[\|X_\xi - X_\theta\|_\infty^p\Big] \lesssim \mathbb{E}[|\xi-\theta|^p]$. □

Moment Calculations for Theorem 2.5 . Fix $t \in [0,T]$ and using Itô's formula with $f(x) = |x|^p$ and X_θ satisfying Equation (2.3), we get that

$$
\begin{aligned}
|X_\theta(t)|^p =& |\theta|^p + p\int_0^t |X_\theta(s)|^{p-2}\langle X_\theta(s), B(s,\omega)X_\theta(s)\rangle ds + p\int_0^t |X_\theta(s)|^{p-2}\langle X_\theta(s), b(s,\omega)\rangle ds\\
&+ p\int_0^t |X_\theta(s)|^{p-2}\langle X_\theta(s), \Sigma(s,\omega)X_\theta(s)dW(s)\rangle\\
&+ p\int_0^t |X_\theta(s)|^{p-2}\langle X_\theta(s), \sigma(s,\omega)dW(s)\rangle\\
&+ \frac{p}{2}\int_0^t |X_\theta(s)|^{p-2}\Big[\Sigma(s,\omega)X_\theta(s)+\sigma(s,\omega)\Big]^2 ds\\
&+ \frac{p(p-2)}{2}\int_0^t |X_\theta(s)|^{p-4}\Big\langle X_\theta(s), \big[\Sigma(s,\omega)X_\theta(s)+\sigma(s,\omega)\big]\Big\rangle^2 ds.
\end{aligned}
$$

Take a supremum over $t\in[0,T]$ and expectations to have

$$
\begin{aligned}
\mathbb{E}\Big[\|X_\theta\|^p\Big] \leq& \mathbb{E}\Big[|\theta|^p\Big]\\
&+ p\mathbb{E}\Big[\int_0^T |X_\theta(s)|^{p-2}\Big\langle X_\theta(s), \big[B(s,\omega)X_\theta(s)+b(s,\omega)\big]\Big\rangle ds\Big] && \text{(A.5)}\\
&+ p\mathbb{E}\Big[\sup_{t\in[0,T]}\int_0^t |X_\theta(s)|^{p-2}\Big\langle X_\theta(s), \big[\Sigma(s,\omega)X_\theta(s)+\sigma(s,\omega)\big]dW(s)\Big\rangle\Big] && \text{(A.6)}\\
&+ \frac{p(p-1)}{2}\mathbb{E}\Big[\int_0^T |X_\theta(s)|^{p-2}\Big[\Sigma(s,\omega)X_\theta(s)+\sigma(s,\omega)\Big]^2 ds\Big].
\end{aligned}
$$

Fix $n\in\mathbb{N}$ to be chosen later. Throughout the next three arguments, we use Young's Inequality. Using the negative semidefinite properties of B, we get that

$$
\text{(A.5)} \leq pL\int_0^T \mathbb{E}\Big[\|X_\theta\|_{\infty,s}^p\Big]ds + \frac{\mathbb{E}\Big[\|X_\theta\|_\infty^p\Big]}{n} + n^{p-1}(p-1)^{p-1}\times\mathbb{E}\Big[\Big(\int_0^T |b(s,\omega)|ds\Big)^p\Big].
$$

Secondly, using the Burkholder-Davis-Gundy Inequality gives that

$$
\begin{aligned}
\text{(A.6)} \leq& \frac{2\mathbb{E}\Big[\|X_\theta\|_\infty^p\Big]}{n} + \frac{p^2C_1^2 n\sqrt{2}}{4}\int_0^T \mathbb{E}\Big[\|X_\theta\|_{\infty,s}^p\Big]\|\Sigma(s,\cdot)\|_{L^\infty}^2 ds\\
&+ C_1^p n^{p-1}p^{\frac{p}{2}}(p-2)^{\frac{p-2}{2}}\frac{2^{\frac{p}{4}}}{2^{p-1}}\cdot\mathbb{E}\Big[\Big(\int_0^T |\sigma(s,\omega)|^2 ds\Big)^{\frac{p}{2}}\Big].
\end{aligned}
$$

Thirdly, we have

$$
\begin{aligned}
\text{(A.6)} \leq& \frac{\mathbb{E}\Big[\|X_\theta\|_\infty^p\Big]}{n} + p(p-1)\int_0^T \mathbb{E}\Big[\|X_\theta\|_{\infty,s}^p\Big]\|\Sigma(s,\cdot)\|_{L^\infty} ds\\
&+ 2[n(p-2)]^{\frac{p-2}{2}}(p-1)^{\frac{p}{2}}\mathbb{E}\Big[\Big(\int_0^T |\sigma(s,\omega)|^2 ds\Big)^{\frac{p}{2}}\Big].
\end{aligned}
$$

