ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Seminar for Applied Mathematics SAM

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

A. Hudde and M. Hutzenthaler and A. Jentzen and S. Mazzonetto

Research Report No. 2019-03
January 2019

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

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

Anselm Hudde[1], Martin Hutzenthaler[1], Arnulf Jentzen[2] & Sara Mazzonetto[1]

[1] Faculty of Mathematics, University of Duisburg-Essen, 45117 Essen, Germany

[2] Seminar für Angewandte Mathematik, ETH Zurich, 8092 Zürich, Switzerland

January 7, 2019

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 proposed Itô-Alekseev-Gröbner formula is a powerful tool for deriving strong approximation rates for perturbations and approximations of stochastic ordinary and partial differential equations.

Contents

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

$$e^{at} - e^{bt} = -\int_0^t \tfrac{d}{ds}\big(e^{a(t-s)}e^{bs}\big)\, ds = \int_0^t e^{a(t-s)}(a-b)e^{bs}\, ds. \tag{1}$$

The nonlinear integration-by-parts formula, which is also referred to as *Alekseev-Gröbner formula* or as nonlinear variation-of-constants formula, generalizes this relation to nonlinear ordinary differential equations and has been established in Alekseev [1] and Gröbner [11]. More formally, the Alekseev-Gröbner formula (cf., e.g., Hairer et

AMS 2010 subject classification: 60H10

Key words and phrases: Itô formula, Alekseev-Gröbner formula, nonlinear variation-of-constants formula, nonlinear integration-by-parts formula, perturbation of stochastic differential equations, strong convergence rate, non-globally monotone coefficients, small-noise analysis

al. [13, Theorem I.14.5]) 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 $X^{\cdot}_{\cdot,\cdot} = (X^x_{s,t})_{s\in[0,t],t\in[0,T],x\in\mathbb{R}^d} \in C(\{(s,t)\in[0,T]^2\colon s\leq t\}\times\mathbb{R}^d,\mathbb{R}^d)$ with $\forall\, s\in[0,T]$, $t\in[s,T]$, $x\in\mathbb{R}^d$: $X^x_{s,t} = x + \int_s^t \mu(r,X^x_{s,r})\,dr$ it holds that

$$X^{Y_0}_{0,T} - Y_T = \int_0^T \Big(\tfrac{\partial}{\partial x} X^{Y_r}_{r,T}\Big)\Big(\mu(r,Y_r) - \tfrac{d}{dr}Y_r\Big)\,dr. \tag{2}$$

Informally speaking, the Alekseev-Gröbner formula expresses the global error (the term $X^{Y_0}_{0,T} - Y_T$ in (2)) in terms of the infinitesimal error (the term $\mu(r,Y_r) - \frac{d}{dr}Y_r$ in (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 [35, Theorem 3], Lie & Norsett [31, Theorem 1], Iserles & Soederlind [22, Theorem 1], and Iserles [21, 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 (8) below) depends both on the past (e.g. the term $\mu(r,Y_r)$) and on the future (e.g. the term $\frac{\partial}{\partial x}X^{Y_r}_{r,T}$). 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.

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 $W\colon[0,T]\times\Omega\to\mathbb{R}^m$ be a standard Brownian motion with continuous sample paths, let $\mathcal{N}=\{A\in\mathcal{F}\colon\mathbb{P}(A)=0\}$, let $\mu\colon[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$, $\sigma\colon[0,T]\times\mathbb{R}^d\to\mathbb{R}^{d\times m}$ be continuous functions, let $X^{\cdot}_{\cdot,\cdot} = (X^x_{s,t})_{s\in[0,t],t\in[0,T],x\in\mathbb{R}^d}\colon\{(s,t)\in[0,T]^2\colon s\leq t\}\times\mathbb{R}^d\times\Omega\to\mathbb{R}^d$ be a continuous random field, assume that for all $s\in[0,T]$, $\omega\in\Omega$ it holds that $(\mathbb{R}^d\ni x\mapsto X^x_{s,T}(\omega)\in\mathbb{R}^d)\in C^2(\mathbb{R}^d,\mathbb{R}^d)$, assume that for all $\omega\in\Omega$ it holds that $(\frac{\partial^2}{\partial x^2}X^{\cdot}_{\cdot,T}(\omega)\in L^{(2)}(\mathbb{R}^d,\mathbb{R}^d))\in C([0,T]\times\mathbb{R}^d,L^{(2)}(\mathbb{R}^d,\mathbb{R}^d))$, assume that for all $s\in[0,T]$, $x\in\mathbb{R}^d$ the stochastic process $[s,T]\times\Omega\ni(t,\omega)\mapsto X^x_{s,t}(\omega)\in\mathbb{R}^d$ is $(\mathfrak{S}(\mathcal{N}\cup\mathfrak{S}(W_r-W_s\colon r\in[s,t])))_{t\in[s,T]}$-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*

$$X^x_{s,t} = x + \int_s^t \mu(r,X^x_{s,r})\,dr + \int_s^t \sigma(r,X^x_{s,r})\,dW_r, \tag{3}$$

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 $X^{X^x_{s,t}}_{t,T} = X^x_{s,T}$, let $A,Y\colon[0,T]\times\Omega\to\mathbb{R}^d$, $B\colon[0,T]\times\Omega\to\mathbb{R}^{d\times m}$ be $(\mathfrak{S}(\mathcal{N}\cup\mathfrak{S}(W_r\colon r\in[0,t])))_{t\in[0,T]}$-predictable stochastic processes, assume that Y has continuous sample paths, assume that $\int_0^T\mathbb{E}\Big[\|A_s\|^p_{\mathbb{R}^d}+\|Y_s\|^p_{\mathbb{R}^d}+\|B_s\|^p_{\mathrm{HS}(\mathbb{R}^m,\mathbb{R}^d)}\Big]\,ds<\infty$, assume that for all $t\in[0,T]$ it holds $\mathbb{P}$-a.s. that

$$Y_t = \xi + \int_0^t A_s\,ds + \int_0^t B_s\,dW_s, \tag{4}$$

assume that

$$\sup_{\substack{s,t\in[0,T]\\ s\leq t}}\mathbb{E}\Big[\Big\|\mu\Big(t,X^{Y_s}_{s,t}\Big)\Big\|^p_{\mathbb{R}^d}+\Big\|\sigma\Big(t,X^{Y_s}_{s,t}\Big)\Big\|^p_{\mathrm{HS}(\mathbb{R}^m,\mathbb{R}^d)}\Big]<\infty, \tag{5}$$

assume that

$$\sup_{\substack{r,s,t\in[0,T]\\ r\leq s\leq t}}\mathbb{E}\Big[\Big\|X^{Y_r}_{r,s}\Big\|^p_{\mathbb{R}^d}+\Big\|\tfrac{\partial}{\partial x}X^{Y_r}_{r,s}\Big\|^{\frac{4p}{p-2(q+2)}}_{L(\mathbb{R}^d,\mathbb{R}^d)}+\Big\|\tfrac{\partial^2}{\partial x^2}X^{Y_r}_{r,s}\Big\|^{\frac{2p}{p-2(q+2)}}_{L^{(2)}(\mathbb{R}^d,\mathbb{R}^d)}\Big]<\infty, \tag{6}$$

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

$$\max\left\{\tfrac{\|f(x)\|_{\mathbb{R}^k}}{1+\|x\|_{\mathbb{R}^d}}, \|f'(x)\|_{L(\mathbb{R}^d,\mathbb{R}^k)}, \|f''(x)\|_{L^{(2)}(\mathbb{R}^d,\mathbb{R}^k)}\right\} \leq c(1+\|x\|_{\mathbb{R}^d}^q). \tag{7}$$

Then

(i) the stochastic process $\big(f'\big(X_{r,T}^{Y_r}\big)\frac{\partial}{\partial x}X_{r,T}^{Y_r}(\sigma(r,Y_r)-B_r)\big)_{r\in[0,T]}$ *is Skorohod-integrable and*

(ii) it holds $\mathbb{P}$*-a.s. that*

$$\begin{aligned} f\big(X_{0,T}^{Y_0}\big) - f(Y_T) &= \int_0^T f'\big(X_{r,T}^{Y_r}\big)\tfrac{\partial}{\partial x}X_{r,T}^{Y_r}\Big(\mu(r,Y_r)-A_r\Big)\,dr + \int_0^T f'\big(X_{r,T}^{Y_r}\big)\tfrac{\partial}{\partial x}X_{r,T}^{Y_r}\Big(\sigma(r,Y_r)-B_r\Big)\,\delta W_r \\ &+ \tfrac{1}{2}\sum_{i,j=1}^d \int_0^T \Big(\sigma(r,Y_r)[\sigma(r,Y_r)]^* - B_r[B_r]^*\Big)_{i,j}\Big(f''\big(X_{r,T}^{Y_r}\big)\big(\tfrac{\partial}{\partial x}X_{r,T}^{Y_r}, \tfrac{\partial}{\partial x}X_{r,T}^{Y_r}\big) + f'\big(X_{r,T}^{Y_r}\big)\tfrac{\partial^2}{\partial x^2}X_{r,T}^{Y_r}\Big)\big(e_i^{(d)}, e_j^{(d)}\big)\,dr. \end{aligned} \tag{8}$$

Theorem 1.1 follows immediately from Theorem 3.1 (applied with $\mathbb{F}_0 = \mathfrak{S}(\mathcal{N})$, $O = \mathbb{R}^d$ in the notation of Theorem 3.1). Here and throughout this article we denote by $\mathfrak{S}(\mathcal{A})$ the smallest σ-algebra generated by $\mathcal{A} \subseteq \mathcal{P}(\Omega)$. 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 $\sigma = 0$, $B = 0$, $k = d$, $= \mathrm{Id}_{\mathbb{R}^d}$ in the notation of Theorem 1.1) implies the Alekseev-Gröbner formula in (2) (cf., e.g., Hairer et al. [13, 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 [37, Theorem IV.3.3]) in the case where the Itô process Y, its drift process A, and its diffusion process B satisfy $\inf_{p\in(4,\infty)}\big(\sup_{s\in[0,T]}\mathbb{E}\big[\|Y_s\|_{\mathbb{R}^d}^p\big] + \int_0^T \mathbb{E}\big[\|A_s\|_{\mathbb{R}^d}^p + \|B_s\|_{\mathrm{HS}(\mathbb{R}^m,\mathbb{R}^d)}^p\big]\,ds\big) < \infty$. This moment requirement is due to the fact that we use the Skorohod integral. An approach with rough path integrals (cf., e.g., Hairer & Friz [9]) 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. [13, Theorem I.14.5]) even in the deterministic case ($\sigma = 0$ and $B = 0$ in the notation of Theorem 1.1) from $f = \mathrm{Id}_{\mathbb{R}^d}$ to general test functions. In Proposition 2.1 below we prove the Itô-Alekseev-Gröbner formula in (2) in the deterministic case with the test function $f\colon \mathbb{R}^d \to \mathbb{R}^k$ 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 [10, (7.48) and the last Display on page 182] or related weak error estimates in [38, 8, 39]). 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 $\mathbb{E}\big[f\big(X_{0,T}^{Y_0}\big)\big] - \mathbb{E}\big[f(Y_T)\big]$.

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., Alos & 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. To demonstrate the applicability of Theorem 1.1, we apply Theorem 1.1 to the stochastic van-der-Pol oscillator in Section 4 below and obtain in Lemma 4.5 that the stochastic van-der-Pol oscillator can be approximated with L^2-rate $\frac{1}{2}$.

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 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 $A_t = \mu(\frac{kT}{N}, Y_{\frac{kT}{N}})$ and $B_t = \sigma(\frac{kT}{N}, Y_{\frac{kT}{N}})$ for all $t \in [\frac{kT}{N}, \frac{(k+1)T}{N})$, $k \in \mathbb{N}_0 \cap [0, N)$),

(ii) strong convergence rates for *Galerkin approximations for stochastic evolution equations (SEEs)* (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 $\tilde{\sigma}\colon [0, T] \times \mathbb{R}^d \to \mathbb{R}^{d \times m}$ 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

$$\langle x - y, \mu(t, x) - \mu(t, y)\rangle_{\mathbb{R}^d} + \tfrac{1}{2}\|\sigma(t, x) - \sigma(t, y)\|^2_{\mathrm{HS}(\mathbb{R}^m, \mathbb{R}^d)} \leq c\|x - y\|^2_{\mathbb{R}^d} \tag{9}$$

cf. also [15] 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.

