On the Itô-Alekseev-Gröbner formula for stochastic differential equations

Anselm Hudde\textsuperscript{1}, Martin Hutzenthaler\textsuperscript{1}, Arnulf Jentzen\textsuperscript{2} & Sara Mazzonetto\textsuperscript{3}

\textsuperscript{1} Faculty of Mathematics, University of Duisburg-Essen, Germany
\textsuperscript{2} Faculty of Mathematics and Computer Science, University of Muenster, Germany
\textsuperscript{3} Faculté des sciences et Technologies, Université de Lorraine, France

May 14, 2021

Abstract

In this article we establish a new formula for the difference of a test function of the solution of a stochastic differential equation and of the test function of an Itô process. The introduced formula essentially generalizes both the classical Alekseev-Gröbner formula from the literature on deterministic differential equations as well as the classical Itô formula from stochastic analysis. The discovered formula, which we suggest to refer to as Itô-Alekseev-Gröbner formula, is a powerful tool for deriving strong approximation rates for perturbations and approximations of stochastic ordinary and partial differential equations.

1 Introduction

The linear integration-by-parts formula states in the simplest case that for all $a, b \in \mathbb{R}$, $t \in [0, \infty)$ it holds that

\[
e^{at} - e^{bt} = -\int_0^t \frac{d}{ds}(e^{a(t-s)}e^{bs}) \, ds = \int_0^t e^{a(t-s)}(a-b)e^{bs} \, ds.
\]

(1)

The nonlinear integration-by-parts formula, which is also referred to as \textit{Alekseev-Gröbner formula} or as nonlinear variation-of-constants formula, generalizes this relation to nonlinear ordinary differential equations and has been established in Alekseev \cite{1} and Gröbner \cite{17}. More formally, the Alekseev-Gröbner formula (cf., e.g., Hairer et al. \cite[Theorem I.14.5]{19}) asserts that for all $d \in \mathbb{N}$, $T \in (0, \infty)$, $\mu \in C^{0,1}([0, T] \times \mathbb{R}^d, \mathbb{R}^d)$, $Y \in C^1([0, T], \mathbb{R}^d)$, and all $X_r = (X^{x}_{s,t})_{s \in [0, t], t \in [0, T], x \in \mathbb{R}^d} \in C((s,t) \in [0, T]^2; s \leq t) \times \mathbb{R}^d, \mathbb{R}^d)$ with $\forall s \in [0, T], t \in [s, T]$, $x \in \mathbb{R}^d$:

\[
X^{Y_0}_{0,T} - Y_T = \int_0^T \left( \frac{\partial}{\partial x} X^{Y_r}_{r,T} \right) \left( \mu(r, Y_r) - \frac{d}{dr} Y_r \right) \, dr.
\]

(2)

Informally speaking, the Alekseev-Gröbner formula expresses the global error (the term $X^{Y_0}_{0,T} - Y_T$ in \cite{2}) in terms of the infinitesimal error (the term $\mu(r, Y_r) - \frac{d}{dr} Y_r$ in \cite{2}) which corresponds to the difference of time
derivatives). For this reason, the Alekseev-Gröbner formula is a powerful tool for studying perturbations of ordinary differential equations; see, e.g., Norsett & Wanner [47, Theorem 3], Lie & Norsett [38, Theorem 1], Iserles & Soederlind [27, Theorem 1], and Iserles [26, Theorem 3.7].

In this article we generalize the Alekseev-Gröbner formula to a stochastic setting and derive the nonlinear integration-by-parts formula for stochastic differential equations (SDEs). Informally speaking, one key difficulty in this generalization is that the integrand on the right-hand side of (2) (and a similar integrand appears in the stochastic integral in (3) below) depends both on the past (e.g. the term $\mu(r, Y_t)$) and on the future (e.g. the term $\partial_{x^r}Y_{r,r,t}$). This precludes a generalization which is solely based on Itô calculus. In this article we apply Malliavin calculus and express anticipating stochastic integrals as Skorohod integrals. The following theorem, Theorem 1.1 formulates our main contribution and establishes – what we call – the Itô-Alekseev-Gröbner formula. For its formulation and throughout this article we use the notation introduced in Subsection 1.1 below.

**Theorem 1.1** (Itô-Alekseev-Gröbner formula). Let $d, m, k \in \mathbb{N}$, $T, c \in (0, \infty)$, $p \in (4, \infty)$, $q \in [0, \frac{p}{2} - 2)$, $\xi \in \mathbb{R}^d$, $e_1 = (1, 0, \ldots, 0), \ldots, e_d = (0, \ldots, 0, 1) \in \mathbb{R}^d$, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $W : [0, T] \times \Omega \rightarrow \mathbb{R}^m$ be a standard Brownian motion with continuous sample paths, let $N = \{A \in \mathcal{F} : \mathbb{P}(A) = 0\}$, let $\mu \in C([0, T] \times \mathbb{R}^d, \mathbb{R}^d)$, $\sigma \in C([0, T] \times \mathbb{R}^d, \mathbb{R}^{d \times m})$, let $X_{r,T} = (X_{r,s,t})_{s \in [0,t], t \in [0,T], x \in \mathbb{R}^d}$, let $(s,t) \in [0,T]^2 : s \leq t \rightarrow \mathbb{R}^d \times \Omega \rightarrow \mathbb{R}^d$ be a continuous random field, assume that for all $s \in [0, T)$, $\omega \in \Omega$ it holds that $(\mathbb{R}^d \ni x \mapsto X_{r,T}^x(\omega) \in \mathbb{R}^d) \in C^2(\mathbb{R}^d, \mathbb{R}^d)$, assume that for all $\omega \in \Omega$ it holds that $\partial_{x^r}^2 X_{r,T}^x(\omega) \in C([0, T] \times \mathbb{R}^d, L^2(\mathbb{R}^d, \mathbb{R}^d))$, assume that for all $s \in [0, T)$, $x \in \mathbb{R}^d$ the stochastic process $[s, T] \times \Omega \ni (t, \omega) \mapsto X_{s,t}^x(\omega) \in \mathbb{R}^d$ is $(\mathcal{G}(N \cup \mathcal{G}(W_r - W_s : r \in [s, t])) \ni (t, \omega) \mapsto (s, t) \rightarrow \mathbb{R}^d$-adapted, assume that for all $s \in [0, T)$, $t \in [s, T)$, $x \in \mathbb{R}^d$ it holds $\mathbb{P}$-a.s. that

$$X_{s,t}^x = x + \int_s^t \mu(r, X_{s,r}^x) \, dr + \int_s^t \sigma(r, X_{s,r}^x) \, dW_r,$$

(3)

assume that for all $s, t \in [0, T)$, $x \in \mathbb{R}^d$ with $s \leq t$ it holds $\mathbb{P}$-a.s. that $X_{t,T}^{X_{s,t}^x} = X_{s,t}^x$, let $A, Y : [0, T] \times \Omega \rightarrow \mathbb{R}^d$, $B : [0, T] \times \Omega \rightarrow \mathbb{R}^{d \times m}$ be $(\mathcal{G}(N \cup \mathcal{G}(W_r : r \in [0, t]))) \ni (t, \omega) \rightarrow \mathbb{R}^d$-predictable stochastic processes, assume that $Y$ has continuous sample paths, assume that $\int_0^T \mathbb{E} \left[ \left\| A_s \right\|_{\mathbb{R}^d}^p + \left\| Y_s \right\|_{\mathbb{R}^d}^p + \left\| B_s \right\|_{L^p(\mathbb{R}^m, \mathbb{R}^d)}^p \right] \, ds < \infty$, assume that for all $t \in [0, T)$ it holds $\mathbb{P}$-a.s. that

$$Y_t = \xi + \int_0^t A_s \, ds + \int_0^t B_s \, dW_s,$$

(4)

assume that

$$\sup_{s,t \in [0,T]} \mathbb{E} \left[ \left\| \mu(\xi, Y_{s,t}^x) \right\|_{\mathbb{R}^d}^p + \left\| \sigma(\xi, Y_{s,t}^x) \right\|_{L^p(\mathbb{R}^m, \mathbb{R}^d)}^p \right] < \infty,$$

(5)

assume that

$$\sup_{r,s,t \in [0,T]} \mathbb{E} \left[ \left\| X_{r,T}^{X_{s,t}^x} \right\|_{\mathbb{R}^d}^p + \left\| \frac{\partial}{\partial x} X_{r,T}^{X_{s,t}^x} \right\|_{L^p(\mathbb{R}^d, \mathbb{R}^d)}^{2p(4 + 2)} + \left\| \frac{\partial^2}{\partial x^2} X_{r,T}^{X_{s,t}^x} \right\|_{L^p(\mathbb{R}^d, \mathbb{R}^d)}^{2p(4 + 2)} \right] < \infty,$$

(6)

and let $f \in C^2(\mathbb{R}^d, \mathbb{R}^k)$ satisfy that for all $x \in \mathbb{R}^d$ it holds that

$$\max \left\{ \left\| f(x) \right\|_{\mathbb{R}^k}, \left\| f'(x) \right\|_{L^2(\mathbb{R}^d, \mathbb{R}^k)}, \left\| f''(x) \right\|_{L^2(\mathbb{R}^d, \mathbb{R}^k)} \right\} \leq c(1 + \left\| x \right\|_{\mathbb{R}^d}^q),$$

(7)

Then the stochastic process $\left( f(X_{r,T}^{Y_{r,T}}) \frac{\partial}{\partial x} X_{r,T}^{Y_{r,T}}(\sigma(r, Y_r) - B_r) \right)_{r \in [0, T]}$ is Skorohod integrable and it holds $\mathbb{P}$-a.s. that

$$f(X_{0,T}^{Y_0}) - f(Y_T) = \int_0^T f'(X_{r,T}^{Y_r}) \frac{\partial}{\partial x} X_{r,T}^{Y_r} \left( \mu(r, Y_r) - A_r \right) \, dr + \int_0^T f'(X_{r,T}^{Y_r}) \frac{\partial}{\partial x} X_{r,T}^{Y_r} \left( \sigma(r, Y_r) - B_r \right) \, dW_r$$

$$+ \frac{1}{2} \sum_{i,j=1}^d \int_0^T \mathbb{E} (\sigma(r, Y_r)[\sigma(r, Y_r)] - B_r[B_r]) \left( f''(X_{r,T}^{Y_r}) \left( \frac{\partial}{\partial x} X_{r,T}^{Y_r}, \frac{\partial}{\partial x} X_{r,T}^{Y_r} \right) + f'(X_{r,T}^{Y_r}) \frac{\partial^2}{\partial x^2} X_{r,T}^{Y_r} \right) \left( e_i^{(d)}, e_j^{(d)} \right) \, dr,$$

(8)
Theorem 1.1 follows immediately from Theorem 3.1 (applied with $F_0 = \mathcal{S}(\mathcal{N})$, $O = \mathbb{R}^d$ in the notation of Theorem 3.1). Theorem 1.1 essentially generalizes the following results from the literature:

(i) Theorem 1.1 essentially generalizes the Alekseev-Gröbner formula. More formally, Theorem 1.1 (applied with $\sigma = 0$, $B = 0$, $k = d$, $= \text{Id}_{\mathbb{R}^d}$ in the notation of Theorem 1.1) implies the Alekseev-Gröbner formula in (2) (cf., e.g., Hairer et al. [19, Theorem I.14.5]) in the case where the solution process is twice continuously differentiable in the space variable.

(ii) Theorem 1.1 essentially generalizes the Itô formula. More formally, Theorem 1.1 (applied with $\mu = 0$, $\sigma = 0$ in the notation of Theorem 1.1) implies the Itô formula for Itô processes (cf., e.g., Revuz & Yor [51, Theorem IV.3.3]) in the case where the Itô process $Y$, its drift process $A$, and its diffusion process $B$ satisfy $\inf_{p \in (4, \infty)} \left( \sup_{s \in [0, T]} \mathbb{E}\left[ \|Y_s\|_{\mathbb{R}^d}^p \right] + \int_0^T \mathbb{E}\left[ \|A_s\|_{\mathbb{R}^d}^p + \|B_s\|_{L(\mathbb{R}^m, \mathbb{R}^d)}^p \right] ds \right) < \infty$. This moment requirement is due to the fact that we use the Skorohod integral. An approach with rough path integrals (cf., e.g., Friz & Hairer [15]) might be suitable to generalize Theorem 1.1 so that this moment condition would not be needed.