When applying Young's Inequality for the case $p=2$, we use the convention that $0^0=1$. Adding these together, we have that there are constants $\widetilde{C_1}, \widetilde{C_2}$ and $\widetilde{C_3}$ such that

$$
\begin{aligned}
\frac{1}{5}\mathbb{E}\Big[\|X_\theta\|_\infty^p\Big] \leq& \mathbb{E}\Big[|\theta|^p\Big] + \widetilde{C_1}\mathbb{E}\Big[\Big(\int_0^T |b(s,\omega)|ds\Big)^p\Big] + \widetilde{C_2}\mathbb{E}\Big[\Big(\int_0^T |\sigma(s,\omega)|^2 ds\Big)^{\frac{p}{2}}\Big]\\
&+ \widetilde{C_3}\int_0^T \big[1+\|\Sigma(s,\cdot)\|_{L^\infty}\big]\mathbb{E}\Big[\|X_\theta\|_{\infty,s}^p\Big]ds.
\end{aligned}
$$

Applying Grönwall Inequality yields

$$\mathbb{E}\Big[\|X_\theta\|_\infty^p\Big] \le 5\Bigg(\mathbb{E}\Big[|\theta|^p\Big] + \widetilde{C_1}\mathbb{E}\Big[\Big(\int_0^T |b(s,\omega)|ds\Big)^p\Big] + \widetilde{C_2}\mathbb{E}\Big[\Big(\int_0^T |\sigma(s,\omega)|^2 ds\Big)^{\frac{p}{2}}\Big]\Bigg)$$
$$\times \exp\Big(5\widetilde{C_3}\int_0^T \big[1+\|\Sigma(s,\cdot)\|_{L^\infty}\big]ds\Big). \qquad \square$$

References

[1] Sandra Cerrai, *Second order PDE's in finite and infinite dimension*, Lecture Notes in Mathematics, vol. 1762, Springer-Verlag, Berlin, 2001, A probabilistic approach. MR-1840644

[2] Dan Crisan and Eamon McMurray, *Smoothing properties of McKean–Vlasov SDEs*, Probability Theory and Related Fields (2017), 1–52. MR-3800831

[3] Kenneth D. Elworthy, *Stochastic flows on Riemannian manifolds*, Diffusion Processes and Related Problems in Analysis, Volume II: Stochastic Flows (Mark A. Pinsky and Volker Wihstutz, eds.), Birkhäuser Boston, Boston, MA, 1992, pp. 37–72. MR-1187985

[4] Eric Fournié, Jean-Michel Lasry, Jérôme Lebuchoux, and Pierre-Louis Lions, *Applications of Malliavin calculus to Monte-Carlo methods in finance. II*, Finance and Stochastics **5** (2001), no. 2, 201–236. MR-1841717

[5] Eric Fournié, Jean-Michel Lasry, Jérôme Lebuchoux, Pierre-Louis Lions, and Nizar Touzi, *Applications of Malliavin calculus to Monte-Carlo methods in finance*, Finance and Stochastics **3** (1999), no. 4, 391–412. MR-1842285

[6] Christel Geiss and Alexander Steinicke, *Malliavin derivative of random functions and applications to Lévy driven BSDEs*, Electronic Journal of Probability **21** (2016). MR-3485352

[7] Martin Hutzenthaler, Arnulf Jentzen, and Peter E. Kloeden, *Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **467** (2011), no. 2130, 1563–1576. MR-2795791

[8] Peter Imkeller, Thibaut Mastrolia, Dylan Possamaï, and Anthony Réveillac, *A note on the Malliavin-Sobolev spaces*, Statist. Probab. Lett. **109** (2016), 45–53. MR-3434959

[9] Hiroshi Kunita, *Stochastic flows and stochastic differential equations*, Cambridge Studies in Advanced Mathematics, vol. 24, Cambridge University Press, Cambridge, 1990. MR-1070361

[10] Xuerong Mao, *Stochastic differential equations and applications*, Horwood, 2008. MR-2380366

[11] Thibaut Mastrolia, Dylan Possamaï, and Anthony Réveillac, *On the Malliavin differentiability of BSDEs*, Ann. Inst. Henri Poincaré Probab. Stat. **53** (2017), no. 1, 464–492. MR-3606749

[12] Charles Mode and Candace Sleeman, *Stochastic processes in genetics and evolution: Computer experiments in the quantification of mutation and selection*, World Scientific Publishing Company, 2012. MR-2952851

[13] David Nualart, *The Malliavin calculus and related topics*, vol. 1995, Springer, 2006. MR-1344217

[14] Bernt Øksendal, *Stochastic differential equations: An introduction with applications*, Hochschultext / Universitext, Springer, 2003. MR-2001996

[15] Sebastian Riedel and Michael Scheutzow, *Rough differential equations with unbounded drift term*, Journal of Differential Equations **262** (2017), no. 1, 283 – 312. MR-3567487

[16] Hiroshi Sugita, *On a characterization of the Sobolev spaces over an abstract Wiener space*, J. Math. Kyoto Univ. **25** (1985), no. 4, 717–725. MR-0810975

[17] Mahdieh Tahmasebi and Shiva Zamani, *Weak differentiability of solutions to SDEs with semi-monotone drifts.* Bulletin of the Iranian Mathematical Society 41.4 (2015): 873-888. MR-3385363

[18] Ali Suleyman Üstünel and Moshe Zakai, *Transformation of measure on Wiener space*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2000. MR-1736980

[19] Xicheng Zhang, *Stochastic differential equations with Sobolev diffusion and singular drift and applications*, Ann. Appl. Probab. **26** (2016), no. 5, 2697–2732. MR-3563191