Our main motivation for the Itô-Alekseev-Gröbner formula are *strong convergence rates* for *time-discrete numerical approximations of SEEs*. In the literature, positive strong convergence rates have been established for SEEs with monotone nonlinearities; see, e.g., [12, 26, 23, 4, 3, 6, 5, 33, 40] for the case of additive noise and [34, 32] for the case of multiplicative noise. To the best of our knowledge, strong convergence rates for time-discrete approximations of SEEs with non-monotone superlinearly growing nonlinearities remain an open problem. This problem becomes now feasible by applying our perturbation result in Theorem 1.1. The details hereof are deferred to future publications. Summarizing, we believe that Theorem 1.1 is an appropriate tool to analyze temporal approximations of semilinear SEEs.

A crucial assumption in Theorem 1.1 is existence of a solution of the SDE (3) 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 [30] construct a two-dimensional example with smooth and globally bounded coefficient functions which is not even strongly complete (that is, the exceptional subset of Ω where (3) fails to hold can not be chosen independently of the starting point); cf. also Hairer et al. [14, 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., [7, 41, 29]. 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 [28, 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.

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 $\mathbb{N} = \{1, 2, 3, \dots\}$ and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. For all $c \in (0, \infty)$ let 0^0, $\frac{0}{0}$, $\frac{c}{0}$, $\frac{-c}{0}$, $0 \cdot \infty$, $0 \cdot (-\infty)$, ∞^c denote the extended real numbers $0^0 = 1$, $\frac{0}{0} = 0$, $\frac{c}{0} = \infty$, $\frac{-c}{0} = -\infty$, $0 \cdot \infty = 0$, $0 \cdot (-\infty) = 0$, and $\infty^c = \infty$. For all $T \in [0, \infty)$ let $\Delta_T \subseteq [0, T]^2$ denote the subset with the property that $\Delta_T = \{(s, t) \in [0, T]^2 \colon s \leq t\}$ and denote by ${}^{T}\!/_{\mathbb{N}}$ the set ${}^{T}\!/_{\mathbb{N}} = \{{}^{T}\!/_{n} \colon n \in \mathbb{N}\}$. For all $h \in (0, \infty)$, $r \in [0, \infty)$ let $\lceil r \rceil_h, \lfloor r \rfloor_h, \lceil r \rceil_0, \lfloor r \rfloor_0 \in [0, \infty)$ be the real numbers with the properties that $\lceil r \rceil_h = \inf\{nh \in [r, \infty) \colon n \in \mathbb{N}_0\}$, $\lfloor r \rfloor_h = \sup\{nh \in [0, r] \colon n \in \mathbb{N}_0\}$, $\lceil r \rceil_0 = r$, and $\lfloor r \rfloor_0 = r$. For a real vector space V and a subset $S \subseteq V$ let $\operatorname{span}(S) \subseteq V$ denote the set with the property that $\operatorname{span}(S) = \{\sum_{i=1}^{n} r_i v_i \colon n \in \mathbb{N}, r_1, \dots, r_n \in \mathbb{R}, v_1, \dots, v_n \in V\}$. 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 $i \in \{1, \dots, d\}$ let $e_i^{(d)}$ denote the i-th unit vector in $\mathbb{R}^d$. For all measurable spaces $(\Omega, \mathcal{F})$, $(\Omega', \mathcal{B})$ let $\mathcal{M}(\mathcal{F}, \mathcal{B})$ be the set $\mathcal{M}(\mathcal{F}, \mathcal{B}) = \{f \colon \Omega \to \Omega' \colon f \text{ is } \mathcal{F}/\mathcal{B}\text{-measurable}\}$. For every measure space $(\Omega, \mathcal{F}, \mu)$, every normed vector space $(V, \|\cdot\|_V)$, and all $p \in [1, \infty)$ let $\mathcal{B}(V)$ denote the Borel-sigma-algebra on V, let $\mathcal{L}^p(\mu; V)$ be the set with the property that $\mathcal{L}^p(\mu; V) = \{f \in \mathcal{M}(\mathcal{F}, \mathcal{B}(V)) \colon \int_\Omega \|f\|_V^p \, d\mu < \infty\}$, let $L^p(\mu; V)$ be the set with the property that $L^p(\mu, V) = \big\{\{f \in \mathcal{L}^p(\mu, V) \colon f = g \ \mu\text{-a.e.}\} \colon g \in \mathcal{L}^p(\mu, V)\big\}$, and let

$$\|\cdot\|_{L^p(\mu;V)} \colon \big(\mathcal{M}(\mathcal{F}, \mathcal{B}(V)) \cup \big\{\{f \in \mathcal{M}(\mathcal{F}, \mathcal{B}) \colon f = g \ \mu\text{-a.e.}\} \colon g \in \mathcal{M}(\mathcal{F}, \mathcal{B})\big\}\big) \to [0, \infty] \tag{10}$$

be the function which satisfies for all $f \in \big(\mathcal{M}(\mathcal{F}, \mathcal{B}(V)) \cup \big\{\{h \in \mathcal{M}(\mathcal{F}, \mathcal{B}) \colon h = g \ \mu\text{-a.e.}\} \colon g \in \mathcal{M}(\mathcal{F}, \mathcal{B})\big\}\big)$ that $\|f\|_{L^p(\mu;V)} = \big(\int_\Omega \|f\|_V^p d\mu\big)^{\frac{1}{p}}$. 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 $C_b^{\infty, \mathcal{F}}(\mathbb{R}^n \times \Omega, \mathbb{R})$ be the set which satisfies that

$$C_b^{\infty, \mathcal{F}}(\mathbb{R}^n \times \Omega, \mathbb{R}) = \left\{ f \colon \mathbb{R}^n \times \Omega \to \mathbb{R} \colon \begin{array}{c} \forall \omega \in \Omega \colon f(\cdot, \omega) \in C_b^\infty(\mathbb{R}^n, \mathbb{R}), \\ \forall x \in \mathbb{R}^d \colon f(x, \cdot) \text{ is } \mathcal{F}/\mathcal{B}(\mathbb{R})\text{-measurable} \end{array} \right\}. \tag{11}$$

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

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. [13, Theorem I.14.5]) (which is the special case $k = d$, $f = \mathrm{Id}_{\mathbb{R}^d}$ of Proposition 2.1) to general test functions.

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

$$f(X^{Y_0}_{0,T}) - f(Y_T) = \int_0^T f'(X^{Y_s}_{s,T}) \tfrac{\partial}{\partial x} X^{Y_s}_{s,T} \Big(\mu(s, Y_s) - \tfrac{d}{ds} Y_s\Big) \, ds. \tag{12}$$

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. [13, 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 $\frac{\partial}{\partial x} X^{\cdot}_{\cdot,\cdot} \in C(\{(s,t) \in [0,T]^2 \colon s \leq t\} \times O, L(\mathbb{R}^d, \mathbb{R}^d))$. Moreover, the assumptions, and Hairer et al. [13, 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 $\frac{\partial}{\partial s} X^{\cdot}_{\cdot,T} \in C([0,T] \times O, \mathbb{R}^d)$, and that for all $s \in [0,T]$, $x \in O$ it holds that

$$\tfrac{\partial}{\partial s} X^x_{s,T} = -\tfrac{\partial}{\partial x} X^x_{s,T} \mu(s,x). \tag{13}$$

Therefore, the chain rule implies that $([0,T] \ni s \mapsto X^{Y_s}_{s,T} \in O) \in C^1([0,T],O)$. Moreover, the fundamental theorem of calculus, the chain rule, and (13) yield that

$$\begin{aligned} f(X^{Y_0}_{0,T}) - f(Y_T) &= -\int_0^T \tfrac{d}{ds}\Big(f\big(X^{Y_s}_{s,T}\big)\Big)\, ds \\ &= -\int_0^T f'(X^{Y_s}_{s,T}) \left(\Big(\tfrac{\partial}{\partial s} X^x_{s,T}\Big)\Big|_{x=Y_s} + \tfrac{\partial}{\partial x} X^{Y_s}_{s,T} \tfrac{d}{ds} Y_s \right) ds \\ &= -\int_0^T f'(X^{Y_s}_{s,T}) \Big(-\tfrac{\partial}{\partial x} X^{Y_s}_{s,T} \mu(s,Y_s) + \tfrac{\partial}{\partial x} X^{Y_s}_{s,T} \tfrac{d}{ds} Y_s \Big) ds \\ &= \int_0^T f'(X^{Y_s}_{s,T}) \tfrac{\partial}{\partial x} X^{Y_s}_{s,T} \Big(\mu(s,Y_s) - \tfrac{d}{ds} Y_s \Big)\, ds. \end{aligned} \tag{14}$$

This finishes the proof of Proposition 2.1. □

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

$$X^x_{s,t} = x + \int_s^t \mu(r, X^x_{s,r})\, dr + \int_s^t \sigma(r, X^x_{s,r})\, dW_r, \tag{15}$$

assume that for all $(s,t) \in \Delta_T$, $x \in O$ it holds $\mathbb{P}$-a.s. that $X^{X^x_{s,t}}_{t,T} = X^x_{s,T}$, assume that for all $(s,x,\omega) \in [0,T] \times O \times \Omega$ it holds that $X^{1,x}_{s,T}(\omega) = \frac{\partial}{\partial x}\big(X^x_{s,T}(\omega)\big)$ and $X^{2,x}_{s,T}(\omega) = \frac{\partial^2}{\partial x^2}\big(X^x_{s,T}(\omega)\big)$, let $Y \in \mathcal{L}^p(\lambda_{[0,T]} \otimes \mathbb{P}; O)$, $A \in \mathcal{L}^p(\lambda_{[0,T]} \otimes \mathbb{P}; \mathbb{R}^d)$, $B \in \mathcal{L}^p(\lambda_{[0,T]} \otimes \mathbb{P}; \mathbb{R}^{d \times m})$ 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

$$Y_t = Y_0 + \int_0^t A_s\, ds + \int_0^t B_s\, dW_s, \tag{16}$$

assume that

$$\sup_{h \in T/\mathbb{N}} \mathbb{E}\left[\int_0^T \Big\| \mu\Big(t, X^{Y_{\lfloor t \rfloor_h}}_{\lfloor t \rfloor_h, t}\Big) \Big\|^p_{\mathbb{R}^d} + \Big\| \sigma\Big(t, X^{Y_{\lfloor t \rfloor_h}}_{\lfloor t \rfloor_h, t}\Big) \Big\|^p_{\mathrm{HS}(\mathbb{R}^m, \mathbb{R}^d)}\, dt \right] < \infty, \tag{17}$$

assume that

$$\sup_{\substack{r,s,t\in[0,T]\\ r\leq s\leq t}} \mathbb{E}\left[\left\|X_{t,T}^{X_{r,s}^{Y_r}}\right\|_{\mathbb{R}^d}^p + \left\|X_{t,T}^{1,X_{r,s}^{Y_r}}\right\|_{L(\mathbb{R}^d,\mathbb{R}^d)}^{\frac{4p}{p-2(q+2)}} + \left\|X_{t,T}^{2,X_{r,s}^{Y_r}}\right\|_{L^{(2)}(\mathbb{R}^d,\mathbb{R}^d)}^{\frac{2p}{p-2(q+2)}}\right] < \infty, \tag{18}$$

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

$$\max\left\{\tfrac{\|f(x)\|_{\mathbb{R}^k}}{1+\|x\|_{\mathbb{R}^d}}, \|f'(x)\|_{L(\mathbb{R}^d,\mathbb{R}^k)}, \|f''(x)\|_{L^{(2)}(\mathbb{R}^d,\mathbb{R}^k)}\right\} \leq c(1+\|x\|_{\mathbb{R}^d}^q). \tag{19}$$