(iii) Theorem 1.1 essentially generalizes the Alekseev-Gröbner formula in (2) (cf., e.g., Hairer et al. [19, Theorem I.14.5]) even in the deterministic case ($\sigma = 0$ and $B = 0$ in the notation of Theorem 1.1) from $f = \text{Id}_{\mathbb{R}^d}$ to general test functions. In Proposition 2.1 below we prove the Itô-Alekseev-Gröbner formula in (8) in the deterministic case with the test function $f : \mathbb{R}^d \to \mathbb{R}^k$ being only in $C^1(\mathbb{R}^d, \mathbb{R}^k)$ instead of in $C^2(\mathbb{R}^d, \mathbb{R}^k)$ as in Theorem 1.1 above. The proof of Proposition 2.1 below is also illustrative to understand the structure of the Itô-Alekseev-Gröbner formula in (8).

(iv) Theorem 1.1 essentially provides a pathwise version of the well-known weak error expansion (cf., e.g., Graham & Talay [16] (7.48) and the last Display on page 182) or related weak error estimates in [53, 13, 54]. More precisely, in the notation of Theorem 1.1 taking expectation of (8), using that the expectation of the Skorohod integral vanishes, and exchanging expectations and temporal integrals results in the standard representation of the weak error $\mathbb{E}\left[ f(X_{1_0,T}^0) \right] - \mathbb{E}\left[ f(Y_{T}) \right]$. We note that (8) roughly follows from the case $k = d$, $f = \text{Id}_{\mathbb{R}^d}$ of Theorem 1.1 and from the Itô formula for anticipating processes established in Alòs & Nualart [2] applied to the anticipating process $[0, T] \ni t \mapsto X_{1_0,T}^0 \in \mathbb{R}^d$. Moreover, using the two-sided stochastic integral of Pardoux & Protter [50], Nualart & Pardoux [49, Proposition 8.2] establish a backward Itô-Ventzell-type formula where the random test function roughly speaking has the form $F(t, x, \omega) = f(t, x, Y_t(\omega))$ where $f$ is deterministic and $Y$ is a semimartingale. This result is not applicable to the situation of Theorem 1.1 since $x \mapsto X^x$ is a random function. However, our proof of Theorem 1.1 uses ideas of the proof of Nualart & Pardoux [49, Proposition 8.2]. Moreover, after the preprint [21] of our paper appeared, Del Moral & Singh [14] establish a backward Itô-Ventzell formula and use this to provide a new proof of Theorem 1.1 in the special case of coefficient functions which have continuous and uniformly bounded spatial derivatives up to third order. In addition, independently of our results (cf. [3] with our preprint [21]), Arnaudon & Del Moral [3] (3.2)] arrive at the Itô-Alekseev-Gröbner formula in a specific situation (e.g. the diffusion terms are equal) by heuristically applying the backward Itô-Ventzell formula which was later established in [14].

Theorem 1.1 implies immediately an $L^2$-estimate. For example the $L^2$-norm of the right-hand side of (8) can be bounded by the triangle inequality. The $L^2$-norm of the Skorohod integral on the right-hand side of (8) can then be calculated by applying the Itô isometry for Skorohod integrals (see, e.g., Alòs & Nualart [2, Lemma 4]). Another approach for obtaining $L^2$-estimates is to apply the Itô formula for Skorohod processes to the squared norm of the right-hand side of (8). However this seems to require additional regularity.

Our main motivation for the Itô-Alekseev-Gröbner formula are strong convergence rates for time-discrete numerical approximations of stochastic evolution equations (SEEs). In the literature, positive strong convergence rates have been established for SEE with monotone nonlinearities (see, e.g., [39, Chapter 4]); see, e.g.,
for the case of additive noise and for lower bounds see, e.g., [12 43 44 45 5]. Recently, the classical 
Alekseev-Gröbner formula has been applied in [25] to establish strong convergence rates for space-time 
discrete approximations for stochastic Burgers equations with additive noise; for lower bounds see, e.g., [12 43 44 45 5]. 
This demonstrates that the Alekseev-Gröbner formula is a successful approach for proving convergence rates in 
the case of SEEs such as the stochastic Burgers equation with additive noise. Now the Itô-Alekseev-Gröbner 
formula in Theorem 1.1 provides an approach to derive strong convergence rates e.g. for stochastic Burgers 
equations also in the case of non-additive noise. Applications of this approach are left to future research.

In addition, Theorem 1.1 can be applied to any approximation of an SDE which is an Itô process with respect 
to the same Wiener process driving the SDE. Possible applications (cf., e.g., [23]) include, in the notation of 
Theorem 1.1

(i) strong convergence rates for time-discrete numerical approximations of SDEs (e.g., the Euler-Maruyama 
approximation with $N \in \mathbb{N}$ time discretization steps is given by $A_t = \mu(kT, Y_{kT})$ and $B_t = \sigma(kT, Y_{kT})$ for 
all $t \in \left[ \frac{kT}{N}, \frac{(k+1)T}{N} \right)$, $k \in \mathbb{N}_0 \cap [0, N]$),

(ii) strong convergence rates for Galerkin approximations for SEEs (see, e.g., [9]) (choose $A_t = P(\mu(t, Y_t))$ 
and $B_t u = P(\sigma(t, Y_t) u)$ for all $u \in \mathbb{R}^m$, $t \in [0, T]$ and some suitable projection operator $P \in L(\mathbb{R}^d)$ where 
d, m $\in \mathbb{N}$; Theorem 1.1 is applied to a finite-dimensional approximation of the exact solution of the SEE 
of which convergence in probability is known), and

(iii) strong convergence rates for small noise perturbations of solutions of deterministic differential equations 
(choose $\sigma = 0$, $A_t = \mu(t, Y_t)$ and $B_t = \varepsilon \tilde{\sigma}(t, Y_t)$ for all $t \in [0, T]$ where $\tilde{\sigma} : [0, T] \times \mathbb{R}^d \to \mathbb{R}^{d \times m}$ is a suitable 
Borel measurable function and where $\varepsilon > 0$ is a sufficiently small parameter).

In the literature, nearly all estimates of perturbation errors exploit the popular global monotonicity assumption 
which, in the notation of Theorem 1.1 assumes existence of a real number $c \in \mathbb{R}$ such that for all $x, y \in \mathbb{R}^d$, 
t $\in [0, T]$ it holds that

$$
\langle x - y, \mu(t, x) - \mu(t, y) \rangle_{\mathbb{R}^d} + \frac{1}{2} \| \sigma(t, x) - \sigma(t, y) \|^2_{L(\mathbb{R}^m, \mathbb{R}^d)} \leq c \| x - y \|^2_{\mathbb{R}^d};
$$

(9)
cf. also [23] and the references therein. We emphasize that many SDEs from the literature do not satisfy [9] 
and that Theorem 1.1 does not require that the global monotonicity assumption is fulfilled.

A crucial assumption in Theorem 1.1 is existence of a solution of the SDE [9] which is twice continuously 
differentiable in the starting point since in the proof of Theorem 1.1 we apply Itô’s formula for independent 
random fields to the random functions $\mathbb{R}^d \ni x \mapsto X^x_{t,T} \in \mathbb{R}^d$, $t \in [0, T]$. This assumption is not satisfied in 
a number of cases. For example Li & Scheutzow [37] construct a two-dimensional example with smooth and 
globally bounded coefficient functions which is not even strongly complete (that is, the exceptional subset of 
$\Omega$ where [3] fails to hold can not be chosen independently of the starting point); cf. also Hairer et al. [20] 
Theorem 1.2. Under suitable assumptions on the coefficients, however, strong completeness and existence of a 
solution of [3] which is continuous in the starting point can be ensured; see, e.g., [10 56 36]. Existence of a 
solution of [3] which satisfies the assumptions of Theorem 1.1 is currently known essentially only in the case 
of twice continuously differentiable coefficient functions whose derivatives up to second order are bounded; see, 
e.g., in Kunita [54 Theorem 1.4.1]. In future research we will generalize this to unbounded twice continuously 
differentiable coefficient functions which satisfy certain growth conditions at infinity; see [22]. Moreover, in 
the case of non-differentiable coefficients a possible approach is to approximate the SDE by SDEs with smooth 
coefficients and to apply Theorem 1.1 to the sequence of smoothened SDE solutions.

We prove Theorem 1.1 as follows. First, we rewrite the left-hand side of equation (8) as telescoping sum; see [21] 
below. Then we apply Itô’s formula to the random functions $\mathbb{R}^d \ni x \mapsto X^x_{t,T} \in \mathbb{R}^d$, $t \in [0, T]$ in
order to expand the local errors. Thereby we obtain Itô integrals which we rewrite as Skorohod integrals by applying Proposition A.8 below. These Skorohod integrals are non-standard since the integrands are in general not measurable with respect to a Wiener process. For this reason we introduce an extended Skorohod integral in the appendix. Moreover, the integrands in the Itô integrals are adapted to different filtrations. We apply Proposition A.7 below in order to carefully rewrite the sum of these integrals as a single Skorohod integral.

1.1 Notation

The following notation is used throughout this article. We denote by \( \mathbb{N} \) and by \( \mathbb{N}_0 \) the sets satisfying that \( \mathbb{N} = \{1, 2, 3, \ldots\} \) and \( \mathbb{N}_0 = \mathbb{N} \cup \{0\} \). For all \( c \in (0, \infty) \) let \( 0^0, \frac{0}{0}, \frac{\infty}{\infty}, \frac{c}{0}, 0 \cdot \infty, 0 \cdot (-\infty), \infty^c \) denote the extended real numbers \( 0^0 = 1, \frac{0}{0} = 0, \frac{\infty}{\infty} = \infty, \frac{c}{0} = -\infty, 0 \cdot \infty = 0, 0 \cdot (-\infty) = 0, \) and \( \infty^c = \infty \). For all \( T \in (0, \infty) \) let \( \Delta_T \subseteq [0, T]^2 \) denote the subset with the property that \( \Delta_T = \{(s, t) \in [0, T]^2 : s \leq t\} \) and denote by \( \mathcal{T} \) the set \( \mathcal{T} = \{ \mathcal{T}_n : n \in \mathbb{N}\} \). For all \( h \in (0, \infty), r \in [0, \infty) \) let \( [r]_h, [r]_h^0, [r]_0, [r]_0^0 \) be the real numbers with the properties that \( [r]_h = \inf\{nh \in [r, \infty) : n \in \mathbb{N}_0\}, \quad [r]_h^0 = \sup\{nh \in [0, r) : n \in \mathbb{N}_0\}, \), \( [r]_0 = r, \) and \( [r]_0^0 = 0 \). For a real vector space \( V \) and a subset \( S \subseteq V \) let \( \text{span}(S) \subseteq V \) denote the set with the property that \( \text{span}(S) = \left\{ \sum_{i=1}^{n} r_i v_i : n \in \mathbb{N}, r_1, \ldots, r_n \in \mathbb{R}, v_1, \ldots, v_n \in V \right\} \). For all \( (s, t) \in \Delta_T \) let \( \lambda_{[s,t]} \) be the Lebesgue-measure restricted to the Borel-sigma-algebra of \([s, t]\). For all \( d \in \mathbb{N}, x \in \mathbb{R}^d \) we write \( \|x\|_{\mathbb{R}^d} \) for the Euclidean norm of \( x \) and for all \( i \in \{1, \ldots, d\} \) let \( e_i^{(d)} \) denote the \( i \)-th unit vector in \( \mathbb{R}^d \). For every set \( \Omega \) we denote by \( \mathcal{G}(\mathcal{E}) \) the smallest \( \sigma \)-algebra generated by \( \mathcal{E} \subseteq \mathcal{P}(\Omega) \). For all measurable spaces \((\Omega, \mathcal{F}), (\Omega', \mathcal{B})\) let \( \mathcal{M}(\mathcal{F}, \mathcal{B}) \) be the set \( \mathcal{M}(\mathcal{F}, \mathcal{B}) = \{f : \Omega \to \Omega' : f \text{ is } \mathcal{F}/\mathcal{B}-\text{measurable}\} \). For every measurable space \((\Omega, \mathcal{F}, \mu), \mu\), every normed vector space \((V, \|\cdot\|_V)\), and all \( p \in [1, \infty) \) let \( \mathcal{L}^p(\mu) \) denote the Borel-sigma-algebra on \( V \), let \( \mathcal{L}^p(\mu, V) \) be the set with the property that \( \mathcal{L}^p(\mu, V) = \{f \in \mathcal{M}(\mathcal{F}, \mathcal{B}(V)) : \int_{\Omega} \|f\|_V^p \, d\mu < \infty\} \), let \( \mathcal{L}^p(\mu) \) be the set with the property that \( \mathcal{L}^p(\mu, V) = \{f \in \mathcal{L}^p(\mu, V) : g = f \text{ } \mu\text{-a.e.: } g \in \mathcal{L}^p(\mu)\} \), and let

\[
\|\cdot\|_{L^p(\mu, V)} : \mathcal{M}(\mathcal{F}, \mathcal{B}(V)) \cup \{\{h \in \mathcal{M}(\mathcal{F}, \mathcal{B}) : h = g \text{ } \mu\text{-a.e.: } g \in \mathcal{M}(\mathcal{F}, \mathcal{B})\}) \to [0, \infty]
\]

be the function which satisfies for all \( f \in \mathcal{M}(\mathcal{F}, \mathcal{B}(V)) \cup \{\{h \in \mathcal{M}(\mathcal{F}, \mathcal{B}) : h = g \text{ } \mu\text{-a.e.: } g \in \mathcal{M}(\mathcal{F}, \mathcal{B})\}) \) that

\[
\|f\|_{L^p(\mu, V)} = \left( \int_{\Omega} \|f\|_V^p \, d\mu \right)^{\frac{1}{p}}. \]