Then

(i) the stochastic process $\left(f'\big(X_{r,T}^{Y_r}\big)X_{r,T}^{1,Y_r}(\sigma(r,Y_r)-B_r)\right)_{r\in[0,T]}$ *is Skorohod-integrable and*

(ii) it holds $\mathbb{P}$*-a.s. that*

$$\begin{aligned} f\big(X_{0,T}^{Y_0}\big) - f(Y_T) &= \int_0^T f'\big(X_{r,T}^{Y_r}\big)X_{r,T}^{1,Y_r}\Big(\mu(r,Y_r)-A_r\Big)\,dr + \int_0^T f'\big(X_{r,T}^{Y_r}\big)X_{r,T}^{1,Y_r}\Big(\sigma(r,Y_r)-B_r\Big)\,\delta W_r \\ &+ \tfrac{1}{2}\sum_{i,j=1}^d \int_0^T \Big(\sigma(r,Y_r)[\sigma(r,Y_r)]^* - B_r[B_r]^*\Big)_{i,j}\Big(f''\big(X_{r,T}^{Y_r}\big)\big(X_{r,T}^{1,Y_r},X_{r,T}^{1,Y_r}\big) + f'\big(X_{r,T}^{Y_r}\big)X_{r,T}^{2,Y_r}\Big)\big(e_i^{(d)},e_j^{(d)}\big)\,dr. \end{aligned} \tag{20}$$

Proof of Theorem 3.1. The fact that for all $\omega \in \Omega$ the function $O \ni x \mapsto X_{T,T}^x(\omega) \in O$ is continuous and equation (15) imply that it holds $\mathbb{P}$-a.s. that $X_{T,T}^{Y_T} = Y_T$. Moreover, we rewrite the left-hand side of equation (20) as telescoping sum and obtain that for all $n \in \mathbb{N}$, $h \in \{\frac{T}{n}\}$ it holds $\mathbb{P}$-a.s. that

$$\begin{aligned} f\big(X_{0,T}^{Y_0}\big) - f\big(Y_T\big) &= f\big(X_{0h,T}^{Y_{0h}}\big) - f\big(X_{nh,T}^{Y_{nh}}\big) = \sum_{i=0}^{n-1}\Big(f\big(X_{ih,T}^{Y_{ih}}\big) - f\big(X_{(i+1)h,T}^{Y_{(i+1)h}}\big)\Big) \\ &= \sum_{i=0}^{n-1}\Big(f\big(X_{ih,T}^{Y_{ih}}\big) - f\big(X_{(i+1)h,T}^{Y_{ih}}\big)\Big) - \sum_{i=0}^{n-1}\Big(f\big(X_{(i+1)h,T}^{Y_{(i+1)h}}\big) - f\big(X_{(i+1)h,T}^{Y_{ih}}\big)\Big). \end{aligned} \tag{21}$$

First, we analyze the second sum on the right-hand side of equation (21)**.** For all $t \in [0,T]$, $x \in O$, $i \in \{1,2\}$ the functions $\Omega \ni \omega \mapsto X_{t,T}^x(\omega) \in O$, $\Omega \ni \omega \mapsto X_{t,T}^{i,x}(\omega) \in L^{(i)}(\mathbb{R}^d,\mathbb{R}^d)$ are $\mathfrak{S}(\mathcal{N} \cup \mathfrak{S}(W_s - W_t\colon s \in [t,T]))$-measurable. This together with the fact that for all $\omega \in \Omega$, $t \in [0,T]$ it holds that $\Big(O \ni x \mapsto f(X_{t,T}^x(\omega)) \in \mathbb{R}^k\Big) \in C^2(O,\mathbb{R}^k)$ implies that for all $t \in [0,T]$ the function $\Omega \ni \omega \mapsto \Big(O \ni x \mapsto f(X_{t,T}^x(\omega)) \in \mathbb{R}^k\Big) \in C^2(O,\mathbb{R}^k)$ is independent of the sigma-algebra $\mathbb{F}_t$. Itô's formula for independent random fields (e.g., Klenke [25, Theorem 25.30 and Remark 25.26]) (applied with the functions $\Omega \ni \omega \mapsto \Big(O \ni x \mapsto f(X_{(i+1)h,T}^x(\omega)) \in \mathbb{R}^k\Big) \in C^2(O,\mathbb{R}^k)$ for $n \in \mathbb{N}$, $i \in \{0,1,\ldots,n-1\}$, $h \in \{\frac{T}{n}\}$) yields 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

$$\begin{aligned} &f\big(X_{(i+1)h,T}^{Y_{(i+1)h}}\big) - f\big(X_{(i+1)h,T}^{Y_{ih}}\big) \\ &= \int_{ih}^{(i+1)h} \tfrac{\partial}{\partial x}\big(f(X_{(i+1)h,T}^x)\big)\big|_{x=Y_r}\,dY_r + \tfrac{1}{2}\sum_{l,j=1}^d \int_{ih}^{(i+1)h} \tfrac{\partial^2}{\partial x^2}\big(f(X_{(i+1)h,T}^x)\big)\big|_{x=Y_r}(e_l^{(d)},e_j^{(d)})\,d\,(\langle Y\rangle_r)_{l,j} \\ &= \int_{ih}^{(i+1)h} f'\big(X_{(i+1)h,T}^{Y_r}\big)X_{(i+1)h,T}^{1,Y_r}A_r\,dr + \int_{ih}^{(i+1)h} f'\big(X_{(i+1)h,T}^{Y_r}\big)X_{(i+1)h,T}^{1,Y_r}B_r\,dW_r \\ &\quad + \tfrac{1}{2}\sum_{l,j=1}^d \int_{ih}^{(i+1)h} (B_r[B_r]^*)_{l,j}\Big(f''\big(X_{(i+1)h,T}^{Y_r}\big)\big(X_{(i+1)h,T}^{1,Y_r},X_{(i+1)h,T}^{1,Y_r}\big) + f'\big(X_{(i+1)h,T}^{Y_r}\big)X_{(i+1)h,T}^{2,Y_r}\Big)\big(e_l^{(d)},e_j^{(d)}\big)\,dr. \end{aligned} \tag{22}$$

Inequalities (19) and (18) imply for all $i \in \{1,2\}$ that

$$\begin{aligned}\sup_{\substack{r,s,t\in[0,T]\\ r\le s\le t}} \Big\| f^{(i)}\big(X^{X_r}_{t,T}\big)\Big\|_{L^{\frac{p}{q}}(\mathbb{P};L^{(i)}(\mathbb{R}^d,\mathbb{R}^k))} &\le c \sup_{\substack{r,s,t\in[0,T]\\ r\le s\le t}} \Big\| 1 + \Big\| X^{X_r}_{t,T}\Big\|^q_{\mathbb{R}^d}\Big\|_{L^{\frac{p}{q}}(\mathbb{P};\mathbb{R})} \\ &\le c\left(1 + \sup_{\substack{r,s,t\in[0,T]\\ r\le s\le t}} \Big\| X^{X_r}_{t,T}\Big\|^q_{L^p(\mathbb{P};\mathbb{R}^d)}\right) < \infty.\end{aligned} \tag{23}$$

Hölder's inequality, inequalities (18), (23), and the assumption $B \in \mathcal{L}^p(\lambda_{[0,T]} \otimes \mathbb{P}; \mathbb{R}^{d\times m})$ imply that for all $n \in \mathbb{N}$, $i \in \{0,1,\dots,n-1\}$, $h \in \{\frac{T}{n}\}$ it holds that

$$\begin{aligned}&\Big\| f'\big(X^{Y_\cdot}_{(i+1)h,T}\big) X^{1,Y_\cdot}_{(i+1)h,T} B_\cdot\Big\|_{L^2(\mathbb{P};L^2(\lambda_{[ih,(i+1)h]};\mathbb{R}^{k\times m}))} \le \Big\| f'\big(X^{Y_\cdot}_{\lceil\cdot\rceil_h,T}\big) X^{1,Y_\cdot}_{\lceil\cdot\rceil_h,T} B_\cdot\Big\|_{L^2(\lambda_{[0,T]}\otimes\mathbb{P};\mathbb{R}^{k\times m})} \\ &\le \Big\| \big\| f'\big(X^{Y_\cdot}_{\lceil\cdot\rceil_h,T}\big)\big\|_{L(\mathbb{R}^d,\mathbb{R}^k)} \big\| X^{1,Y_\cdot}_{\lceil\cdot\rceil_h,T}\big\|_{L(\mathbb{R}^d,\mathbb{R}^d)} \|B_\cdot\|_{\mathrm{HS}(\mathbb{R}^m,\mathbb{R}^d)}\Big\|_{L^2(\lambda_{[0,T]}\otimes\mathbb{P};\mathbb{R})} \\ &\le \big\| f'\big(X^{Y_\cdot}_{\lceil\cdot\rceil_h,T}\big)\big\|_{L^{\frac{p}{q}}(\lambda_{[0,T]}\otimes\mathbb{P};L(\mathbb{R}^d,\mathbb{R}^k))} \big\| X^{1,Y_\cdot}_{\lceil\cdot\rceil_h,T}\big\|_{L^{\frac{2p}{p-2(q+1)}}(\lambda_{[0,T]}\otimes\mathbb{P};L(\mathbb{R}^d,\mathbb{R}^d))} \|B\|_{L^p(\lambda_{[0,T]}\otimes\mathbb{P};\mathbb{R}^{d\times m})} \\ &\le T^{\frac{p-2}{2p}} \Big(\sup_{(r,s)\in\Delta_T} \big\| f'\big(X^{Y_r}_{s,T}\big)\big\|_{L^{\frac{p}{q}}(\mathbb{P};L(\mathbb{R}^d,\mathbb{R}^k))}\Big) \Big(\sup_{(r,s)\in\Delta_T} \big\| X^{1,Y_r}_{s,T}\big\|_{L^{\frac{2p}{p-2(q+1)}}(\mathbb{P};L(\mathbb{R}^d,\mathbb{R}^d))}\Big) \|B\|_{L^p(\lambda_{[0,T]}\otimes\mathbb{P};\mathbb{R}^{d\times m})} < \infty.\end{aligned} \tag{24}$$

For all $n \in \mathbb{N}$, $i \in \{0,1,\dots,n-1\}$, $h \in \{\frac{T}{n}\}$ the stochastic process $\big(f'\big(X^{Y_r}_{(i+1)h,T}\big) X^{1,Y_r}_{(i+1)h,T} B_r\big)_{r\in[ih,(i+1)h]}$ is predictable with respect to the filtration

$$\big(\mathfrak{S}\big(\mathbb{F}_r \cup \mathfrak{S}\big(\{W_s - W_{(i+1)h} \colon s \in [(i+1)h, T]\}\big)\big)\big)_{r\in[ih,(i+1)h]}. \tag{25}$$

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 $(f'\big(X^{Y_r}_{\lceil r\rceil_h,T}\big) X^{1,Y_r}_{\lceil r\rceil_h,T} B_r)_{r\in[0,T]}$ is Skorohod integrable and that for all $n \in \mathbb{N}$, $h \in \{\frac{T}{n}\}$ it holds $\mathbb{P}$-a.s. that

$$\begin{aligned}&\sum_{i=0}^{n-1} \int_{ih}^{(i+1)h} f'\big(X^{Y_r}_{(i+1)h,T}\big) X^{1,Y_r}_{(i+1)h,T} B_r \, dW_r \\ &= \sum_{i=0}^{n-1} \int_{ih}^{(i+1)h} f'\big(X^{Y_r}_{(i+1)h,T}\big) X^{1,Y_r}_{(i+1)h,T} B_r \, \delta W_r^{\mathfrak{S}(\mathbb{F}_{ih}\cup\mathfrak{S}(\{W_s - W_{(i+1)h} \colon s\in[(i+1)h,T]\}))} \\ &= \sum_{i=0}^{n-1} \int_0^T \mathbb{1}_{[ih,(i+1)h]}(r) f'\big(X^{Y_r}_{\lceil r\rceil_h,T}\big) X^{1,Y_r}_{\lceil r\rceil_h,T} B_r \, \delta W_r^{\mathbb{F}_0} \\ &= \int_0^T f'\big(X^{Y_r}_{\lceil r\rceil_h,T}\big) X^{1,Y_r}_{\lceil r\rceil_h,T} B_r \, \delta W_r.\end{aligned} \tag{26}$$

Equations (22) and (26) imply that for all $n \in \mathbb{N}$, $h \in \{\frac{T}{n}\}$ it holds $\mathbb{P}$-a.s. that

$$\begin{aligned}&\sum_{i=0}^{n-1} \Big(f\big(X^{Y_{(i+1)h}}_{(i+1)h,T}\big) - f\big(X^{Y_{ih}}_{(i+1)h,T}\big)\Big) \\ &= \int_0^T f'\big(X^{Y_r}_{\lceil r\rceil_h,T}\big) X^{1,Y_r}_{\lceil r\rceil_h,T} A_r \, dr + \int_0^T f'\big(X^{Y_r}_{\lceil r\rceil_h,T}\big) X^{1,Y_r}_{\lceil r\rceil_h,T} B_r \, \delta W_r \\ &\quad + \tfrac{1}{2} \sum_{l,j=1}^d \int_0^T (B_r[B_r]^*)_{l,j} \Big(f''\big(X^{Y_r}_{\lceil r\rceil_h,T}\big)\big(X^{1,Y_r}_{\lceil r\rceil_h,T}, X^{1,Y_r}_{\lceil r\rceil_h,T}\big) + f'\big(X^{Y_r}_{\lceil r\rceil_h,T}\big) X^{2,Y_r}_{\lceil r\rceil_h,T}\Big)(e_l^{(d)}, e_j^{(d)}) \, dr.\end{aligned} \tag{27}$$

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 $\mathbb{P}\big(X^x_{s,T} = X^{X^x_{s,t}}_{t,T}\big) = 1$. This and the fact that X is a continuous random field imply for all $(s,t) \in \Delta_T$ that $\mathbb{P}\Big(X^{Y_s}_{s,T} = X^{X^{Y_s}_{s,t}}_{t,T}\Big) = 1$. For all $t \in [0,T]$, $x \in O$, $i \in \{1,2\}$ the functions $\Omega \ni \omega \mapsto X^x_{t,T}(\omega) \in O$, $\Omega \ni \omega \mapsto X^{i,x}_{t,T}(\omega) \in L^{(i)}(\mathbb{R}^d, \mathbb{R}^d)$ are $\mathfrak{S}(\mathcal{N} \cup \mathfrak{S}(W_s - W_t \colon s \in [t,T]))$-measurable. This together with the fact that for all $\omega \in \Omega$, $t \in [0,T]$ it holds that $\Big(O \ni x \mapsto f(X^x_{t,T}(\omega)) \in \mathbb{R}^k\Big) \in C^2(O, \mathbb{R}^k)$ implies that for all $t \in [0,T]$ the function $\Omega \ni \omega \mapsto \Big(O \ni x \mapsto f(X^x_{t,T}(\omega)) \in \mathbb{R}^k\Big) \in C^2(O, \mathbb{R}^k)$ is independent of the sigma-algebra $\mathbb{F}_t$. Itô's formula for independent random fields (e.g., Klenke [25, Theorem 25.30 and Remark 25.26]) (applied with the functions $\Omega \ni \omega \mapsto \Big(O \ni x \mapsto f(X^x_{(i+1)h,T}(\omega)) \in \mathbb{R}^k\Big) \in C^2(O, \mathbb{R}^k)$ for $n \in \mathbb{N}$, $i \in \{0,1,\ldots,n-1\}$, $h \in \{\frac{T}{n}\}$) yields 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

$$
\begin{aligned}
&f\big(X^{Y_{ih}}_{ih,T}\big) - f\big(X^{Y_{ih}}_{(i+1)h,T}\big) = f\Big(X^{X^{Y_{ih}}_{ih,(i+1)h}}_{(i+1)h,T}\Big) - f\big(X^{Y_{ih}}_{(i+1)h,T}\big) \\
&= \int_{ih}^{(i+1)h} f'\Big(X^{X^{Y_{ih}}_{ih,r}}_{(i+1)h,T}\Big) X^{1,X^{Y_{ih}}_{ih,r}}_{(i+1)h,T}\, dX^{Y_{ih}}_{ih,r} \\
&\quad + \tfrac{1}{2} \sum_{l,j=1}^{d} \int_{ih}^{(i+1)h} \left(f''\Big(X^{X^{Y_{ih}}_{ih,r}}_{(i+1)h,T}\Big)\Big(X^{1,X^{Y_{ih}}_{ih,r}}_{(i+1)h,T}, X^{1,X^{Y_{ih}}_{ih,r}}_{(i+1)h,T}\Big) + f'\Big(X^{X^{Y_{ih}}_{ih,r}}_{(i+1)h,T}\Big) X^{2,X^{Y_{ih}}_{ih,r}}_{(i+1)h,T} \right)\big(e^{(d)}_l, e^{(d)}_j\big)\, d\big(\langle X^{Y_{ih}}_{ih,\cdot}\rangle_r\big)_{l,j} \\
&= \int_{ih}^{(i+1)h} f'\Big(X^{X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}}_{\lceil r\rceil_h,T}\Big) X^{1,X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}}_{\lceil r\rceil_h,T} \mu\big(r, X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}\big)\, dr + \int_{ih}^{(i+1)h} f'\Big(X^{X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}}_{\lceil r\rceil_h,T}\Big) X^{1,X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}}_{\lceil r\rceil_h,T} \sigma\big(r, X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}\big)\, dW_r \\
&\quad + \tfrac{1}{2} \sum_{l,j=1}^{d} \int_{ih}^{(i+1)h} \Big(\sigma\Big(r, X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}\Big)\Big[\sigma\Big(r, X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}\Big)\Big]^*\Big)_{l,j} \\
&\qquad \cdot \left(f''\Big(X^{X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}}_{\lceil r\rceil_h,T}\Big)\Big(X^{1,X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}}_{\lceil r\rceil_h,T}, X^{1,X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}}_{\lceil r\rceil_h,T}\Big) + f'\Big(X^{X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}}_{\lceil r\rceil_h,T}\Big) X^{2,X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}}_{\lceil r\rceil_h,T} \right)\big(e^{(d)}_l, e^{(d)}_j\big)\, dr.
\end{aligned}
\tag{28}
$$

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{aligned}
&\left\| f'\left(X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot,T} \right) X^{1,X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}}_{\lceil\cdot\rceil_h,T} \sigma\left(\cdot, X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}\right) \right\|_{L^2(\mathbb{P};L^2(\lambda_{[ih,(i+1)h]};\mathbb{R}^{k\times m}))} \\
&\leq \left\| f'\left(X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot,T} \right) X^{1,X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}}_{\lceil\cdot\rceil_h,T} \sigma\left(\cdot, X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}\right) \right\|_{L^2(\lambda_{[0,T]}\otimes\mathbb{P};\mathbb{R}^{k\times m})} \\
&\leq \left\| \left\| \left\| f'\left(X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot,T} \right) \right\|_{L(\mathbb{R}^d,\mathbb{R}^k)} \left\| X^{1,X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}}_{\lceil\cdot\rceil_h,T} \right\|_{L(\mathbb{R}^d,\mathbb{R}^d)} \left\| \sigma\left(\cdot, X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}\right) \right\|_{\mathrm{HS}(\mathbb{R}^m,\mathbb{R}^d)} \right\|_{L^2(\lambda_{[0,T]}\otimes\mathbb{P};\mathbb{R})} \\
&\leq \left\| f'\left(X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot,T} \right) \right\|_{L^{\frac{p}{q}}(\lambda_{[0,T]}\otimes\mathbb{P};L(\mathbb{R}^d,\mathbb{R}^k))} \left\| X^{1,X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}}_{\lceil\cdot\rceil_h,T} \right\|_{L^{\frac{2p}{p-2(q+1)}}(\lambda_{[0,T]}\otimes\mathbb{P};L(\mathbb{R}^d,\mathbb{R}^d))} \left\| \sigma\left(\cdot, X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}\right) \right\|_{L^p(\lambda_{[0,T]}\otimes\mathbb{P};\mathbb{R}^{d\times m})} \\
&\leq T^{\frac{p-2}{2p}} \left(\sup_{\substack{r,s,t\in[0,T]\\ r\leq s\leq t}} \left\| f'\left(X^{X^{Y_r}_{r,s}}_{t,T} \right) \right\|_{L^{\frac{p}{q}}(\mathbb{P};L(\mathbb{R}^d,\mathbb{R}^k))} \right) \left(\sup_{\substack{r,s,t\in[0,T]\\ r\leq s\leq t}} \left\| X^{1,X^{Y_r}_{r,s}}_{t,T} \right\|_{L^{\frac{2p}{p-2(q+1)}}(\mathbb{P};L(\mathbb{R}^d,\mathbb{R}^d))} \right) \\
&\qquad \cdot \left(\sup_{\kappa\in T/N} \left\| \sigma\left(\cdot, X^{Y_{\lfloor\cdot\rfloor_\kappa}}_{\lfloor\cdot\rfloor_\kappa,\cdot}\right) \right\|_{L^p(\lambda_{[0,T]}\otimes\mathbb{P};\mathbb{R}^{d\times m})} \right) < \infty.
\end{aligned} \tag{29}
$$

For all $n \in \mathbb{N}$, $i \in \{0, 1, \ldots, n-1\}$, $h \in \{\frac{T}{n}\}$ the process $\left(f'\left(X^{Y_{ih}}_{ih,r,(i+1)h,T} \right) X^{1,X^{Y_{ih}}_{ih,r}}_{(i+1)h,T} \sigma\left(r, X^{Y_{ih}}_{ih,r}\right) \right)_{r\in[ih,(i+1)h]}$ is predictable with respect to the filtration (25). Proposition A.8 together with inequality (29), Proposition A.7, and linearity of the Skorohod integral assert that the process $f'\left(X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot,T} \right) X^{1,X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}}_{\lceil\cdot\rceil_h,T} \sigma\left(\cdot, X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}\right)$ is Skorohod integrable and that for all $n \in \mathbb{N}$, $h \in \{\frac{T}{n}\}$ it holds $\mathbb{P}$-a.s. that

$$
\begin{aligned}
&\sum_{i=0}^{n-1} \int_{ih}^{(i+1)h} f'\left(X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r,\lceil r\rceil_h,T} \right) X^{1,X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}}_{\lceil r\rceil_h,T} \sigma\left(r, X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}\right) dW_r \\
&= \sum_{i=0}^{n-1} \int_{ih}^{(i+1)h} f'\left(X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r,\lceil r\rceil_h,T} \right) X^{1,X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}}_{\lceil r\rceil_h,T} \sigma\left(r, X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}\right) \delta W_r^{\mathfrak{S}(\mathbb{F}_{ih}\cup\mathfrak{S}(W_s - W_{(i+1)h}\colon s\in[(i+1)h,T]))} \\
&= \sum_{i=0}^{n-1} \int_0^T \mathbb{1}_{[ih,(i+1)h]}(r) f'\left(X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r,\lceil r\rceil_h,T} \right) X^{1,X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}}_{\lceil r\rceil_h,T} \sigma\left(r, X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}\right) \delta W_r^{\mathbb{F}_0} \\
&= \int_0^T f'\left(X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r,\lceil r\rceil_h,T} \right) X^{1,X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}}_{\lceil r\rceil_h,T} \sigma\left(r, X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}\right) \delta W_r.
\end{aligned} \tag{30}
$$