For all \( d, m \in \mathbb{N} \) and all \( A \in \mathbb{R}^{d \times m} \) we denote by \( A^* \) the transpose of \( A \). For every measurable space \((\Omega, \mathcal{F})\) and every \( n \in \mathbb{N} \) let \( C^\infty_{\mathcal{F}}(\mathbb{R}^n \times \Omega, \mathbb{R}) \) be the set which satisfies that

\[
C^\infty_{\mathcal{F}}(\mathbb{R}^n \times \Omega, \mathbb{R}) = \left\{ f : \mathbb{R}^n \times \Omega \to \mathbb{R} : \forall \omega \in \Omega: f(\cdot, \omega) \in C^\infty(\mathbb{R}^n, \mathbb{R}), \forall x \in \mathbb{R}^d : f(x, \cdot) \text{ is } \mathcal{F}/\mathcal{B}(\mathbb{R})\text{-measurable} \right\}.
\]

For all \( d, k \in \mathbb{N} \) we denote by \( L^{(2)}(\mathbb{R}^d, \mathbb{R}^k) \) the set of bilinear functions from \((\mathbb{R}^d)^2\) to \( \mathbb{R}^k \).

2 The Itô-Alekseev-Gröbner formula in the deterministic case

The following proposition, Proposition 2.1, generalizes the Alekseev-Gröbner formula (cf., e.g., Hairer et al. [19, Theorem I.14.5]) (which is the special case \( k = d \), \( f = \text{Id}_{\mathbb{R}^d} \) of Proposition 2.1) to general test functions.

Proposition 2.1 (Deterministic Itô-Alekseev-Gröbner formula). Let \( d, k, \in \mathbb{N}, T \in (0, \infty), \) \( O \subseteq \mathbb{R}^d \) be a non-empty open set, let \( \mu \in C^{0,1}([0, T] \times O, \mathbb{R}^d), Y \in C^1([0, T], O), X_{\cdot} = (X^{x}_{\cdot})_{x \in [0, T], t \in \mathbb{R}, \xi \in O} \subseteq C^1([0, T]^2 : s \leq t) \times O, f \in C^1(O, \mathbb{R}^k), \) and assume for all \( s \in [0, T], t \in [s, T], x \in O \) that \( X^{x}_{s,t} = x + \int_{s \wedge T}^{t \wedge T} \mu(r, X^{x}_{s,r}) \, dr \). Then

\[
f(X^{Y^{0}_{0,T}}_{0,T} - f(Y_T)) = \int_{0}^{T} f'(X^{Y^{0}_{0,T}}_{s,T}) \frac{\partial}{\partial x} X^{Y^{0}_{s,T}}_{s,T} \left( \mu(s, Y_s) - \frac{d}{ds} Y_s \right) \, ds. \tag{12}
\]
Proof of Proposition 2.1 The assumptions and the fundamental theorem of calculus imply for all $s \in [0,T)$, $t \in [s,T]$, $x \in O$ that $([s,T] \ni u \mapsto X^x_{s,u} \in O) \in C^1(O,O)$ and that $\frac{\partial}{\partial t} X^x_{s,t} = \mu(t, X^x_{s,t})$. This, the assumptions, and Hairer et al. [19] Theorem I.14.3) prove that for all $s \in [0,T)$, $t \in [s,T]$ it holds that $(O \ni x \mapsto X^x_{s,t} \in O) \in C^1(O,O)$ and that $\frac{\partial}{\partial t} X^x_{s,t} \in C([0,T] \times O, L(\mathbb{R}^d, \mathbb{R}^d))$. Moreover, the assumptions, and Hairer et al. [19] Theorem I.14.4] show that for all $x \in O$ it holds that $([0,T] \ni s \mapsto X^x_{s,T} \in O) \in C^1([0,T],O)$, that $\frac{\partial}{\partial t} X^x_{s,T} = - \frac{\partial}{\partial x} X^x_{s,T}\mu(s,x)$. Therefore, the chain rule implies that $([0,T] \ni s \mapsto X^x_{s,T} \in O) \in C^1([0,T],O)$. Moreover, the fundamental theorem of calculus, the chain rule, and (13) yield that

$$f(X^x_{0,T}) - f(Y_T) = -\int_0^T \frac{d}{ds}f(X^x_{s,T}) \, ds$$

$$= -\int_0^T \frac{d}{ds}f'(X^x_{s,T}) \left(\frac{\partial}{\partial s}X^x_{s,T}\right)_{|s=Y_s} + \frac{\partial}{\partial s}X^x_{s,T} \frac{d}{ds}Y_s \, ds$$

$$= -\int_0^T f'(X^x_{s,T}) \left(- \frac{\partial}{\partial s}X^x_{s,T}\mu(s,Y_s) + \frac{\partial}{\partial s}X^x_{s,T} \frac{d}{ds}Y_s\right) \, ds$$

$$= \int_0^T f'(X^x_{s,T}) \frac{\partial}{\partial x}X^x_{s,T} \left(\mu(s,Y_s) - \frac{d}{ds}Y_s\right) \, ds.$$ 

This finishes the proof of Proposition 2.1.

\[\square\]

3 The Itô-Aleksiev-Görner formula in the general case

The following theorem, Theorem 3.1 is the main result of this article. We note that throughout this article we use notation introduced in Subsection 1.1 in the Appendix.

**Theorem 3.1** (Itô-Aleksiev-Görner formula). Let $d,m,k \in \mathbb{N}$, $T,c \in (0,\infty)$, $p \in (4,\infty)$, $q \in [0,\frac{p}{2} - 2)$, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $W : [0,T] \times \Omega \to \mathbb{R}^m$ be a standard Brownian motion, let $\mathcal{N} = \{A \in \mathcal{F} : \mathbb{P}(A) = 0\}$, let $\mathbb{F} = (\mathbb{F}_t)_{t \in [0,T]}$ be a filtration on $(\Omega, \mathcal{F})$ which satisfies that $\mathbb{F}_0$ and $\mathbb{G}(W_s : s \in [0,T])$ are independent and which satisfies for all $t \in [0,T]$ that $\mathbb{F}_t = \mathbb{G}(\mathbb{F}_0 \cup \mathbb{G}(W_s : s \in [0,t]) \cup \mathcal{N})$, let $O \subseteq \mathbb{R}^d$ be a non-empty open set, let $\mu : [0,T] \times O \to \mathbb{R}^d$, $\sigma : [0,T] \times O \to \mathbb{R}^{d \times m}$ be continuous functions, let $X^\omega_\cdot : \Delta_T \times O \times \Omega \to O$, $X^\omega_{1/\cdot} : [0,T] \times O \times \Omega \to L(\mathbb{R}^d, \mathbb{R}^d)$, and $X^\omega_{2/\cdot} : [0,T] \times O \times \Omega \to L^2(\mathbb{R}^d, \mathbb{R}^d)$ be continuous random fields, assume that for all $s \in [0,T)$, $\omega \in \Omega$ it holds that $(O \ni x \mapsto X^\omega_{x/\cdot}(\omega) \in O) \in C^2(O,O)$, assume that for all $s \in [0,T]$, $x \in O$ the stochastic process $[s,T] \times \Omega \ni (t, \omega) \mapsto X^x_{s,t}(\omega) \in O$ is $(\mathbb{F}_t)_{t \in [0,T]}$-adapted, assume that for all $s \in [0,T]$, $t \in [s,T]$, $x \in O$ it holds $\mathbb{P}$-a.s. that

$$X^x_{s,t} = x + \int_s^t \mu(r, X^x_{s,r}) \, dr + \int_s^t \sigma(r, X^x_{s,r}) \, dW_r,$$

(15)

assume that for all $(s,t) \in \Delta_T$, $x \in O$ it holds $\mathbb{P}$-a.s. that $X^x_{t,T} = X^x_{s,T}$, assume that for all $(s,x,\omega) \in [0,T] \times O \times \Omega$ it holds that $X^1_{1/\cdot}(\omega) = \frac{\partial}{\partial t}(X^x_{1/\cdot}(\omega))$ and $X^2_{2/\cdot}(\omega) = \frac{\partial^2}{\partial t^2}(X^x_{2/\cdot}(\omega))$, let $Y \in L^p(\mathbb{L}^1_{[0,T]} \otimes \mathbb{P} ; \mathbb{P})$, $A \in L^p(\mathbb{L}^1_{[0,T]} \otimes \mathbb{P} ; \mathbb{R}^{d \times m})$, $B \in L^p(\mathbb{L}^1_{[0,T]} \otimes \mathbb{P} ; \mathbb{R}^{d})$ be stochastic processes, assume that $Y$ has continuous sample paths, assume that $Y$ and $B$ are $\mathbb{F}$-predictable, assume that for all $t \in [0,T]$ it holds $\mathbb{P}$-a.s. that

$$Y_t = Y_0 + \int_0^t A_s \, ds + \int_0^t B_s \, dW_s,$$

(16)
assume that
\[
\sup_{h \in T_n} \mathbb{E} \left[ \int_0^T \left\| \mu(t, X^{Y_{ih,T}}_t) \right\|^p_{\mathbb{R}^d} + \left\| \sigma(t, X^{Y_{ih,T}}_t) \right\|^p_{\text{HS}(\mathbb{R}^m, \mathbb{R}^d)} \, dt \right] < \infty, \tag{17}
\]
assume that
\[
\sup_{r,s,t \in [0,T]} \mathbb{E} \left[ \left\| X^{Y_{rs,T}}_t \right\|^p_{\mathbb{R}^d} + \left\| X^{Y_{rs,T}}_t \right\|^p_{\mathbb{R}^d} \right] < \infty, \tag{18}
\]
and let \( f \in C^2(O, \mathbb{R}^k) \) satisfy that for all \( x \in O \) it holds that
\[
\max \left\{ \frac{\| f(x) \|_{L^1[\mathbb{R}^d, \mathbb{R}^k])}}{1 + \| x \|_{\mathbb{R}^d}}, \frac{\| f''(x) \|_{L^1[\mathbb{R}^d, \mathbb{R}^k])}}{1 + \| x \|_{\mathbb{R}^d}} \right\} \leq c(1 + \| x \|_{\mathbb{R}^d}). \tag{19}
\]
Then the stochastic process \( \left( f'(X^{Y_{r,T}}_{r,T})X^{Y_{r,T}}_{r,T}(\sigma(r, Y_r) - B_r) \right)_{r \in [0,T]} \) is Skorohod-integrable and it holds \( \mathbb{P}\text{-a.s. that} \)
\[
f(X^Y_{0,T}) - f(Y_T) = \int_0^T f'(X^Y_{r,T})X^{Y_{r,T}}_{r,T} \left( \mu(r, Y_r) - A_r \right) \, dr + \int_0^T f'(X^Y_{r,T})X^{Y_{r,T}}_{r,T} \left( \sigma(r, Y_r) - B_r \right) \, \delta W_r \\
+ \frac{1}{2} \sum_{i,j=1}^d \int_0^T \left( \sigma(r, Y_r)[\sigma(r, Y_r)]^* - B_r[B_r]^* \right)_{i,j} \left( f''(X^Y_{r,T})X^{Y_{r,T}}_{r,T}, X^{Y_{r,T}}_{r,T} \right) + f'(X^Y_{r,T})X^{2Y_{r,T}}_{r,T} \left( e^{(d)}_i, e^{(d)}_j \right) \, dr. \tag{20}
\]

Proof of Theorem 3.4 The fact that for all \( \omega \in \Omega \) the function \( O \ni x \mapsto X^{x}_{x,T}(\omega) \in O \) is continuous and equation (15) imply that it holds \( \mathbb{P}\text{-a.s. that} \) \( X^{Y_{r,T}}_{T,T} = Y_T \). Moreover, we rewrite the left-hand side of equation (20) as telescoping sum and obtain that for all \( n \in \mathbb{N}, h \in \left\{ \frac{T}{n} \right\} \) it holds \( \mathbb{P}\text{-a.s. that} \)
\[
f(X^Y_{0,T}) - f(Y_T) = f(X^Y_{0,h,T}) - f(X^Y_{h,h,T}) - \sum_{i=0}^{n-1} \left( f(X^Y_{i,h,T}) - f(X^Y_{(i+1),h,T}) \right) \\
- \sum_{i=0}^{n-1} \left( f(X^Y_{(i+1),h,T}) - f(X^Y_{(i+1),h,T}) \right). \tag{21}
\]

First, we analyze the second sum on the right-hand side of equation (21). For all \( t \in [0,T], x \in O, i \in \{1, 2, \ldots, n \} \) the functions \( \Omega \ni \omega \mapsto X^{x}_{t,T}(\omega) \in O, \Omega \ni \omega \mapsto X^{i,x}_{t,T}(\omega) \in L^1(\mathbb{R}^d, \mathbb{R}^d) \) are \( \mathcal{G}(\mathcal{N}) \cup \mathcal{G}(W_s - W_t : s \in [t, T]) \)-measurable. This together with the fact that for all \( \omega \in \Omega, t \in [0,T] \) it holds that \( \left( O \ni x \mapsto f(X^x_{t,T}(\omega)) \right) \in \mathbb{R}^k \) \( \in C^2(O, \mathbb{R}^k) \) implies that for all \( t \in [0,T] \) the function \( \Omega \ni \omega \mapsto \left( O \ni x \mapsto f(X^x_{t,T}(\omega)) \right) \in \mathbb{R}^k \) \( \in C^2(O, \mathbb{R}^k) \) is independent of the sigma-algebra \( \mathcal{F}_t \). Itô’s formula for independent random fields (e.g., Klenke [31 Theorem 25.30 and Remark 25.26]) (applied with the functions \( \Omega \ni \omega \mapsto \left( O \ni x \mapsto f(X^x_{t,T}(\omega)) \right) \in \mathbb{R}^k \) \( \in C^2(O, \mathbb{R}^k) \) for \( n \in \mathbb{N}, i \in \{0, 1, \ldots, n-1\}, h \in \left\{ \frac{T}{n} \right\} \) yields that for all \( n \in \mathbb{N}, i \in \{0, 1, \ldots, n-1\}, h \in \left\{ \frac{T}{n} \right\} \) it holds \( \mathbb{P}\text{-a.s. that} \)
\[
f(X^Y_{(i+1),h,T}) - f(X^Y_{(i+1),h,T}) \\
= \int_{ih}^{(i+1)h} \frac{\partial}{\partial x} \left( f(X^Y_{(i+1),h,T}) \right) \bigg|_{x=Y_r} \, dY_r + \frac{1}{2} \sum_{l,j=1}^{(i+1)h} \int_{ih}^{(i+1)h} \frac{\partial^2}{\partial x^2} \left( f(X^Y_{(i+1),h,T}) \right) \bigg|_{x=Y_r} \left( e^{(d)}_l, e^{(d)}_j \right) \, d((Y)_r)_{l,j} \\
= \int_{ih}^{(i+1)h} f'(X^Y_{(i+1),h,T})X^{Y_{(i+1),h,T}}_{(i+1),h,T} \, dA_r + \int_{ih}^{(i+1)h} f'(X^Y_{(i+1),h,T})X^{Y_{(i+1),h,T}}_{(i+1),h,T} \, dB_r \, dW_r \\
+ \frac{1}{2} \sum_{l,j=1}^{(i+1)h} \int_{ih}^{(i+1)h} \left( B_r[B_r]^* \right)_{l,j} \left( f''(X^Y_{(i+1),h,T})X^{Y_{(i+1),h,T}}_{(i+1),h,T}, X^{Y_{(i+1),h,T}}_{(i+1),h,T} + f'(X^Y_{(i+1),h,T})X^{2Y_{(i+1),h,T}}_{(i+1),h,T} \left( e^{(d)}_l, e^{(d)}_j \right) \, dr. \tag{22}
\]
Inequalities (18) and (23) imply for all \( i \in \{1, 2\} \) that
\[
\sup_{r,s,t \in [0,T]} \left\| f^{(i)}(X_{t,T}^r) \right\|_{L^p(P: L^p([0,T]; \mathbb{R}^k))} \leq c \sup_{r,s,t \in [0,T]} \left\| X_{t,T}^{Y_{r,s}} \right\|_{L^p(P: \mathbb{R}^d)}^q \leq c \left( 1 + \sup_{r,s,t \in [0,T]} \left\| X_{t,T}^{Y_{r,s}} \right\|_{L^p(P: \mathbb{R}^d)}^q \right) < \infty.
\]

Hölder's inequality, inequalities (18), (23), and the assumption \( B \in \mathcal{C}^\infty([0,T] \times \mathbb{R}^d \times \mathbb{R}^d) \) imply that for all \( n \in \mathbb{N}, i \in \{0,1,\ldots,n-1\}, h \in \left\{ \frac{T}{n} \right\} \) it holds that
\[
\sup_{r,s \in \Delta_T} \left\| f'(X_{r,s}^{Y_{r,s}}) X_{r,s}^{Y_{r,s}} \right\|_{L^p(P: \mathbb{R}^d)} \leq \left( \sup_{r,s \in \Delta_T} \left\| X_{r,s}^{Y_{r,s}} \right\|_{L^p(P: \mathbb{R}^d)} \right) \left( \sup_{r,s \in \Delta_T} \left\| f'(X_{r,s}^{Y_{r,s}}) \right\|_{L^p(P: \mathbb{R}^d)} \right) < \infty.
\]

For all \( n \in \mathbb{N}, i \in \{0,1,\ldots,n-1\}, h \in \left\{ \frac{T}{n} \right\} \) the stochastic process \( f'(X_{(i+1)h,T}^{Y_{(i+1)h}}) X_{(i+1)h,T}^{Y_{(i+1)h} B_r} \) is predictable with respect to the filtration
\[
\mathcal{F}_r \cup \mathcal{G}(\{W_s - W_{(i+1)h} : s \in [(i+1)h,T]\}) \cup \{W_s - W_{(i+1)h} : s \in [(i+1)h,T]\}.
\]

Proposition A.8 together with inequality (24), Proposition A.7 and linearity of the Skorohod integral yield that for all \( h \in T/\mathbb{N} \) it holds that \( f'(X_{(i+1)h,T}^{Y_{(i+1)h}}) X_{(i+1)h,T}^{Y_{(i+1)h} B_r} \) is Skorohod-integrable and that for all \( n \in \mathbb{N}, h \in \left\{ \frac{T}{n} \right\} \) it holds \( \mathbb{P}\)-a.s. that
\[
\sum_{i=0}^{n-1} \int_{(i+1)h}^{(i+1)h} f'(X_{r,T}^{Y_{r}}) X_{r,T}^{Y_{r}} \, dB_r \, dW_r
\]
\[
= \sum_{i=0}^{n-1} \int_{i}^{i+1} f'(X_{r,T}^{Y_{r}}) X_{r,T}^{Y_{r}} \, dB_r \, dW_r
\]
\[
= \sum_{i=0}^{n-1} \int_{(i+1)h}^{(i+1)h} \mathbb{1}_{[(i+1)h,T]}(r) f'(X_{r,T}^{Y_{r}}) X_{r,T}^{Y_{r}} \, dB_r \, dW_r
\]
\[
= \int_{0}^{T} f'(X_{r,T}^{Y_{r}}) X_{r,T}^{Y_{r}} \, dB_r \, dW_r.
\]

Equations (22) and (26) imply that for all \( n \in \mathbb{N}, h \in \left\{ \frac{T}{n} \right\} \) it holds \( \mathbb{P}\)-a.s. that
\[
\sum_{i=0}^{n-1} \left( f(X_{(i+1)h,T}^{Y_{(i+1)h}}) - f(X_{ih,T}^{Y_{ih}}) \right)
\]
\[
= \int_{0}^{T} f'(X_{r,T}^{Y_{r}}) X_{r,T}^{Y_{r}} \, dB_r \, dW_r + \int_{0}^{T} f''(X_{r,T}^{Y_{r}}) X_{r,T}^{Y_{r}} \, dB_r \, dW_r
\]
\[
+ \frac{1}{2} \sum_{i,j=1}^{d} \int_{0}^{T} (B_r | B_s|^*)_{i,j} \left( f''(X_{r,T}^{Y_{r}}) (X_{r,T}^{Y_{r}}, X_{r,T}^{Y_{r}}), X_{r,T}^{Y_{r}}, X_{r,T}^{Y_{r}}) + f'(X_{r,T}^{Y_{r}}) X_{r,T}^{Y_{r}} \right) (e_i^{(d)}, e_j^{(d)}) \, dr.
\]
Next we analyze the first sum on the right-hand side of equation (21). For all \((s, t) \in \Delta_T, x \in O\) it holds that \(\mathbb{P}\left(X^x_{s,t} = X^x_{t,T}\right) = 1\). This and the fact that \(X\) is a continuous random field imply for all \((s, t) \in \Delta_T\) that \(\mathbb{P}\left(Y_{s,t} = Y_{t,T}\right) = 1\). For all \(t \in [0, T], x \in O, i \in \{1, 2\}\) the functions \(\Omega \ni \omega \mapsto X^x_{t,T}(\omega) \in O, \Omega \ni \omega \mapsto X_j^x(\omega) \in L^1(\mathbb{R}^d, \mathbb{R}^d)\) are \(\mathcal{G}(\mathcal{N} \cup \mathcal{G}(W_s - W_t; s \in [t, T]))\)-measurable. This together with the fact that for all \(\omega \in \Omega, t \in [0, T]\) it holds that \((O \ni x \mapsto f(X^x_{t,T}(\omega)) \in \mathbb{R}^k)\) is \(C^2(O, \mathbb{R}^k)\) implies that for all \(t \in [0, T]\) the function \(\Omega \ni \omega \mapsto \left(O \ni x \mapsto f(X^x_{t,T}(\omega)) \in \mathbb{R}^k\right) \in C^2(O, \mathbb{R}^k)\) is independent of the sigma-algebra \(\mathcal{F}_t\). Itô's formula for independent random fields (e.g., Klenke [31] Theorem 25.30 and Remark 25.26) implies for all \(\omega \in \Omega \ni \omega \mapsto \left(O \ni x \mapsto f(X^x_{t,T}(\omega)) \in \mathbb{R}^k\right) \in C^2(O, \mathbb{R}^k)\) for \(n \in \mathbb{N}, i \in \{0, 1, \ldots, n - 1\}, h \in \{\frac{T}{n}\}\) that for all \(n \in \mathbb{N}, i \in \{0, 1, \ldots, n - 1\}, h \in \{\frac{T}{n}\}\) it holds \(\mathbb{P}\)-a.s. that

\[
\begin{align*}
&f(X^x_{i+1, h, T}) - f(X^x_{ih, h, T}) = f(X^x_{i+1, h, T}) - f(X^x_{ih, h, T}) \\
&= \int_{h}^{(i+1)h} f'(X^x_{i, h, T}) X^1, X^x_{i, h, T} dX^x_{ih, T} \\
&\quad + \frac{1}{2} \sum_{i,j=1}^{d} \left( f''(X^x_{i, h, T}) \left( X^1, X^x_{i, h, T}, X^1, X^x_{i, h, T} \right) + f''(X^x_{i, h, T}) \left( X^1, X^x_{i, h, T}, X^1, X^x_{i, h, T} \right) \right) \left( \epsilon_i^{(d)}, \epsilon_j^{(d)} \right) d((X^x_{ih, h, r})_{i,j}) \\
&\quad + \frac{1}{2} \sum_{i,j=1}^{d} \left( \sigma(r, X^x_{r, h, T}) \left( \sigma(r, X^x_{r, h, T}) \right) \right) \left( \sigma(r, X^x_{r, h, T}) \right) \left( \epsilon_i^{(d)}, \epsilon_j^{(d)} \right) dW_{r} \\
&\quad + \frac{1}{2} \sum_{i,j=1}^{d} f''(X^x_{i, h, T}) \left( X^1, X^x_{i, h, T}, X^1, X^x_{i, h, T} \right) \left( \epsilon_i^{(d)}, \epsilon_j^{(d)} \right) dr.
\end{align*}
\]