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

$$\begin{aligned}
&\sum_{i=0}^{n-1} \Big(f\big(X_{ih,T}^{Y_{ih}}\big) - f\big(X_{(i+1)h,T}^{Y_{ih}}\big) \Big) \\
&= \int_0^T f'\Big(X_{\lceil r \rceil_h, T}^{X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}} \Big) X_{\lceil r \rceil_h, T}^{1, X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}} \mu\Big(r, X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}} \Big)\, dr + \int_0^T f'\Big(X_{\lceil r \rceil_h, T}^{X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}} \Big) X_{\lceil r \rceil_h, T}^{1, X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}} \sigma\Big(r, X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}} \Big)\, \delta W_r \\
&+ \tfrac{1}{2} \sum_{l,j=1}^{d} \int_0^T \Big(\sigma\big(r, X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}\big) \big[\sigma\big(r, X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}\big)\big]^* \Big)_{l,j} \\
&\qquad \cdot \left(f''\Big(X_{\lceil r \rceil_h, T}^{X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}} \Big) \Big(X_{\lceil r \rceil_h, T}^{1, X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}}, X_{\lceil r \rceil_h, T}^{1, X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}} \Big) + f'\Big(X_{\lceil r \rceil_h, T}^{X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}} \Big) X_{\lceil r \rceil_h, T}^{2, X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}} \right) \big(e_l^{(d)}, e_j^{(d)}\big)\, dr.
\end{aligned} \tag{31}$$

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

$$\begin{aligned}
&f\big(X_{0,T}^{Y_0}\big) - f\big(Y_T\big) \\
&= \int_0^T f'\Big(X_{\lceil r \rceil_h, T}^{X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}} \Big) X_{\lceil r \rceil_h, T}^{1, X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}} \mu\big(r, X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}} \big) - f'\big(X_{\lceil r \rceil_h, T}^{Y_r} \big) X_{\lceil r \rceil_h, T}^{1, Y_r} A_r\, dr \\
&+ \int_0^T f'\Big(X_{\lceil r \rceil_h, T}^{X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}} \Big) X_{\lceil r \rceil_h, T}^{1, X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}} \sigma\Big(r, X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}} \Big) - f'\big(X_{\lceil r \rceil_h, T}^{Y_r} \big) X_{\lceil r \rceil_h, T}^{1, Y_r} B_r\, \delta W_r \\
&+ \tfrac{1}{2} \sum_{l,j=1}^{d} \int_0^T \Big(\sigma\big(r, X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}\big) \big[\sigma\big(r, X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}\big)\big]^* \Big)_{l,j} \\
&\qquad \cdot \left(f''\Big(X_{\lceil r \rceil_h, T}^{X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}} \Big) \Big(X_{\lceil r \rceil_h, T}^{1, X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}}, X_{\lceil r \rceil_h, T}^{1, X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}} \Big) + f'\Big(X_{\lceil r \rceil_h, T}^{X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}} \Big) X_{\lceil r \rceil_h, T}^{2, X_{\lfloor r \rfloor_h, r}^{Y_{\lfloor r \rfloor_h}}} \right) \big(e_l^{(d)}, e_j^{(d)}\big)\, dr \\
&- \tfrac{1}{2} \sum_{l,j=1}^{d} \int_0^T \big(B_r [B_r]^* \big)_{l,j} \Big(f''\big(X_{\lceil r \rceil_h, T}^{Y_r} \big) \big(X_{\lceil r \rceil_h, T}^{1, Y_r}, X_{\lceil r \rceil_h, T}^{1, Y_r} \big) + f'\big(X_{\lceil r \rceil_h, T}^{Y_r} \big) X_{\lceil r \rceil_h, T}^{2, Y_r} \Big) \big(e_l^{(d)}, e_j^{(d)}\big)\, dr.
\end{aligned} \tag{32}$$

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

$$\begin{aligned}
&\sup_{h \in {}^{T}\!/_{\mathbb{N}}} \left\| f'\Big(X_{\lceil \cdot \rceil_h, T}^{X_{\lfloor \cdot \rfloor_h, \cdot}^{Y_{\lfloor \cdot \rfloor_h}}} \Big) X_{\lceil \cdot \rceil_h, T}^{1, X_{\lfloor \cdot \rfloor_h, \cdot}^{Y_{\lfloor \cdot \rfloor_h}}} \mu\Big(\cdot, X_{\lfloor \cdot \rfloor_h, \cdot}^{Y_{\lfloor \cdot \rfloor_h}} \Big) - f'\big(X_{\lceil \cdot \rceil_h, T}^{Y_\cdot} \big) X_{\lceil \cdot \rceil_h, T}^{1, Y_\cdot} A_\cdot \right\|_{L^2(\lambda_{[0,T]} \otimes \mathbb{P}; \mathbb{R}^k)} \\
&\qquad \le T^{\frac{p-2}{2p}} \Big(\sup_{\substack{r,s,t \in [0,T] \\ r \le s \le t}} \Big\| f'\Big(X_{t,T}^{X_{r,s}^{Y_r}} \Big) \Big\|_{L^{\frac{p}{q}}(\mathbb{P}; L(\mathbb{R}^d, \mathbb{R}^k))} \Big) \Big(\sup_{\substack{r,s,t \in [0,T] \\ r \le s \le t}} \Big\| X_{t,T}^{1, X_{r,s}^{Y_r}} \Big\|_{L^{\frac{2p}{p-2(q+1)}}(\mathbb{P}; L(\mathbb{R}^d, \mathbb{R}^d))} \Big) \\
&\qquad\qquad \cdot \Big(\sup_{h \in {}^{T}\!/_{\mathbb{N}}} \Big\| \mu\big(\cdot, X_{\lfloor \cdot \rfloor_h, \cdot}^{Y_{\lfloor \cdot \rfloor_h}} \big) \Big\|_{L^p(\lambda_{[0,T]} \otimes \mathbb{P}; \mathbb{R}^d)} + \|A\|_{L^p(\lambda_{[0,T]} \otimes \mathbb{P}; \mathbb{R}^d)} \Big) < \infty.
\end{aligned} \tag{33}$$

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

$$
\begin{aligned}
&\sup_{h \in T/\mathbb{N}} \left\| \left(f''\left(X^{Y_{\lfloor \cdot \rfloor_h}}_{\lfloor \cdot \rfloor_h, \cdot, T} \right) \left(X^{1, X^{Y_{\lfloor \cdot \rfloor_h}}_{\lfloor \cdot \rfloor_h, \cdot}}_{\lceil \cdot \rceil_h, T}, X^{1, X^{Y_{\lfloor \cdot \rfloor_h}}_{\lfloor \cdot \rfloor_h, \cdot}}_{\lceil \cdot \rceil_h, T} \right) + f'\left(X^{Y_{\lfloor \cdot \rfloor_h}}_{\lfloor \cdot \rfloor_h, \cdot, T} \right) X^{2, X^{Y_{\lfloor \cdot \rfloor_h}}_{\lfloor \cdot \rfloor_h, \cdot}}_{\lceil \cdot \rceil_h, T} \right) \left(e^{(d)}_l, e^{(d)}_j \right) \right\|_{L^{\frac{2p}{p-4}}(\lambda_{[0,T]} \otimes \mathbb{P}; \mathbb{R}^k)} \\
&\leq \sup_{h \in T/\mathbb{N}} \left\| \left\| f''\left(X^{Y_{\lfloor \cdot \rfloor_h}}_{\lfloor \cdot \rfloor_h, \cdot, T} \right) \right\|_{L^{(2)}(\mathbb{R}^d, \mathbb{R}^k)} \left\| X^{1, X^{Y_{\lfloor \cdot \rfloor_h}}_{\lfloor \cdot \rfloor_h, \cdot}}_{\lceil \cdot \rceil_h, T} \right\|^2_{L(\mathbb{R}^d, \mathbb{R}^d)} \right. \\
&\qquad \left. + \left\| f'\left(X^{Y_{\lfloor \cdot \rfloor_h}}_{\lfloor \cdot \rfloor_h, \cdot, T} \right) \right\|_{L(\mathbb{R}^d, \mathbb{R}^k)} \left\| X^{2, X^{Y_{\lfloor \cdot \rfloor_h}}_{\lfloor \cdot \rfloor_h, \cdot}}_{\lceil \cdot \rceil_h, T} \right\|_{L^{(2)}(\mathbb{R}^d, \mathbb{R}^d)} \right\|_{L^{\frac{2p}{p-4}}(\lambda_{[0,T]} \otimes \mathbb{P}; \mathbb{R})} \\
&\leq \sup_{h \in T/\mathbb{N}} \left(\left\| f''\left(X^{Y_{\lfloor \cdot \rfloor_h}}_{\lfloor \cdot \rfloor_h, \cdot, T} \right) \right\|_{L^{\frac{p}{q}}(\lambda_{[0,T]} \otimes \mathbb{P}; L^{(2)}(\mathbb{R}^d, \mathbb{R}^k))} \left\| X^{1, X^{Y_{\lfloor \cdot \rfloor_h}}_{\lfloor \cdot \rfloor_h, \cdot}}_{\lceil \cdot \rceil_h, T} \right\|^2_{L^{\frac{4p}{p-2(q+2)}}(\lambda_{[0,T]} \otimes \mathbb{P}; L(\mathbb{R}^d, \mathbb{R}^d))} \right. \\
&\qquad \left. + \left\| f'\left(X^{Y_{\lfloor \cdot \rfloor_h}}_{\lfloor \cdot \rfloor_h, \cdot, T} \right) \right\|_{L^{\frac{p}{q}}(\lambda_{[0,T]} \otimes \mathbb{P}; L(\mathbb{R}^d, \mathbb{R}^k))} \left\| X^{2, X^{Y_{\lfloor \cdot \rfloor_h}}_{\lfloor \cdot \rfloor_h, \cdot}}_{\lceil \cdot \rceil_h, T} \right\|_{L^{\frac{2p}{p-2(q+2)}}(\lambda_{[0,T]} \otimes \mathbb{P}; L^{(2)}(\mathbb{R}^d, \mathbb{R}^d))} \right) \\
&\leq T^{\frac{p-4}{2p}} \left(\left(\sup_{\substack{r,s,t \in [0,T] \\ r \leq s \leq t}} \left\| f''\left(X^{X^{Y_r}_{r,s}}_{t,T} \right) \right\|_{L^{\frac{p}{q}}(\mathbb{P}; L^{(2)}(\mathbb{R}^d, \mathbb{R}^k))} \right) \left(\sup_{\substack{r,s,t \in [0,T] \\ r \leq s \leq t}} \left\| X^{1, X^{Y_r}_{r,s}}_{t,T} \right\|^2_{L^{\frac{4p}{p-2(q+2)}}(\mathbb{P}; L(\mathbb{R}^d, \mathbb{R}^d))} \right) \right. \\
&\qquad \left. + \left(\sup_{\substack{r,s,t \in [0,T] \\ r \leq s \leq t}} \left\| f'\left(X^{X^{Y_r}_{r,s}}_{t,T} \right) \right\|_{L^{\frac{p}{q}}(\mathbb{P}; L(\mathbb{R}^d, \mathbb{R}^k))} \right) \left(\sup_{\substack{r,s,t \in [0,T] \\ r \leq s \leq t}} \left\| X^{2, X^{Y_r}_{r,s}}_{t,T} \right\|_{L^{\frac{2p}{p-2(q+2)}}(\mathbb{P}; L^{(2)}(\mathbb{R}^d, \mathbb{R}^d))} \right) \right) < \infty
\end{aligned}
\tag{34}
$$

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

$$
\sup_{h \in T/\mathbb{N}} \left\| \left(f''(X^{Y_{\lfloor \cdot \rfloor_h}}_{\lceil \cdot \rceil_h, T}) \left(X^{1, Y_{\lfloor \cdot \rfloor_h}}_{\lceil \cdot \rceil_h, T}, X^{1, Y_{\lfloor \cdot \rfloor_h}}_{\lceil \cdot \rceil_h, T} \right) + f'(X^{Y_{\lfloor \cdot \rfloor_h}}_{\lceil \cdot \rceil_h, T}) X^{2, Y_{\lfloor \cdot \rfloor_h}}_{\lceil \cdot \rceil_h, T} \right) \left(e^{(d)}_i, e^{(d)}_j \right) \right\|_{L^{\frac{2p}{p-4}}(\lambda_{[0,T]} \otimes \mathbb{P}; \mathbb{R}^k)} < \infty. \tag{35}
$$