Hölder’s inequality and inequalities (23), (18), (17) imply that for all \(n \in \mathbb{N}, i \in \{0, 1, \ldots, n - 1\}, h \in \{\frac{T}{n}\}\) it holds that

\[
\begin{align*}
&\left\| f'(X^x_{i, h, T}) X^1, X^x_{i, h, T} \sigma(\cdot, X^x_{i, h, T}) \right\|_{L^2(\mathbb{P}; L^2(\mathbb{R}^m \times \mathbb{R}^k \times \mathbb{R}^m \times \mathbb{R}^k))} \\
&\leq \left\| f'(X^x_{i, h, T}) X^1, X^x_{i, h, T} \sigma(\cdot, X^x_{i, h, T}) \right\|_{L^2(\lambda_{\mathbb{I}^\mathbb{H} \times \mathbb{H} \times \mathbb{I}^\mathbb{H} \times \mathbb{H}}; \mathbb{R}^m \times \mathbb{R}^k \times \mathbb{R}^m \times \mathbb{R}^k)} \\
&\leq \left\| f'(X^x_{i, h, T}) \right\|_{L^2(\mathbb{P}; L^2(\mathbb{R}^m \times \mathbb{R}^k \times \mathbb{R}^m \times \mathbb{R}^k))} \left\| X^1, X^x_{i, h, T} \sigma(\cdot, X^x_{i, h, T}) \right\|_{L^2(\mathbb{H} \mathcal{S}(\mathbb{R}^m \times \mathbb{R}^d))} \\
&\leq \left\| f'(X^x_{i, h, T}) \right\|_{L^2(\mathbb{P}; L^2(\mathbb{R}^m \times \mathbb{R}^d))} \left\| X^1, X^x_{i, h, T} \sigma(\cdot, X^x_{i, h, T}) \right\|_{L^2(\mathbb{P}; L^2(\mathbb{R}^m \times \mathbb{R}^d)} \\
&\leq \left\| f'(X^x_{i, h, T}) \right\|_{L^2(\mathbb{P}; L^2(\mathbb{R}^m \times \mathbb{R}^d))} \left\| X^1, X^x_{i, h, T} \sigma(\cdot, X^x_{i, h, T}) \right\|_{L^2(\mathbb{P}; L^2(\mathbb{R}^m \times \mathbb{R}^d)} \\
&\leq T^{-\frac{2p}{2p-2}} \left( \sup_{r,s,t \in [0,T]} \left\| f'(X^x_{r,s}, T) \right\|_{L^p(\mathbb{P}; L^p(\mathbb{R}^m \times \mathbb{R}^d))} \right) \left( \sup_{r,s,t \in [0,T]} \left\| X^1, X^x_{r,s} \right\|_{L^p(\mathbb{P}; L^p(\mathbb{R}^m \times \mathbb{R}^d))} \right).
\end{align*}
\]
For all $n \in \mathbb{N}$, $i \in \{0,1,\ldots,n-1\}$, $h \in \{\frac{t}{n}\}$ the process \( f'(X_{i}^{Y_{ih},r}) X_{i}^{1,Y_{ih},r} \sigma(r, X_{i}^{Y_{ih},r}) \) is predictable with respect to the filtration \( \mathcal{F}_{(i+1)h} \). Proposition A.8 together with inequality (29), Proposition A.7, and linearity of the Skorohod integral assert that the process \( f'(X_{i}^{Y_{ih},r}) X_{i}^{1,Y_{ih},r} \sigma(r, X_{i}^{Y_{ih},r}) \) is Skorohod-integrable and that for all $n \in \mathbb{N}$, $h \in \{\frac{T}{n}\}$ it holds $\mathbb{P}$-a.s. that

\[
\sum_{i=0}^{n-1} f'(X_{i}^{Y_{ih},r}) X_{i}^{1,Y_{ih},r} \sigma(r, X_{i}^{Y_{ih},r}) dW_{r} = \sum_{i=0}^{n-1} f'(X_{i}^{Y_{ih},r}) X_{i}^{1,Y_{ih},r} \sigma(r, X_{i}^{Y_{ih},r}) \delta W_{r}^{\mathbb{P}}(\cup_{s \in [(i+1)h,T]} W_{r} - W_{(i+1)h})
\]

(30)

Equations (28) and (30) imply that for all $n \in \mathbb{N}$, $h \in \{\frac{T}{n}\}$ it holds $\mathbb{P}$-a.s. that

\[
\sum_{i=0}^{n-1} \left( f(X_{i+1}^{Y_{ih},r}) - f(X_{ih}^{Y_{ih},r}) \right)
= \sum_{i=0}^{n-1} \mu \left( X_{i}^{Y_{ih},r} \right) dr + \sum_{i=0}^{n-1} \mu_{e} \left( X_{i}^{Y_{ih},r} \right) \sigma(r, X_{i}^{Y_{ih},r}) \delta W_{r}
\]

\[
+ \frac{1}{2} \sum_{i,j=1}^{d} \left( \sigma(r, X_{i}^{Y_{ih},r}) \right)_{i,j} \left( f''(X_{i}^{Y_{ih},r}) \right)_{i,j} \left( X_{i}^{1,Y_{ih},r} \right)_{i,j} + \frac{1}{2} \sum_{i,j=1}^{d} \left( \sigma(r, X_{i}^{Y_{ih},r}) \right)_{i,j} \left( f''(X_{i}^{Y_{ih},r}) \right)_{i,j} \left( X_{i}^{2,Y_{ih},r} \right)_{i,j}
\]

(31)

Equations (21), (31), and (27) imply that for all $h \in T/\mathbb{N}$ it holds $\mathbb{P}$-a.s. that

\[
f(X_{0,T}) - f(Y_{T})
= \sum_{i=0}^{T} \mu \left( X_{i}^{Y_{ih},r} \right) dr + \sum_{i=0}^{T} \mu_{e} \left( X_{i}^{Y_{ih},r} \right) \sigma(r, X_{i}^{Y_{ih},r}) \delta W_{r}
\]

\[
+ \frac{1}{2} \sum_{i,j=1}^{d} \left( \sigma(r, X_{i}^{Y_{ih},r}) \right)_{i,j} \left( f''(X_{i}^{Y_{ih},r}) \right)_{i,j} \left( X_{i}^{1,Y_{ih},r} \right)_{i,j} + \frac{1}{2} \sum_{i,j=1}^{d} \left( \sigma(r, X_{i}^{Y_{ih},r}) \right)_{i,j} \left( f''(X_{i}^{Y_{ih},r}) \right)_{i,j} \left( X_{i}^{2,Y_{ih},r} \right)_{i,j}
\]

(32)
$$- \frac{1}{2} \sum_{i,j=1}^{d} \int_{0}^{T} (B_{i}(B_{j})*)_{l,j} \left( f''(X_{1,Y,r}^{i}, T) X_{1,Y,r}^{i} + f'(X_{1,Y,r}^{i}, T) X_{2,Y,r}^{i} \right) (e_{i}^{(d)}, e_{j}^{(d)}) \, dr.$$ 

Next we want to let $T/n \to h \to 0$ in (32) in a suitable sense and first justify this. Hölder's inequality, inequalities (23), (18), (17), and the fact that $A \in L^{p}(\lambda_{[0,T]} \otimes \mathbb{P}; \mathbb{R}^{d})$ imply that

$$\sup_{h \in T/n} \left\| f'(X_{1,Y,r}^{i}, T) X_{1,Y,r}^{i} \right\|_{L^{2}(\lambda_{[0,T]} \otimes \mathbb{P}; \mathbb{R}^{d})} + \left\| f'(X_{1,Y,r}^{i}, T) X_{2,Y,r}^{i} \right\|_{L^{2}(\lambda_{[0,T]} \otimes \mathbb{P}; \mathbb{R}^{d})} < \infty.$$ 

Hölder’s inequality and inequalities (23) and (18) imply that for all $l, j \in \{1, \ldots, d\}$ it holds that

\begin{align*}
\sup_{h \in T/n} \left\| f''(X_{1,Y,r}^{i}, T) \left( X_{1,Y,r}^{i} X_{1,Y,r}^{j} - f'(X_{1,Y,r}^{i}, T) X_{1,Y,r}^{j} \right) \right\|_{L^{2}(\lambda_{[0,T]} \otimes \mathbb{P}; \mathbb{R}^{d})} \\
\leq \sup_{h \in T/n} \left\| f''(X_{1,Y,r}^{i}, T) \right\|_{L^{2}(\mathbb{R}^{d}, \mathbb{R}^{d})} + \left\| f'(X_{1,Y,r}^{i}, T) \right\|_{L^{2}(\mathbb{R}^{d}, \mathbb{R}^{d})} < \infty.
\end{align*} 

and, analogously, that for all $i, j \in \{1, \ldots, d\}$ it holds that

\begin{align*}
\sup_{h \in T/n} \left\| f''(X_{1,Y,r}^{i}, T) \left( X_{1,Y,r}^{i} X_{2,Y,r}^{j} + f'(X_{1,Y,r}^{i}, T) X_{2,Y,r}^{j} \right) \right\|_{L^{2}(\lambda_{[0,T]} \otimes \mathbb{P}; \mathbb{R}^{d})} < \infty.
\end{align*} 

The fact that for all $C \in \mathbb{R}^{d \times m}$ it holds that $\sum_{i,j=1}^{d} |(CC^{*})_{i,j}| \leq d\|C\|_{HS(\mathbb{R}^{m}, \mathbb{R}^{d})}^{2}$, Hölder’s inequality, assump-
tion (17) and inequality (34) imply that
\[
\sup_{h \in \mathbb{T}/n} \frac{1}{2} \sum_{l,j=1}^{d} \left( \sigma(\cdot, X_{l,h}^{Y_{l,h}}) \left[ \sigma(\cdot, X_{l,h}^{Y_{l,h}}) \right]^* \right)_{l,j} \\
\cdot \left( f'' \left( X_{l,h}^{Y_{l,h}}, X_{l,h}^{Y_{l,h}} \right) \left[ f'' \left( X_{l,h}^{Y_{l,h}}, X_{l,h}^{Y_{l,h}} \right) \right] + f' \left( X_{l,h}^{Y_{l,h}}, X_{l,h}^{Y_{l,h}} \right) \left[ f' \left( X_{l,h}^{Y_{l,h}}, X_{l,h}^{Y_{l,h}} \right) \right] \right) \left( e_{l}^{(d)}, e_{j}^{(d)} \right) \left\| L^2(\lambda_{[0,T]} \otimes \mathbb{P} \otimes \mathbb{R}^k) \right\|
\leq \sup_{h \in \mathbb{T}/n} \frac{1}{2} \left\| \sigma(\cdot, X_{l,h}^{Y_{l,h}}) \right\|_{\text{HS}(\mathbb{R}^m, \mathbb{R}^d)}^2 \\
\cdot \sum_{l,j=1}^{d} \left( f'' \left( X_{l,h}^{Y_{l,h}}, X_{l,h}^{Y_{l,h}} \right) \left[ f'' \left( X_{l,h}^{Y_{l,h}}, X_{l,h}^{Y_{l,h}} \right) \right] + f' \left( X_{l,h}^{Y_{l,h}}, X_{l,h}^{Y_{l,h}} \right) \left[ f' \left( X_{l,h}^{Y_{l,h}}, X_{l,h}^{Y_{l,h}} \right) \right] \right) \left( e_{l}^{(d)}, e_{j}^{(d)} \right) \left\| L^2(\lambda_{[0,T]} \otimes \mathbb{P} \otimes \mathbb{R}^k) \right\|
\leq \sup_{h \in \mathbb{T}/n} \frac{d}{2} \left\| \sigma(\cdot, X_{l,h}^{Y_{l,h}}) \right\|_{L^p(\lambda_{[0,T]} \otimes \mathbb{P} ; \mathbb{R}^{d \times m})}^2 \\
\cdot \sum_{l,j=1}^{d} \left( f'' \left( X_{l,h}^{Y_{l,h}}, X_{l,h}^{Y_{l,h}} \right) \left[ f'' \left( X_{l,h}^{Y_{l,h}}, X_{l,h}^{Y_{l,h}} \right) \right] + f' \left( X_{l,h}^{Y_{l,h}}, X_{l,h}^{Y_{l,h}} \right) \left[ f' \left( X_{l,h}^{Y_{l,h}}, X_{l,h}^{Y_{l,h}} \right) \right] \right) \left( e_{l}^{(d)}, e_{j}^{(d)} \right) \left\| L^{2p} \left( \lambda_{[0,T]} \otimes \mathbb{P} \otimes \mathbb{R}^k \right) \right\|
< \infty.
\] (36)

Analogously, the fact that for all \( C \in \mathbb{R}^{d \times m} \) it holds that \( \sum_{i,j=1}^{d} |(CC^*)_{i,j}| \leq d \| C \|_{\text{HS}(\mathbb{R}^m, \mathbb{R}^d)}^2 \), Hölder’s inequality, the assumption \( B \in L^p(\lambda_{[0,T]} \otimes \mathbb{P} ; \mathbb{R}^{d \times m}) \), and inequality (35) yield that
\[
\sup_{h \in \mathbb{T}/n} \frac{1}{2} \left\| \sum_{l,j=1}^{d} (B[B]^*)_{l,j} \left( f'' \left( X_{l,h}^{Y_{l}}, X_{l,h}^{Y_{l}} \right) \left[ f'' \left( X_{l,h}^{Y_{l}}, X_{l,h}^{Y_{l}} \right) \right] + f' \left( X_{l,h}^{Y_{l}}, X_{l,h}^{Y_{l}} \right) \left[ f' \left( X_{l,h}^{Y_{l}}, X_{l,h}^{Y_{l}} \right) \right] \right) \left( e_{l}^{(d)}, e_{j}^{(d)} \right) \left\| L^2(\lambda_{[0,T]} \otimes \mathbb{P} \otimes \mathbb{R}^k) \right\|
\leq \frac{d}{2} \| B \|_{L^p(\lambda_{[0,T]} \otimes \mathbb{P} \otimes \mathbb{R}^{d \times m})}^2 \\
\cdot \sum_{l,j=1}^{d} \sup_{h \in \mathbb{T}/n} \left\| \left( f'' \left( X_{l,h}^{Y_{l}}, X_{l,h}^{Y_{l}} \right) \left[ f'' \left( X_{l,h}^{Y_{l}}, X_{l,h}^{Y_{l}} \right) \right] + f' \left( X_{l,h}^{Y_{l}}, X_{l,h}^{Y_{l}} \right) \left[ f' \left( X_{l,h}^{Y_{l}}, X_{l,h}^{Y_{l}} \right) \right] \right) \left( e_{l}^{(d)}, e_{j}^{(d)} \right) \left\| L^{2p} \left( \lambda_{[0,T]} \otimes \mathbb{P} \otimes \mathbb{R}^k \right) \right\|
< \infty.
\] (37)

Next Klenke [31] Corollary 6.21 and Theorem 6.25] together with the uniform \( L^2 \)-bounds in (33), (36), and (37), continuity of \( f' \) and of \( f'' \), path continuity of \( Y \) and of \( \Delta_T \times O \ni (s, t, x) \mapsto X_{s,t}^x \in O \), and \( \inf_{r \in [0,T]} \mathbb{P}(X_{r,t}^r = \ldots)
Equation (32) and inequalities (40), (33), (36), and (37) imply that there exists a constant $K$ such that
\[
\lim_{T \to \infty} \left| \int_0^T f^\prime \left( X_{T,h}^{Y_T} \right) \mu(r, X_{r,T}^{Y_T}) - f^\prime \left( X_{T,h}^{Y_T} \right) X_{r,T}^{1,Y_T} A_r \ dr \right| = 0.
\]

Inequality (19) implies that for all $x, y \in O$ it holds that
\[
\| f(x) - f(y) \|_{\mathbb{R}^k} \leq \| f(x) \|_{\mathbb{R}^k} + \| f(y) \|_{\mathbb{R}^k} \leq c(1 + \| x \|_{\mathbb{R}^d})(1 + \| y \|_{\mathbb{R}^d})^q + c(1 + \| y \|_{\mathbb{R}^d})(1 + \| y \|_{\mathbb{R}^d})^q.
\]

Inequality (39), Hölder’s inequality, the fact that $2q + 2 < p$, the fact that $P(X_{0,T}^{Y_T} = X_{0,T}^{X_{0,T}^{Y_T}}) = 1 = P(Y_T = X_{T,T}^{Y_T})$, and inequality (18) show that
\[
\left\| f(X_{0,T}^{Y_T}) - f(Y_T) \right\|_{L^2(\mathbb{P};\mathbb{R}^k)} \leq c(1 + \| X_{0,T}^{Y_T} \|_{\mathbb{R}^d})^{1+q} \leq c(1 + \| Y_T \|_{\mathbb{R}^d})^{1+q} \leq \left( \sup_{r,s,t \in [0,T]} 2c \left( 1 + \left\| X_{r,T}^{Y_{r,s}} \right\|_{LP(\mathbb{P};\mathbb{R}^d)} \right) \right)^{q+1} < \infty.
\]

Equation (32) and inequalities (40), (33), (36), and (37) imply that there exists a constant $K \in [0, \infty)$ such
that for all $h \in T/\mathbb{N}$ it holds that
\[
\left\| \int_0^T f' \left( X_{[r],h,T}^{Y_{[r],h}} \right) X_{[r],h,T}^{1,1} \sigma(r, X_{[r],h,T}^{Y_{[r],h}}) - f' \left( X_{[r],h,T}^{Y_{[r],h}} \right) X_{[r],h,T}^{1,1} B_r \delta W_r \right\|_{L^2(\mathbb{P}; \mathbb{R}^k)} \\
\leq \left\| f(X_{0,T}^Y) - f(Y_T) \right\|_{L^2(\mathbb{P}; \mathbb{R}^k)} + \left\| \int_0^T f' \left( X_{[r],h,T}^{Y_{[r],h}} \right) X_{[r],h,T}^{1,1} \mu(r, X_{[r],h,T}^{Y_{[r],h}}) - f' \left( X_{[r],h,T}^{Y_{[r],h}} \right) X_{[r],h,T}^{1,1} A_r \, dr \right\|_{L^2(\mathbb{P}; \mathbb{R}^k)} \\
+ \left\| \frac{1}{2} \sum_{l,j=1}^d \int_0^T \left( \sigma(r, X_{[r],h,T}^{Y_{[r],h}}) [\sigma(r, X_{[r],h,T}^{Y_{[r],h}})]^* \right)_{l,j} \right. \\
\left. \cdot \left( f'' \left( X_{[r],h,T}^{Y_{[r],h}} \right) \left( X_{[r],h,T}^{1,1}, X_{[r],h,T}^{1,1} \right) + f' \left( X_{[r],h,T}^{Y_{[r],h}} \right) X_{[r],h,T}^{2,1} \right) \left( e_l^{(d)}, e_j^{(d)} \right) - (B_r[B_r]^*)_{l,j} \left( f'' \left( X_{[r],h,T}^{Y_{[r],h}} \right) \left( X_{[r],h,T}^{1,1}, X_{[r],h,T}^{1,1} \right) + f' \left( X_{[r],h,T}^{Y_{[r],h}} \right) X_{[r],h,T}^{2,1} \right) \left( e_l^{(d)}, e_j^{(d)} \right) \right\|_{L^2(\mathbb{P}; \mathbb{R}^k)} < K. \tag{41}
\]

The fact that $Y, X, X^1$ are continuous random fields, continuity of $f'$, and the fact that $\inf_{r \in [0,T]} \mathbb{P}(X_{[r],h,T}^{Y_{[r],h}} = Y_r = 1)$ yield that for all $r \in [0,T]$ it holds $\mathbb{P}$-a.s. that
\[
\lim_{T/h \downarrow 0} \left( f' \left( X_{[r],h,T}^{Y_{[r],h}} \right) X_{[r],h,T}^{1,1} \sigma(r, X_{[r],h,T}^{Y_{[r],h}}) - f' \left( X_{[r],h,T}^{Y_{[r],h}} \right) X_{[r],h,T}^{1,1} B_r \right) = f' \left( X_{[r],h,T}^{Y_{[r],h}} \right) X_{[r],h,T}^{1,1} \sigma(r, Y_r) - B_r. \tag{42}
\]

This, Fatou’s lemma, and the inequalities (29) and (24) yield that the sequence
\[
\left( f' \left( X_{[r],h,T}^{Y_{[r],h}} \right) X_{[r],h,T}^{1,1} \sigma(\cdot, X_{[r],h,T}^{Y_{[r],h}}) - f' \left( X_{[r],h,T}^{Y_{[r],h}} \right) X_{[r],h,T}^{1,1} B_r \right) \left( \sigma(\cdot, Y_r) - B_r \right) \tag{43}
\]
is bounded in $L^2(\lambda_{[0,T]} \otimes \mathbb{P}; \mathbb{R}^{k \times m})$. This, the fact that every bounded sequence in the separable Hilbert space $L^2(\lambda_{[0,T]} \otimes \mathbb{P}; \mathbb{R}^{k \times m})$ has a weakly converging subsequence (e.g., Kato [29] Lemma 5.1.4), and the convergence (42) ensure that the sequence (43) converges to 0 in the weak topology of $L^2(\lambda_{[0,T]} \otimes \mathbb{P}; \mathbb{R}^{k \times m})$ as $T/h \downarrow 0$. This, the fact that the processes
\[
\left( f' \left( X_{[r],h,T}^{Y_{[r],h}} \right) X_{[r],h,T}^{1,1} \sigma(r, X_{[r],h,T}^{Y_{[r],h}}) - f' \left( X_{[r],h,T}^{Y_{[r],h}} \right) X_{[r],h,T}^{1,1} B_r \right) \tag{44}
\]
are Skorohod-integrable, (41), and Lemma [A.9] imply that the stochastic process
\[
(f' \left( X_{[r],h,T}^{Y_{[r],h}} \right) X_{[r],h,T}^{1,1} \sigma(r, Y_r) - B_r)_{r \in [0,T]} \tag{45}
\]
is Skorohod-integrable and that for every $\mathbb{F}_T/\mathcal{B}([-1,1]^k)$-measurable function $Z : \Omega \rightarrow [-1,1]^k$ it holds that
\[
\lim_{T/h \downarrow 0} \mathbb{E} \left[ \left( Z \cdot \int_0^T f' \left( X_{[r],h,T}^{Y_{[r],h}} \right) X_{[r],h,T}^{1,1} \sigma(r, X_{[r],h,T}^{Y_{[r],h}}) - f' \left( X_{[r],h,T}^{Y_{[r],h}} \right) X_{[r],h,T}^{1,1} B_r \delta W_r \\
- \int_0^T f' \left( X_{[r],h,T}^{Y_{[r],h}} \right) X_{[r],h,T}^{1,1} \sigma(r, Y_r) - f' \left( X_{[r],h,T}^{Y_{[r],h}} \right) X_{[r],h,T}^{1,1} B_r \delta W_r \right)_{\mathbb{R}^k} \right] = 0. \tag{46}
\]
Equation (32) and the convergences (38) and (46) imply that for every \( F_T/\mathcal{B}([-1,1]^k) \)-measurable function \( Z: \Omega \to [-1,1]^k \) it holds that

\[
\mathbb{E} \left[ \int_0^T f(X_{r,T}^s)^2 \left( \mu(r,Y_r) - A_r \right) dr + \int_0^T f'(X_{r,T}^s)^2 \left( \sigma(r,Y_r) - B_r \right) \delta W_r 
+ \sum_{i,j=1}^d \int_0^T \left( \sigma(r,Y_r)[\sigma(r,Y_r)]^* - B_r[B_r]^* \right) \left( f''(X_{r,T}^s) (X_{r,T}^s, X_{r,T}^s) + f'(X_{r,T}^s) X_{r,T}^s \right) (e_i(d), e_j(d)) \right] = 0.
\]

This implies equation (20). The proof of Theorem 3.1 is thus completed. \( \square \)

**Appendix: The Skorohod integral with respect to Brownian motion and additional independent information**

In this appendix we introduce the Skorohod integral with respect to a Brownian motion \( W \) and an additional sigma-algebra \( \mathbb{F}_0 \) which is independent of \( W \). As a motivation, note that for every probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and every standard Brownian motion \( W: [0,3] \times \Omega \to \mathbb{R} \) the Itô integrals \( f_1^3 \sin(W_s(W_2 - W_1)) dW_s \) and \( f_2^3 \sin(W_s(W_3 - W_2)) dW_s \) are well-defined (however with respect to different filtrations) but their sum cannot be written as Itô integral \( f_2^3 \sin(W_s(W_{[s]+1} - W_{[s]})) dW_s \) (which is not well-defined). In this appendix we provide sufficient results to rewrite Itô integrals as Skorohod integrals and then to write the sum of these as a single Skorohod integral.