The fact that for all $C \in \mathbb{R}^{d \times m}$ it holds that $\sum_{i,j=1}^d |(CC^*)_{i,j}| \leq d \|C\|^2_{\mathrm{HS}(\mathbb{R}^m, \mathbb{R}^d)}$, Hölder's inequality, assump-

tion (17) and inequality (34) imply that

$$
\begin{aligned}
&\sup_{h\in T/\mathbb{N}} \tfrac{1}{2}\Bigg\| \sum_{l,j=1}^{d} \left(\sigma\Big(\cdot, X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}\Big)\Big[\sigma\Big(\cdot, X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}\Big)\Big]^*\right)_{l,j} \\
&\quad\cdot \left(f''\left(X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}\right)\left(X^{1,X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}}_{\lceil\cdot\rceil_h,T}, X^{1,X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}}_{\lceil\cdot\rceil_h,T}\right) + f'\left(X^{X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}}_{\lceil\cdot\rceil_h,T}\right) X^{2,X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}}_{\lceil\cdot\rceil_h,T}\right)\big(e^{(d)}_l, e^{(d)}_j\big)\Bigg\|_{L^2(\lambda_{[0,T]}\otimes\mathbb{P};\mathbb{R}^k)} \\
&\le \sup_{h\in T/\mathbb{N}} \tfrac{1}{2}\Bigg\| d\left\|\sigma\Big(\cdot, X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}\Big)\right\|^2_{\mathrm{HS}(\mathbb{R}^m,\mathbb{R}^d))} \\
&\quad\cdot \sum_{l,j=1}^{d} \left| f''\left(X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}\right)\left(X^{1,X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}}_{\lceil\cdot\rceil_h,T}, X^{1,X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}}_{\lceil\cdot\rceil_h,T}\right) + f'\left(X^{X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}}_{\lceil\cdot\rceil_h,T}\right) X^{2,X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}}_{\lceil\cdot\rceil_h,T}\right| \big(e^{(d)}_l, e^{(d)}_j\big)\Bigg\|_{L^2(\lambda_{[0,T]}\otimes\mathbb{P};\mathbb{R}^k)} \\
&\le \sup_{h\in T/\mathbb{N}} \tfrac{d}{2}\Bigg(\left\|\sigma\Big(\cdot, X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}\Big)\right\|^2_{L^p(\lambda_{[0,T]}\otimes\mathbb{P};\mathbb{R}^{d\times m})} \\
&\quad\cdot \sum_{l,j\in\{1,\dots,d\}} \left\| f''\left(X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}\right)\left(X^{1,X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}}_{\lceil\cdot\rceil_h,T}, X^{1,X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}}_{\lceil\cdot\rceil_h,T}\right) + f'\left(X^{X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}}_{\lceil\cdot\rceil_h,T}\right) X^{2,X^{Y_{\lfloor\cdot\rfloor_h}}_{\lfloor\cdot\rfloor_h,\cdot}}_{\lceil\cdot\rceil_h,T}\right\| \big(e^{(d)}_l, e^{(d)}_j\big)\Bigg\|_{L^{\frac{2p}{p-4}}(\lambda_{[0,T]}\otimes\mathbb{P};\mathbb{R}^k)}\Bigg) \\
&< \infty.
\end{aligned}
\tag{36}
$$

Analogously, the fact that for all $C \in \mathbb{R}^{d\times m}$ it holds that $\sum_{i,j=1}^{d} |(CC^*)_{i,j}| \le d\|C\|^2_{\mathrm{HS}(\mathbb{R}^m,\mathbb{R}^d)}$, Hölder's inequality, the assumption $B \in \mathcal{L}^p(\lambda_{[0,T]}\otimes\mathbb{P};\mathbb{R}^{d\times m})$, and inequality (35) yield that

$$
\begin{aligned}
&\sup_{h\in T/\mathbb{N}} \tfrac{1}{2}\Bigg\| \sum_{l,j=1}^{d} (B_\cdot[B_\cdot]^*)_{l,j}\left(f''(X^{Y_\cdot}_{\lceil\cdot\rceil_h,T})\big(X^{1,Y_\cdot}_{\lceil\cdot\rceil_h,T}, X^{1,Y_\cdot}_{\lceil\cdot\rceil_h,T}\big) + f'(X^{Y_\cdot}_{\lceil\cdot\rceil_h,T}) X^{2,Y_\cdot}_{\lceil\cdot\rceil_h,T}\right)\big(e^{(d)}_l, e^{(d)}_j\big)\Bigg\|_{L^2(\lambda_{[0,T]}\otimes\mathbb{P};\mathbb{R}^k)} \\
&\le \frac{d}{2}\|B\|^2_{L^p(\lambda_{[0,T]}\otimes\mathbb{P};\mathbb{R}^{d\times m})} \\
&\quad\cdot \sum_{l,j\in\{1,\dots,d\}} \sup_{h\in T/\mathbb{N}} \Bigg\| \left(f''(X^{Y_\cdot}_{\lceil\cdot\rceil_h,T})\big(X^{1,Y_\cdot}_{\lceil\cdot\rceil_h,T}, X^{1,Y_\cdot}_{\lceil\cdot\rceil_h,T}\big) + f'(X^{Y_\cdot}_{\lceil\cdot\rceil_h,T}) X^{2,Y_\cdot}_{\lceil\cdot\rceil_h,T}\right)\big(e^{(d)}_l, e^{(d)}_j\big)\Bigg\|_{L^{\frac{2p}{p-4}}(\lambda_{[0,T]}\otimes\mathbb{P};\mathbb{R}^k)} < \infty.
\end{aligned}
\tag{37}
$$

Next Klenke [25, 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 $\Delta_T \times O \ni (s,t,x) \mapsto X^x_{s,t} \in O$, and $\inf_{r\in[0,T]} \mathbb{P}(X^{Y_r}_{r,r} =$

$Y_r) = 1$ imply that

$$
\begin{aligned}
&\lim_{T/\mathbb{N}\ni h\searrow 0} \Bigg\| \int_0^T f'\Bigg(X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}_{\lceil r\rceil_h,T} \Bigg) X^{1,X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}}_{\lceil r\rceil_h,T} \mu(r, X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}) - f'\big(X^{Y_r}_{\lceil r\rceil_h,T}\big) X^{1,Y_r}_{\lceil r\rceil_h,T} A_r \, dr \\
&- \int_0^T f'\big(X^{Y_r}_{r,T}\big) X^{1,Y_r}_{r,T} \Big(\mu\Big(r, Y_r\Big) - A_r\Big) \, dr \\
&+ \tfrac{1}{2} \sum_{l,j=1}^{d} \int_0^T \Big(\sigma\big(r, X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}\big) \big[\sigma\big(r, X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}\big)\big]^*\Big)_{l,j} \\
&\quad \cdot \Bigg(f''\Bigg(X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}_{\lceil r\rceil_h,T} \Bigg) \Bigg(X^{1,X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}}_{\lceil r\rceil_h,T}, X^{1,X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}}_{\lceil r\rceil_h,T} \Bigg) + f'\Bigg(X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}_{\lceil r\rceil_h,T} \Bigg) X^{2,X^{Y_{\lfloor r\rfloor_h}}_{\lfloor r\rfloor_h,r}}_{\lceil r\rceil_h,T} \Bigg) (e_l^{(d)}, e_j^{(d)}) \, dr \\
&- \tfrac{1}{2} \sum_{l,j=1}^{d} \int_0^T (B_r[B_r]^*)_{l,j} \Big(f''\big(X^{Y_r}_{\lceil r\rceil_h,T}\big) \big(X^{1,Y_r}_{\lceil r\rceil_h,T}, X^{1,Y_r}_{\lceil r\rceil_h,T}\big) + f'\big(X^{Y_r}_{\lceil r\rceil_h,T}\big) X^{2,Y_r}_{\lceil r\rceil_h,T} \Big) (e_l^{(d)}, e_j^{(d)}) \, dr \\
&- \tfrac{1}{2} \sum_{l,j=1}^{d} \int_0^T (\sigma(r, Y_r)[\sigma(r, Y_r)]^* - B_r[B_r]^*)_{l,j} \Big(f''\big(X^{Y_r}_{r,T}\big) \big(X^{1,Y_r}_{r,T}, X^{1,Y_r}_{r,T}\big) + f'\big(X^{Y_r}_{r,T}\big) X^{2,Y_r}_{r,T} \Big) (e_l^{(d)}, e_j^{(d)}) \, dr \Bigg\|_{L^1(\mathbb{P};\mathbb{R}^k)} \\
&= 0.
\end{aligned}
\tag{38}
$$

Inequality (19) implies that for all $x, y \in O$ it holds that

$$
\|f(x) - f(y)\|_{\mathbb{R}^k} \leq \|f(x)\|_{\mathbb{R}^k} + \|f(y)\|_{\mathbb{R}^k} \leq c(1 + \|x\|_{\mathbb{R}^d})(1 + \|x\|_{\mathbb{R}^d})^q + c(1 + \|y\|_{\mathbb{R}^d})(1 + \|y\|_{\mathbb{R}^d})^q. \tag{39}
$$

Inequality (39), Hölder's inequality, the fact that $2q + 2 < p$, the fact that $\mathbb{P}\Big(X^{Y_0}_{0,T} = X^{X^{Y_0}_{0,0}}_{0,T}\Big) = 1 = \mathbb{P}\Big(Y_T = X^{X^{Y_T}_{T,T}}_{T,T}\Big)$, and inequality (18) show that

$$
\begin{aligned}
&\big\|f\big(X^{Y_0}_{0,T}\big) - f\big(Y_T\big)\big\|_{L^2(\mathbb{P};\mathbb{R}^k)} \\
&\leq c\big\|\big(1 + \big\|X^{Y_0}_{0,T}\big\|_{\mathbb{R}^d}\big)^{1+q}\big\|_{L^2(\mathbb{P};\mathbb{R})} + c\big\|\big(1 + \|Y_T\|_{\mathbb{R}^d}\big)^{1+q}\big\|_{L^2(\mathbb{P};\mathbb{R})} \\
&\leq c\big(1 + \big\|X^{Y_0}_{0,T}\big\|_{L^{2q+2}(\mathbb{P};\mathbb{R}^d)}\big)^{q+1} + c\big(1 + \|Y_T\|_{L^{2q+2}(\mathbb{P};\mathbb{R}^d)}\big)^{q+1} \\
&\leq \sup_{\substack{r,s,t\in[0,T] \\ r\leq s\leq t}} 2c\Big(1 + \Big\|X^{Y_r}_{r,s}_{t,T}\Big\|_{L^p(\mathbb{P};\mathbb{R}^d)}\Big)^{q+1} < \infty.
\end{aligned}
\tag{40}
$$