**Setting A.1.** Let \( d, m \in \mathbb{N} \), let \( S, T \in \mathbb{R} \) satisfy \( S < T \), let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space, let \( W: [S,T] \times \Omega \to \mathbb{R}^m \) be a stochastic process such that \((W_{S+t} - W_S)_{t \in [0,T-S]} \) is a standard Brownian motion with continuous sample paths, let \( \mathbb{F}_S \subseteq \mathcal{F} \) be a sigma-algebra which is independent of \( \mathcal{G}(W_t - W_S: t \in [S,T]) \), let \( \mathcal{N} = \{A \in \mathcal{F}: \mathbb{P}(A) = 0\} \), let \( \mathbb{F}_T \subseteq \mathcal{F} \) be the sigma-algebra which satisfies that \( \mathbb{F}_T = \mathcal{G}(\mathbb{F}_S \cup \mathcal{G}(W_t - W_S: t \in [S,T]) \cup \mathcal{N}) \), let \( \mathcal{S}(P, \mathbb{F}_S, W; \mathbb{R}^d) \subseteq L^2(\mathbb{P}|_{\mathbb{F}_T}; \mathbb{R}^d) \) be the subset with the property that

\[
\mathcal{S}(P, \mathbb{F}_S, W; \mathbb{R}^d) = \left\{ F \in L^2(\mathbb{P}|_{\mathbb{F}_T}; \mathbb{R}^d): \exists n \in \mathbb{N}, \exists \phi_1, \ldots, \phi_n \in L^2(\lambda|_{[S,T]}; \mathbb{R}^m), \forall f \in C^\infty_b(C^b(S \cup \mathcal{N})(\mathbb{R} \times \Omega, \mathbb{R}), \forall h \in \mathbb{R}^d \text{ such that it holds } \mathbb{P}\text{-a.s. that } F = f(j_{S}^T \phi_1(r) dW_r, \ldots, j_{S}^T \phi_n(r) dW_r) h \right\},
\]

and for all \( s, t \in [S,T] \) satisfying that \( s < t \) let \( \mathbb{F}_{[S,s] \cup [t,T]} \subseteq \mathcal{F} \) be the sigma-algebra with the property that \( \mathbb{F}_{[S,s] \cup [t,T]} = \mathcal{G}(\mathbb{F}_S \cup \mathcal{G}(W_t - W_S: r \in [S,s]) \cup \mathcal{G}(W_t - W_S: r \in [t,T]) \cup \mathcal{N}) \).

**Definition A.2.** Assume Setting A.1. The extended Malliavin differential operator

\[
\mathcal{D}(P, \mathbb{F}_S, W; \mathbb{R}^d) : \mathcal{D}^{1,2}(P, \mathbb{F}_S, W; \mathbb{R}^d) \to L^2(\mathbb{P}|_{\mathbb{F}_T}; L^2(\lambda|_{[S,T]}; \mathbb{R}^{d \times m}))
\]

is the closed linear operator with the property that for all \( F \in \mathcal{S}(P, \mathbb{F}_S, W; \mathbb{R}^d) \) with the property that \( \exists n \in \mathbb{N}, \exists \phi_1, \ldots, \phi_n \in L^2(\lambda|_{[S,T]}; \mathbb{R}^m), \forall f \in C^\infty_b(C^b(S \cup \mathcal{N})(\mathbb{R} \times \Omega, \mathbb{R}), \forall h \in \mathbb{R}^d \text{ such that it holds } \mathbb{P}\text{-a.s. that } F = f(j_{S}^T \phi_1(r) dW_r, \ldots, j_{S}^T \phi_n(r) dW_r) h \) it holds \( \mathbb{P}\text{-a.e. that}

\[
\mathcal{D}(P, \mathbb{F}_S, W; \mathbb{R}^d) F = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \left(T_S \phi_1(s) dW_s, \ldots, T_S \phi_n(s) dW_s\right) \phi_i h
\]

(50)
and where $\mathbb{D}^{(1,2)}(\mathbb{P}, \mathbb{F}, S, W; \mathbb{R}^d)$ is the closure of $\text{span}(\mathbb{S}(\mathbb{P}, \mathbb{F}, S, W; \mathbb{R}^d)) \subseteq L^2(\mathbb{P}|_{\mathcal{F}_T}; \mathbb{R}^d)$ with respect to the norm

$$\| \cdot \|_{\mathbb{D}^{(1,2)}(\mathbb{P}, \mathbb{F}, S, W; \mathbb{R}^d)} = \left( \mathbb{E}\left[ \| \cdot \|_{\mathbb{R}^d}^2 + \| \mathcal{D}(\mathbb{P}, \mathbb{F}, S, W; \mathbb{R}^d) \cdot \|_{L^2(\lambda_{[S,T]}; \mathbb{R}^{d\times m})}^2 \right] \right)^{1/2}. $$

We write $\mathcal{D} = \mathcal{D}(\mathbb{P}, \mathbb{G}(\mathcal{N}), W; \mathbb{R}^d)$ and denote $\mathcal{D}$ as the classical Malliavin derivative.

The following lemma, Lemma A.3, shows that the extended Malliavin derivative is well-defined (in particular, the left-hand side of (50) does not depend on the representative and such a closed linear operator exists). The proof of Lemma A.3 is almost literally identical to the proofs of Proposition 4.2 and Proposition 4.4 in Kruse [33] and therefore omitted.

**Lemma A.3.** Assume Setting [A.1] Then the operator

$$\mathcal{D}(\mathbb{P}, \mathbb{F}, S, W; \mathbb{R}^d): \mathbb{D}^{(1,2)}(\mathbb{P}, \mathbb{F}, S, W; \mathbb{R}^d) \rightarrow L^2(\mathbb{P}|_{\mathcal{F}_T}; L^2(\lambda_{[S,T]}; \mathbb{R}^{d\times m}))$$

is well-defined.

The following lemma, Lemma A.4, shows that the set $\mathbb{S}(\mathbb{P}, \mathbb{F}, S, W; \mathbb{R}^d)$ is sufficiently rich. The proof of Lemma A.4 is standard and therefore omitted.

**Lemma A.4.** Assume Setting [A.1] Then span $\left( \mathbb{S}(\mathbb{P}, \mathbb{F}, S, W; \mathbb{R}^d) \right)$ is dense in $L^2(\mathbb{P}|_{\mathcal{F}_T}; \mathbb{R}^d)$.

In particular, Lemma A.4 implies that the extended Malliavin differential operator is densely defined. Next we introduce the adjoint of the densely defined extended Malliavin differential operator.

**Definition A.5.** Assume Setting [A.1] The extended Skorohod integral is the linear operator

$$\delta(\mathbb{P}, \mathbb{F}, S, W; \mathbb{R}^d): \text{Dom}_{\delta}(\mathbb{P}, \mathbb{F}, S, W; \mathbb{R}^d) \rightarrow L^2(\mathbb{P}|_{\mathcal{F}_T}; \mathbb{R}^d)$$

which satisfies that $X \in L^2(\mathbb{P}|_{\mathcal{F}_T}; L^2(\lambda_{[S,T]}; \mathbb{R}^{d\times m}))$ is in the domain $\text{Dom}_{\delta}(\mathbb{P}, \mathbb{F}, S, W; \mathbb{R}^d)$ if and only if there exists a $c \in [0, \infty)$ with the property that for all $F \in \text{span} \left( \mathbb{S}(\mathbb{P}, \mathbb{F}, S, W; \mathbb{R}^d) \right)$ it holds that

$$\mathbb{E}\left[ \langle \mathcal{D}(\mathbb{P}, \mathbb{F}, S, W; \mathbb{R}^d)F, X \rangle_{L^2(\lambda_{[S,T]}; \mathbb{R}^{d\times m})} \right] \leq c \| F \|_{L^2(\mathbb{P}; \mathbb{R}^d)}$$

and which satisfies that for all $X \in \text{Dom}_{\delta}(\mathbb{P}, \mathbb{F}, S, W; \mathbb{R}^d), F \in \mathbb{S}(\mathbb{P}, \mathbb{F}, S, W; \mathbb{R}^d)$ it holds that

$$\mathbb{E}\left[ \langle F, \delta(\mathbb{P}, \mathbb{F}, S, W; \mathbb{R}^d)(X) \rangle_{\mathbb{R}^d} \right] = \mathbb{E}\left[ \langle \mathcal{D}(\mathbb{P}, \mathbb{F}, S, W; \mathbb{R}^d)F, X \rangle_{L^2(\lambda_{[S,T]}; \mathbb{R}^{d\times m})} \right].$$

We say that $X$ is $(\mathbb{P}, \mathbb{F}, S, W; \mathbb{R}^d)$-Skorohod-integrable if and only if $X \in \text{Dom}_{\delta}(\mathbb{P}, \mathbb{F}, S, W; \mathbb{R}^d)$. For all $X \in \text{Dom}_{\delta}(\mathbb{P}, \mathbb{F}, S, W; \mathbb{R}^d)$ we denote by $\int_S^T X_r \delta W_r^{\mathbb{F}S}$ the equivalence class satisfying that

$$\int_S^T X_r \delta W_r^{\mathbb{F}S} = \delta(\mathbb{P}, \mathbb{F}, S, W; \mathbb{R}^d)(X).$$

For all $X \in \text{Dom}_{\delta}(\mathbb{P}, \mathbb{G}(\mathcal{N}), W; \mathbb{R}^d)$ we denote by $\int_S^T X_r \delta W_r$ the equivalence class satisfying that

$$\int_S^T X_r \delta W_r = \int_S^T X_r \delta W_r^{\mathbb{G}(\mathcal{N})}$$

and we refer to $\int_S^T X_r \delta W_r$ as the classical Skorohod integral.

The following lemma will be applied in the proof of Proposition A.7.
Lemma A.6. Assume Setting [A.1] and let \( s, t \in [S, T] \) satisfy that \( s < t \). Then
\[
\mathbb{D}^{(1,2)}(\mathcal{P}, \mathcal{F}_S, W; \mathbb{R}^d) \subseteq \mathbb{D}^{(1,2)}(\mathcal{P}, \mathcal{F}_{[S,s] \cup [t,T]}, W|_{[s,t] \times \Omega}; \mathbb{R}^d) \tag{58}
\]
and for all \( F \in \mathbb{D}^{(1,2)}(\mathcal{P}, \mathcal{F}_S, W; \mathbb{R}^d) \) it holds \( \lambda_{[s,t]} \otimes \mathbb{P} \)-a.e. that
\[
\left( \mathcal{D}(\mathcal{P}, \mathcal{F}_S, W; \mathbb{R}^d)F \right) \big|_{[s,t] \times \Omega} = \mathcal{D}(\mathcal{P}, \mathcal{F}_{[S,s] \cup [t,T]}, W|_{[s,t] \times \Omega}; \mathbb{R}^d)F. \tag{59}
\]