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

$$
\begin{aligned}
&\left\| \int_0^T f'\left(X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r, \lceil r \rceil_h, T} \right) X^{1, X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r}}_{\lceil r \rceil_h, T} \sigma\Big(r, X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r}\Big) - f'\Big(X^{Y_r}_{\lceil r \rceil_h, T}\Big) X^{1, Y_r}_{\lceil r \rceil_h, T} B_r \, \delta W_r \right\|_{L^2(\mathbb{P}; \mathbb{R}^k)} \\
&\leq \big\| f\big(X^{Y_0}_{0,T}\big) - f\left(Y_T\right) \big\|_{L^2(\mathbb{P}; \mathbb{R}^k)} + \left\| \int_0^T f'\left(X^{X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r}}_{\lceil r \rceil_h, T} \right) X^{1, X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r}}_{\lceil r \rceil_h, T} \mu(r, X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r}) - f'\big(X^{Y_r}_{\lceil r \rceil_h, T}\big) X^{1, Y_r}_{\lceil r \rceil_h, T} A_r \, dr \right\|_{L^2(\mathbb{P}; \mathbb{R}^k)} \\
&+ \left\| \tfrac{1}{2} \sum_{l,j=1}^{d} \int_0^T \Big(\sigma\big(r, X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r}\big) \big[\sigma\big(r, X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r}\big)\big]^* \Big)_{l,j} \right. \\
&\qquad \cdot \left(f''\left(X^{X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r}}_{\lceil r \rceil_h, T} \right) \left(X^{1, X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r}}_{\lceil r \rceil_h, T}, X^{1, X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r}}_{\lceil r \rceil_h, T} \right) + f'\left(X^{X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r}}_{\lceil r \rceil_h, T} \right) X^{2, X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r}}_{\lceil r \rceil_h, T} \right)(e^{(d)}_l, e^{(d)}_j) \\
&\qquad \left. - \big(B_r [B_r]^*\big)_{l,j} \left(f''\big(X^{Y_r}_{\lceil r \rceil_h, T}\big)\big(X^{1, Y_r}_{\lceil r \rceil_h, T}, X^{1, Y_r}_{\lceil r \rceil_h, T}\big) + f'\big(X^{Y_r}_{\lceil r \rceil_h, T}\big) X^{2, Y_r}_{\lceil r \rceil_h, T} \right)(e^{(d)}_l, e^{(d)}_j) \, dr \right\|_{L^2(\mathbb{P}; \mathbb{R}^k)} < K.
\end{aligned}
\tag{41}
$$

The fact that Y, X, X^1 are continuous random fields, continuity of f', and the fact that $\inf_{r \in [0,T]} \mathbb{P}(X^{Y_r}_{r,r} = Y_r) = 1$ yield that for all $r \in [0,T]$ it holds $\mathbb{P}$-a.s. that

$$
\begin{aligned}
&\lim_{{}^{T}\!/_{\mathbb{N}} \ni h \searrow 0} \left(f'\left(X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r, \lceil r \rceil_h, T} \right) X^{1, X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r}}_{\lceil r \rceil_h, T} \sigma\Big(r, X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r}\Big) - f'\Big(X^{Y_r}_{\lceil r \rceil_h, T}\Big) X^{1, Y_r}_{\lceil r \rceil_h, T} B_r \right) \\
&= f'\big(X^{Y_r}_{r,T}\big) X^{1, Y r}_{r,T} \Big(\sigma(r, Y_r) - B_r \Big).
\end{aligned}
\tag{42}
$$

This, Fatou's lemma, and the inequalities (29) and (24) yield that the sequence

$$
\left(f'\left(X^{Y_{\lfloor \cdot \rfloor_h}}_{\lfloor \cdot \rfloor_h, \cdot, \lceil \cdot \rceil_h, T} \right) X^{1, X^{Y_{\lfloor \cdot \rfloor_h}}_{\lfloor \cdot \rfloor_h, \cdot}}_{\lceil \cdot \rceil_h, T} \sigma\Big(\cdot, X^{Y_{\lfloor \cdot \rfloor_h}}_{\lfloor \cdot \rfloor_h, \cdot}\Big) - f'\Big(X^{Y_\cdot}_{\lceil \cdot \rceil_h, T}\Big) X^{1, Y_\cdot}_{\lceil \cdot \rceil_h, T} B_\cdot - f'\big(X^{Y_\cdot}_{\cdot, T}\big) X^{1, Y_\cdot}_{\cdot, T} \Big(\sigma(\cdot, Y_\cdot) - B_\cdot \Big) \right)_{h \in {}^{T}\!/_{\mathbb{N}}}
\tag{43}
$$

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 [24, 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 ${}^{T}\!/_{\mathbb{N}} \ni h \searrow 0$. This, the fact that the processes

$$
\left(f'\Big(X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r, \lceil r \rceil_h, T}\Big) X^{1, X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r}}_{\lceil r \rceil_h, T} \sigma\Big(r, X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r}\Big) - f'\big(X^{Y_r}_{\lceil r \rceil_h, T}\big) X^{1, Y_r}_{\lceil r \rceil_h, T} B_r \right)_{r \in [0,T]}, \quad h \in T/\mathbb{N},
\tag{44}
$$

are Skorohod-integrable, (41), and Lemma A.9 imply that the stochastic process

$$
\big(f'\big(X^{Y_r}_{r,T}\big) X^{1, Y_r}_{r,T} \left(\sigma(r, Y_r) - B_r \right) \big)_{r \in [0,T]}
\tag{45}
$$

is Skorohod-integrable and that for every $\mathbb{F}_T/\mathcal{B}([-1,1]^k)$-measurable function $Z \colon \Omega \to [-1,1]^k$ it holds that

$$
\begin{aligned}
\lim_{{}^{T}\!/_{\mathbb{N}} \ni h \searrow 0} \mathbb{E}\Bigg[\Bigg\langle Z, \int_0^T & f'\left(X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r, \lceil r \rceil_h, T} \right) X^{1, X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r}}_{\lceil r \rceil_h, T} \sigma\Big(r, X^{Y_{\lfloor r \rfloor_h}}_{\lfloor r \rfloor_h, r}\Big) - f'\Big(X^{Y_r}_{\lceil r \rceil_h, T}\Big) X^{1, Y_r}_{\lceil r \rceil_h, T} B_r \, \delta W_r \\
& - \int_0^T f'\big(X^{Y_r}_{r,T}\big) X^{1, Y_r}_{r,T} \sigma(r, Y_r) - f'\big(X^{Y_r}_{r,T}\big) X^{1, Y_r}_{r,T} B_r \, \delta W_r \Bigg\rangle_{\mathbb{R}^k} \Bigg] = 0.
\end{aligned}
\tag{46}
$$

Equation (32) and the convergences (38) and (46) imply that for every $\mathbb{F}_T/\mathcal{B}([-1,1]^k)$-measurable function $Z\colon \Omega \to [-1,1]^k$ it holds that

$$\begin{aligned}\mathbb{E}\Bigg[\Bigg\langle Z, \int_0^T f'\big(X_{r,T}^{Y_r}\big) X_{r,T}^{1,Y_r}\Big(\mu(r,Y_r) - A_r\Big)\,dr + \int_0^T f'\big(X_{r,T}^{Y_r}\big) X_{r,T}^{1,Y_r}\Big(\sigma(r,Y_r) - B_r\Big)\,\delta W_r \\ + \tfrac{1}{2}\sum_{l,j=1}^{d}\int_0^T \Big(\sigma(r,Y_r)[\sigma(r,Y_r)]^* - B_r[B_r]^*\Big)_{l,j}\Big(f''\big(X_{r,T}^{Y_r}\big)\big(X_{r,T}^{1,Y_r}, X_{r,T}^{1,Y_r}\big) + f'\big(X_{r,T}^{Y_r}\big) X_{r,T}^{2,Y_r}\Big)\big(e_l^{(d)}, e_j^{(d)}\big)\,dr \qquad (47) \\ - f\big(X_{0,T}^{Y_0}\big) + f(Y_T)\Bigg\rangle_{\mathbb{R}^k}\Bigg] = 0.\end{aligned}$$

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

4 Stochastic van-der-Pol oscillators with additive random forcing

In this subsection we illustrate the power of the Itô-Alekseev-Gröbner formula by applying it to a numerical approximation process to prove that the approximation processes (48) converge with L^2-rate $1/2$ to the exact solution of the stochastic van-der-Pol oscillator which is given by the SDE (49). Our proof is considerably shorter than the analysis in [15, 20] which is needed to prove the analogous statement with the approach of [15]. Here, we assume for simplicity of exposition that the diffusion coefficient is constant (but this is not the reason for the much simpler analysis compared to [15]). First, we introduce the setting for the stochastic van-der-Pol oscillator with additive noise, then we provide three auxiliary results (Lemma 4.2, Lemma 4.3, and Lemma 4.4), and finally we prove Lemma 4.5, the main result of this section, by an application of Theorem 1.1.

Setting 4.1. *Let $T \in (0,\infty)$, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $W\colon [0,T]\times\Omega \to \mathbb{R}$ be a standard Brownian motion with continuous sample paths, let $\mathcal{N} = \{A \in \mathcal{F}\colon \mathbb{P}(A) = 0\}$, let $\alpha, \beta, \gamma, \delta \in (0,\infty)$, let $\mu\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the function which satisfies for all $(x_1, x_2) \in \mathbb{R}^2$ that $\mu(x_1,x_2) = (x_2, (\gamma - \alpha x_1^2)x_2 - \delta x_1)$, let $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$, let $Y^N\colon [0,T]\times\Omega \to \mathbb{R}^2$, $N \in \mathbb{N}$, be stochastic processes such that for all $N \in \mathbb{N}$, $k \in \{0,\ldots,N-1\}$, $\varepsilon \in [0, \frac{T}{N}]$ it holds that $Y_0^N = \xi$ and that*

$$Y_{\frac{kT}{N}+\varepsilon}^N = Y_{\frac{kT}{N}}^N + \mu(Y_{\frac{kT}{N}}^N)\varepsilon \mathbb{1}_{\{\|\mu(Y_{\frac{kT}{N}}^N)\|_{\mathbb{R}^2}^2 < \frac{N}{T}\}} + (W_{\frac{kT}{N}+\varepsilon} - W_{\frac{kT}{N}})\binom{0}{\beta}, \qquad (48)$$

and let $X_{\cdot,\cdot}^{\cdot}\colon \Delta_T \times \mathbb{R}^2 \times \Omega \to \mathbb{R}^2$, $X_{\cdot,\cdot}^{1,\cdot}\colon \Delta_T \times \mathbb{R}^2 \times \Omega \to L(\mathbb{R}^2,\mathbb{R}^2)$, and $X_{\cdot,\cdot}^{2,\cdot}\colon \Delta_T \times \mathbb{R}^2 \times \Omega \to L^{(2)}(\mathbb{R}^2,\mathbb{R}^2)$ be continuous random fields such that for all $s \in [0,T]$, $\omega \in \Omega$ the function $(\mathbb{R}^2 \ni x \mapsto X_{s,T}^x(\omega) \in \mathbb{R}^2) \in C^2(\mathbb{R}^2,\mathbb{R}^2)$, such that for all $(s,x) \in [0,T]\times\mathbb{R}^2$, $i \in \{1,2\}$ the stochastic processes $[s,T]\times\Omega \ni (t,\omega) \mapsto X_{s,t}^x \in \mathbb{R}^2$, $[s,T]\times\Omega \ni (t,\omega) \mapsto X_{s,t}^{i,x} \in L^{(i)}(\mathbb{R}^2,\mathbb{R}^2)$ are $(\mathfrak{S}(\mathcal{N}\cup\mathfrak{S}(W_r - W_s\colon r \in [s,t])))_{t\in[s,T]}$-adapted and for all $t \in [s,T]$ it holds $\mathbb{P}$-a.s. that

$$X_{s,t}^x = x + \int_s^t \mu(X_{s,r}^x)\,dr + \int_s^t \binom{0}{\beta}\,dW_r, \qquad (49)$$

such that for all $r,s,t \in [0,T]$, $x \in \mathbb{R}^2$ with the property that $r \le s \le t$ it holds $\mathbb{P}$-a.s. that $X_{s,t}^{X_{r,s}^x} = X_{r,t}^x$ and such that for all $(s,t,x,\omega) \in \Delta_T \times \mathbb{R}^2 \times \Omega$ it holds that $X_{s,t}^{1,x}(\omega) = \frac{\partial}{\partial x} X_{s,t}^x(\omega)$ and $X_{s,t}^{2,x}(\omega) = \frac{\partial^2}{\partial x^2} X_{s,t}^x(\omega)$.

We note that we can not employ the classical Euler-Maruyama scheme since the Euler-Maruyama approximations diverge in the strong and weak sense for one-dimensional SDEs with superlinearly growing coefficient functions and possibly also in the case of the stochastic van-der-Pol oscillator; see [17, 19]. Instead we consider a tamed Euler scheme. A first tamed Euler scheme was introduced in [18] and a large class of tamed Euler schemes (including (48)) was investigated in [16].