Proof of Lemma A.6. Throughout this proof let \( F \in \mathcal{S}(\mathcal{P}, \mathcal{F}_S, W; \mathbb{R}^d) \), let \( n \in \mathbb{N} \), \( \phi_1, \ldots, \phi_n \in \mathcal{L}^2(\lambda_{[S,T]}; \mathbb{R}^m) \), \( f \in C_b^{\infty}(\mathcal{F}^{[S,T]})(\mathbb{R}^n \times \Omega, \mathbb{R}) \), and \( h \in \mathbb{R}^d \) satisfy that it holds \( \mathbb{P} \)-a.s. that
\[
F = f\left( \int_S^T \phi_1(r) dW_r, \ldots, \int_S^T \phi_n(r) dW_r \right) h, \tag{60}
\]
and let \( g \in C^{\infty}_b(\mathcal{F}^{[S,s] \cup [t,T]})(\mathbb{R}^n \times \Omega, \mathbb{R}) \) be a function such that for all \( (x_1, \ldots, x_n) \in \mathbb{R}^n \) it holds \( \mathbb{P} \)-a.s. that
\[
g(x_1, \ldots, x_n) = f\left( x_1 + \int_S^s \phi_1(r) dW_r + \int_t^T \phi_1(r) dW_r, \ldots, x_n + \int_S^s \phi_n(r) dW_r + \int_t^T \phi_n(r) dW_r \right). \tag{61}
\]
Then it holds \( \mathbb{P} \)-a.s. that
\[
F = g\left( \int_S^T \phi_1(r) dW_r, \ldots, \int_S^T \phi_n(r) dW_r \right) h. \tag{62}
\]
This implies that \( F \in \mathcal{S}(\mathcal{P}, \mathcal{F}_{[S,s] \cup [t,T]}, W|_{[s,t] \times \Omega}; \mathbb{R}^d) \). Next for all \( i \in \{1, \ldots, n\} \) it holds \( \mathbb{P} \)-a.s. that
\[
\frac{\partial f}{\partial x_i} \left( \int_S^T \phi_1(r) dW_r, \ldots, \int_S^T \phi_n(r) dW_r \right) = \frac{\partial g}{\partial x_i} \left( \int_S^T \phi_1(r) dW_r, \ldots, \int_S^T \phi_n(r) dW_r \right). \tag{63}
\]
It follows that it holds \( \lambda_{[s,t]} \otimes \mathbb{P} \)-a.e. that
\[
\left( \mathcal{D}(\mathcal{P}, \mathcal{F}_S, W; \mathbb{R}^d)F \right) \big|_{[s,t] \times \Omega} = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \left( \int_S^T \phi_1(r) dW_r, \ldots, \int_S^T \phi_n(r) dW_r \right) \left( \phi_i|_{[s,t]} \right) h = \sum_{i=1}^n \frac{\partial g}{\partial x_i} \left( \int_S^T \phi_1(r) dW_r, \ldots, \int_S^T \phi_n(r) dW_r \right) \left( \phi_i|_{[s,t]} \right) h \tag{64}
\]
Equation [64] implies that
\[
\left\| F \right\|^2_{\mathbb{D}^{(1,2)}(\mathcal{P}, \mathcal{F}_{[S,s] \cup [t,T]}, W|_{[s,t] \times \Omega}; \mathbb{R}^d)} = E \left[ \left\| F \right\|^2_{\mathbb{R}^d} + \left\| \mathcal{D}(\mathcal{P}, \mathcal{F}_{[S,s] \cup [t,T]}, W|_{[s,t] \times \Omega}; \mathbb{R}^d)F \right\|^2_{L^2(\lambda_{[s,t]}; \mathbb{R}^d)} \right] \tag{65}
\]
Since \( F \in \mathcal{S}(\mathcal{P}, \mathcal{F}_S, W; \mathbb{R}^d) \) was chosen arbitrarily it follows that
\[
\text{span}(\mathcal{S}(\mathcal{P}, \mathcal{F}_S, W; \mathbb{R}^d)) \subseteq \text{span}(\mathcal{S}(\mathcal{P}, \mathcal{F}_{[S,s] \cup [t,T]}, W|_{[s,t] \times \Omega}; \mathbb{R}^d)). \tag{66}
\]
This and inequality [65] yield the inclusion [58], and equation [64] implies equation [59]. The proof of Lemma A.6 is thus completed. \( \square \)
The following result, Proposition A.7, shows how to change the domain of integration for Skorohod integrals.

**Proposition A.7.** Assume Setting A.1, let \( X \in L^0(\mathbb{P}; L^2(\lambda|_{S,T}; \mathbb{R}^{d \times m})) \), and let \( s, t \in [S, T] \) satisfy that \( s < t \). Then the following two statements are equivalent:

(i) It holds that \( X|_{[s,t] \times \Omega} \) is \((\mathbb{P}, \mathbb{F}|_{[S,s] \cup [t,T]}, W|_{[s,t] \times \Omega}; \mathbb{R}^d)\)-Skorohod-integrable.

(ii) It holds that \( 1_{[s,t]}X \) is \((\mathbb{P}, \mathbb{F}, W; \mathbb{R}^d)\)-Skorohod-integrable.

If any of these two statements is true, then it holds \( \mathbb{P}\text{-a.s.} \) that

\[
\int_s^t X_r \, \delta W_r^F|_{[S,s] \cup [t,T]} = \int_S^T 1_{[s,t]}(r) X_r \, \delta W_r^F.
\]  

**(Proof of Proposition A.7)** (i) implies (ii): Assume that the process \( X|_{[s,t] \times \Omega} \) is \((\mathbb{P}, \mathbb{F}|_{[S,s] \cup [t,T]}, W|_{[s,t] \times \Omega}; \mathbb{R}^d)\)-Skorohod-integrable. This implies that \( 1_{[s,t]}X \in L^2(\mathbb{P}|_{S,T}; L^2(\lambda|_{S,T}; \mathbb{R}^{d \times m})) \). Lemma A.6 and the definition of the Skorohod integral, and the Cauchy-Schwarz inequality imply for all \( F \in \mathcal{D}^{(1,2)}(\mathbb{P}, \mathbb{F}, W; \mathbb{R}^d) \) that

\[
\mathbb{E}\left[ \left( \langle D(\mathbb{P}, \mathbb{F}, W; \mathbb{R}^d)F, 1_{[s,t]}X \rangle \right|_{L^2(\lambda|_{S,T}; \mathbb{R}^{d \times m})} \right] 
= \mathbb{E}\left[ \left( \langle (D(\mathbb{P}, \mathbb{F}, W; \mathbb{R}^d)F)|_{[s,t] \times \Omega}, X|_{[s,t] \times \Omega} \rangle \right|_{L^2(\lambda|_{S,T}; \mathbb{R}^{d \times m})} \right] 
= \mathbb{E}\left[ \left( \langle D(\mathbb{P}, \mathbb{F}|_{[S,s] \cup [t,T]}, W|_{[s,t] \times \Omega}; \mathbb{R}^d)F, X|_{[s,t] \times \Omega} \rangle \right|_{L^2(\lambda|_{S,T}; \mathbb{R}^{d \times m})} \right] 
= \mathbb{E}\left[ \left( \langle F, \int_s^t X_r \, \delta W_r^F|_{[S,s] \cup [t,T]} \rangle \right|_{\mathbb{R}^d} \right] 
\leq \left\| \int_s^t X_r \, \delta W_r^F|_{[S,s] \cup [t,T]} \right\|_{L^2(\mathbb{P}; \mathbb{R}^d)} \cdot \| F \|_{L^2(\mathbb{P}; \mathbb{R}^d)} < \infty.
\]  

We conclude that \( 1_{[s,t]}X \) is \((\mathbb{P}, \mathbb{F}, W; \mathbb{R}^d)\)-Skorohod-integrable.

(ii) implies (i): Assume that \( 1_{[s,t]}X \) is \((\mathbb{P}, \mathbb{F}, W; \mathbb{R}^d)\)-Skorohod-integrable. This implies that it holds that \( X|_{[s,t] \times \Omega} \in L^2(\mathbb{P}|_{S,T}; L^2(\lambda|_{S,T}; \mathbb{R}^{d \times m})) \). Lemma A.6 and the definition of the Skorohod integral yield for all \( F \in \mathcal{D}^{(1,2)}(\mathbb{P}, \mathbb{F}, W; \mathbb{R}^d) \) that \( F \in \mathcal{D}^{(1,2)}(\mathbb{P}, \mathbb{F}|_{[S,s] \cup [t,T]}, W|_{[s,t] \times \Omega}; \mathbb{R}^d) \) and that

\[
\mathbb{E}\left[ \left( \langle (D(\mathbb{P}, \mathbb{F}|_{[S,s] \cup [t,T]}, W|_{[s,t] \times \Omega}; \mathbb{R}^d)F), X|_{[s,t] \times \Omega} \rangle \right|_{L^2(\lambda|_{S,T}; \mathbb{R}^{d \times m})} \right] 
= \mathbb{E}\left[ \left( \langle D(\mathbb{P}, \mathbb{F}, W; \mathbb{R}^d)F|_{[s,t] \times \Omega}, X|_{[s,t] \times \Omega} \rangle \right|_{L^2(\lambda|_{S,T}; \mathbb{R}^{d \times m})} \right] 
= \mathbb{E}\left[ \left( \langle D(\mathbb{P}, \mathbb{F}, W; \mathbb{R}^d)F, 1_{[s,t]}X \rangle \right|_{L^2(\lambda|_{S,T}; \mathbb{R}^{d \times m})} \right] 
= \mathbb{E}\left[ \left( \langle F, \int_S^T 1_{[s,t]}(r) X_r \, \delta W_r^F \rangle \right|_{\mathbb{R}^d} \right] 
\leq \left\| \int_S^T 1_{[s,t]}(r) X_r \, \delta W_r^F \right\|_{L^2(\mathbb{P}; \mathbb{R}^d)} \cdot \| F \|_{L^2(\mathbb{P}; \mathbb{R}^d)} < \infty.
\]  

Lemma A.4 shows that \( \text{span}(\mathcal{S}(\mathbb{P}, \mathbb{F}, W; \mathbb{R}^d)) \) is dense in \( L^2(\mathbb{P}|_{S,T}; \mathbb{R}^d) \). This, (68), (69), and the definition of the Skorohod integral imply that \( X|_{[s,t] \times \Omega} \) is \((\mathbb{P}, \mathbb{F}|_{[S,s] \cup [t,T]}, W|_{[s,t] \times \Omega}; \mathbb{R}^d)\)-Skorohod-integrable and that it holds \( \mathbb{P}\text{-a.s.} \) that

\[
\int_s^t X_r \, \delta W_r^F|_{[S,s] \cup [t,T]} = \int_S^T 1_{[s,t]}(r) X_r \, \delta W_r^F.
\]  

The proof of Proposition A.7 is thus completed. \( \square \)
It is well-known (e.g., Nualart [48 Proposition 1.3.11]) that the classical Skorohod integral generalizes the Itô integral restricted to square-integrable integrands which are adapted to the Brownian filtration. The following result, Proposition A.8, generalizes this. The proof of Lemma A.9 is analogous to the proof of Nualart [48 Proposition 1.3.11] and is therefore omitted.

**Proposition A.8.** Assume Setting A.1, let \( s, t \in [S, T] \) satisfy \( s < t \), let \( \tilde{F} = (\tilde{F}_r)_{r \in [s, t]} \) be a filtration with the property that for all \( r \in [s, t] \) it holds that \( \tilde{F}_r = \mathcal{G}(\mathcal{W}_u - \mathcal{W}_s : u \in [s, r]) \cup \mathcal{F}_{[s, s] \cup \{t, T\}} \) and let \( X \in \mathcal{L}^2(\mathbb{P} ; \mathcal{L}^2(\lambda_{[s, t]} ; \mathbb{R}^{d \times m})) \) be \( \tilde{F} \)-predictable. Then \( X \) is \((\mathbb{P}, \mathcal{F}_{[S, s] \cup \{t, T\}}, W|_{[s, t] \times \Omega}, \mathbb{R}^d)\)-Skorohod-integrable and it holds \( \mathbb{P} \)-a.s. that

\[
\int_s^t X_r \delta W_{r} = \int_s^t X_r d\mathcal{W}_r. \tag{71}
\]

The next result, Lemma A.9, proves that if a sequence of integrals converges weakly and has uniformly bounded Skorohod integrals, then the limit is Skorohod-integrable and the sequence of Skorohod integrals of the sequence converges weakly. Lemma A.9 follows immediately from the definition of the Skorohod integral and its proof is therefore omitted.

**Lemma A.9.** Assume Setting A.1, let \( X \in \mathcal{L}^2(\mathbb{P} |_{\mathcal{F}_T} ; \mathcal{L}^2(\lambda_{[S, T]} ; \mathbb{R}^{d \times m})) \), and let \((X_n)_{n \in \mathbb{N}} \subseteq \text{Dom}_\delta(\mathbb{P}, \mathcal{F}_S, W; \mathbb{R}^d)\) be a sequence which satisfies that \( \sup_{n \in \mathbb{N}} \| \delta(\mathbb{P}, \mathcal{F}_S, W; \mathbb{R}^d)(X_n) \|_{\mathcal{L}^2(\mathbb{P} |_{\mathcal{F}_T} ; \mathbb{R}^d)} < \infty \) and which converges to \( X \) in the weak topology of \( \mathcal{L}^2(\mathbb{P} |_{\mathcal{F}_T} ; \mathcal{L}^2(\lambda_{[S, T]} ; \mathbb{R}^{d \times m})) \). Then \( X \in \text{Dom}_\delta(\mathbb{P}, \mathcal{F}_S, W; \mathbb{R}^d) \) and \( (\delta(\mathbb{P}, \mathcal{F}_S, W; \mathbb{R}^d)(X_n))_{n \in \mathbb{N}} \) converges to \( \delta(\mathbb{P}, \mathcal{F}_S, W; \mathbb{R}^d)(X) \) in the weak topology of \( \mathcal{L}^2(\mathbb{P} |_{\mathcal{F}_T} ; \mathbb{R}^d) \).

**Acknowledgements**

This project has been partially supported by the Deutsche Forschungsgesellschaft (DFG) via RTG 2131 *High-dimensional Phenomena in Probability – Fluctuations and Discontinuity.*

**References**


[34] Kunita, H. *Lectures on stochastic flows and applications*, vol. 78 of *Tata Institute of Fundamental Research Lectures on Mathematics and Physics*. Published for the Tata Institute of Fundamental Research, Bombay, 1986.