Lemma 4.2. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $X\colon \Omega \to \mathbb{R}$ be a standard normally distributed $\mathcal{F}/\mathcal{B}(\mathbb{R})$-measurable function, and let $a, b, c \in \mathbb{R}$ satisfy that $2b^2c < 1$. Then it holds that

$$\mathbb{E}\Big[\exp\Big(c(a+bX)^2\Big)\Big] = \frac{1}{\sqrt{1-2b^2c}}\exp\Big(a^2\Big(c + \frac{2(bc)^2}{1-2b^2c}\Big)\Big). \tag{50}$$

Proof of Lemma 4.2. It holds for all $y \in \mathbb{R}$ that

$$c(a+by)^2 - \tfrac{y^2}{2} = ca^2 + 2abcy - y^2\big(\tfrac{1-2b^2c}{2}\big) = a^2\big(c + \tfrac{2(bc)^2}{1-2b^2c}\big) - \big(y - \tfrac{2abc}{1-2b^2c}\big)^2\big(\tfrac{1-2b^2c}{2}\big). \tag{51}$$

This, the definition of the standard normal distribution, equation (51), and the substitution rule imply that

$$\begin{aligned}
\mathbb{E}\Big[\exp\Big(c(a+bX)^2\Big)\Big] &= \int_{-\infty}^{\infty} \tfrac{1}{\sqrt{2\pi}}\exp\Big(c(a+by)^2 - \tfrac{y^2}{2}\Big)dy \\
&= \exp\Big(a^2\Big(c + \tfrac{2(bc)^2}{1-2b^2c}\Big)\Big)\int_{-\infty}^{\infty} \tfrac{1}{\sqrt{2\pi}}\exp\bigg(-\frac{\big(y - \frac{2abc}{1-2b^2c}\big)^2}{\big(\frac{2}{1-2b^2c}\big)}\bigg)dy \\
&= \tfrac{1}{\sqrt{1-2b^2c}}\exp\Big(a^2\Big(c + \tfrac{2(bc)^2}{1-2b^2c}\Big)\Big)\int_{-\infty}^{\infty} \tfrac{1}{\sqrt{2\pi}}\exp\Big(\tfrac{-y^2}{2}\Big)dy \\
&= \tfrac{1}{\sqrt{1-2b^2c}}\exp\Big(a^2\Big(c + \tfrac{2(bc)^2}{1-2b^2c}\Big)\Big).
\end{aligned} \tag{52}$$

The proof of Lemma 4.2 is thus completed. □

Lemma 4.3. Assume Setting 4.1, let $c \in (0, \exp(-T(1+3\beta^2+\delta+2\gamma))]$, and let $N \in \mathbb{N}$ satisfy that $N \geq \max\{6\beta^2 T, T\}$. Then it holds for all $r \in [0,T]$ that

$$\mathbb{E}\Big[\exp\Big(c\big\|Y_r^N\big\|_{\mathbb{R}^2}^2\Big)\Big] \leq \exp\Big((2\beta^2+1)T + \|\xi\|_{\mathbb{R}^2}^2\Big). \tag{53}$$

Proof of Lemma 4.3. Throughout the proof of Lemma 4.3 let $(\theta_k)_{k\in\{0,\ldots,N\}} \subseteq (0,\infty)$ satisfy that for all $k \in \{0,\ldots,N-1\}$ it holds that $\theta_N = c$ and $\theta_k = \theta_{k+1}(1+3\beta^2\frac{T}{N})\big(1+\frac{T}{N}(1+\delta+2\gamma)\big)$. Equation (48) and the fact that for all $k \in \{0,\ldots,N-1\}$, $\varepsilon \in [0, \frac{T}{N}]$ the sigma-algebras $\mathfrak{S}\big(Y^N_{\frac{kT}{N}}\big)$ and $\mathfrak{S}\big(W_{\frac{kT}{N}+\varepsilon} - W_{\frac{kT}{N}}\big)$ are independent and the random variables $W_{\frac{kT}{N}+\varepsilon} - W_{\frac{kT}{N}}$, $\sqrt{\varepsilon}\frac{W_T}{\sqrt{T}}$ are identically distributed yield that for all $k \in \{0,\ldots,N-1\}$, $\varepsilon \in [0, \frac{T}{N}]$ it holds that

$$\begin{aligned}
&\mathbb{E}\Big[\exp\Big(\theta_{k+1}\big\|Y^N_{\frac{kT}{N}+\varepsilon}\big\|_{\mathbb{R}^2}^2\Big)\Big] \\
&= \mathbb{E}\bigg[\mathbb{E}\Big[\exp\Big(\theta_{k+1}\big\|Y^N_{\frac{kT}{N}} + \varepsilon\mathbb{1}_{\{\|\mu(Y^N_{kT/N})\|_{\mathbb{R}^2}^2 < \frac{N}{T}\}}\mu(Y^N_{\frac{kT}{N}}) + (W_{\frac{kT}{N}+\varepsilon} - W_{\frac{kT}{N}})\binom{0}{\beta}\big\|_{\mathbb{R}^2}^2\Big)\Big|\mathfrak{S}\big(Y^N_{\frac{kT}{N}}\big)\Big]\bigg] \\
&= \mathbb{E}\bigg[\mathbb{E}\Big[\exp\big(\theta_{k+1}\big\|v + \big(W_{\frac{kT}{N}+\varepsilon} - W_{\frac{kT}{N}}\big)\binom{0}{\beta}\big\|_{\mathbb{R}^2}^2\big)\Big]\Big|_{v = Y^N_{\frac{kT}{N}} + \varepsilon\mathbb{1}_{\{\|\mu(Y^N_{kT/N})\|_{\mathbb{R}^2}^2 < \frac{N}{T}\}}\mu(Y^N_{\frac{kT}{N}})}\bigg] \\
&= \mathbb{E}\bigg[\exp(\theta_{k+1}v_1^2)\mathbb{E}\Big[\exp\Big(\theta_{k+1}\big(v_2 + \sqrt{\varepsilon}\beta\tfrac{W_T}{\sqrt{T}}\big)^2\Big)\Big]\Big|_{(v_1,v_2) = Y^N_{\frac{kT}{N}} + \varepsilon\mathbb{1}_{\{\|\mu(Y^N_{kT/N})\|_{\mathbb{R}^2}^2 < \frac{N}{T}\}}\mu(Y^N_{\frac{kT}{N}})}\bigg].
\end{aligned} \tag{54}$$

Induction, the fact that $\theta_N = c$, the fact that for all $x \in [0,\infty)$ it holds that $1+x \leq \exp(x)$, and the assumption $c \leq \exp(-T(1+3\beta^2+\delta+2\gamma))$ yield that for all $k \in \{0,\ldots,N\}$ it holds that

$$\begin{aligned}
\theta_k &= \theta_N\Big(\big(1+3\beta^2\tfrac{T}{N}\big)\big(1+\tfrac{T}{N}(1+\delta+2\gamma)\big)\Big)^{N-k} \leq c\exp\big((N-k)\tfrac{T}{N}(1+3\beta^2+\delta+2\gamma)\big) \\
&\leq \exp\big(-T(1+3\beta^2+\delta+2\gamma)\big)\exp\big(T(1+3\beta^2+\delta+2\gamma)\big) = 1.
\end{aligned} \tag{55}$$

Inequality (55), the fact that for all $\varepsilon \in \big[0, \frac{T}{N}\big]$, $k \in \{0, \dots, N-1\}$ it holds that $2(\sqrt{\varepsilon}\beta)^2\theta_{k+1} \leq \frac{2T\beta^2}{N} \leq \frac{1}{3}$, Lemma 4.2, the fact that $\frac{W_T}{\sqrt{T}}$ is standard normally distributed, and the fact that for all $x \in \big[0, \frac{1}{2}\big]$ it holds that $\frac{1}{1-x} \leq \exp(2x)$ imply that for all $k \in \{0, \dots, N-1\}$, $\varepsilon \in [0, \frac{T}{N}]$, $v = (v_1, v_2) \in \mathbb{R}^2$ it holds that

$$\begin{aligned}\exp(\theta_{k+1}v_1^2)\mathbb{E}\Big[\exp\Big(\theta_{k+1}\big(v_2 + \sqrt{\varepsilon}\beta\tfrac{W_T}{\sqrt{T}}\big)^2\Big)\Big] &= \tfrac{1}{\sqrt{1-2\varepsilon\beta^2\theta_{k+1}}}\exp\Big(\theta_{k+1}v_1^2 + v_2^2\Big(\theta_{k+1} + \tfrac{2\varepsilon\beta^2\theta_{k+1}^2}{1-2\varepsilon\beta^2\theta_{k+1}}\Big)\Big)\\ &\leq \sqrt{\exp(4\varepsilon\beta^2\theta_{k+1})}\exp\Big(\theta_{k+1}\big(1 + \tfrac{2\varepsilon\beta^2}{1-2\varepsilon\beta^2\theta_{k+1}}\big)\|v\|_{\mathbb{R}^2}^2\Big)\\ &\leq \exp\Big(2\tfrac{T}{N}\beta^2\Big)\exp\Big(\theta_{k+1}(1 + 3\beta^2\tfrac{T}{N})\|v\|_{\mathbb{R}^2}^2\Big).\end{aligned} \tag{56}$$

Equation (54) and inequality (56) imply that for all $k \in \{0, \dots, N-1\}$, $\varepsilon \in [0, \frac{T}{N}]$ it holds that

$$\mathbb{E}\Big[\exp\Big(\theta_{k+1}\big\|Y^N_{\frac{kT}{N}+\varepsilon}\big\|_{\mathbb{R}^2}^2\Big)\Big] \leq \exp\big(2\tfrac{T}{N}\beta^2\big)\,\mathbb{E}\left[\exp\left(\theta_{k+1}(1+3\beta^2\tfrac{T}{N})\Big\|Y^N_{\frac{kT}{N}} + \varepsilon\mathbb{1}_{\left\{\|\mu(Y^N_{\frac{kT}{N}})\|_{\mathbb{R}^2}^2 < \frac{N}{T}\right\}}\mu(Y^N_{\frac{kT}{N}})\Big\|_{\mathbb{R}^2}^2\right)\right]. \tag{57}$$

Young's inequality shows that for all $x = (x_1, x_2) \in \mathbb{R}^2$ it holds that

$$\langle x, \mu(x)\rangle_{\mathbb{R}^2} = x_1x_2 + x_2\big((\gamma - \alpha x_1^2)x_2 - \delta x_1\big) \leq x_1x_2 + \gamma x_2^2 - \delta x_1x_2 \leq \tfrac{1}{2}\big(1 + \delta + 2\gamma\big)\|x\|_{\mathbb{R}^2}^2. \tag{58}$$

This implies that for all $k \in \{0, \dots, N-1\}$, $\varepsilon \in (0, \frac{T}{N}]$ it holds that

$$\begin{aligned}&\Big\|Y^N_{\frac{kT}{N}} + \varepsilon\mathbb{1}_{\{\|\mu(Y^N_{\frac{kT}{N}})\|_{\mathbb{R}^2}^2 < \frac{N}{T}\}}\mu(Y^N_{\frac{kT}{N}})\Big\|_{\mathbb{R}^2}^2\\ &= \big\|Y^N_{\frac{kT}{N}}\big\|_{\mathbb{R}^2}^2 + 2\varepsilon\mathbb{1}_{\{\|\mu(Y^N_{\frac{kT}{N}})\|_{\mathbb{R}^2}^2 < \frac{N}{T}\}}\big\langle Y^N_{\frac{kT}{N}}, \mu\big(Y^N_{\frac{kT}{N}}\big)\big\rangle_{\mathbb{R}^2} + \varepsilon^2\mathbb{1}_{\{\|\mu(Y^N_{\frac{kT}{N}})\|_{\mathbb{R}^2}^2 < \frac{N}{T}\}}\big\|\mu(Y^N_{\frac{kT}{N}})\big\|_{\mathbb{R}^2}^2\\ &\leq \big\|Y^N_{\frac{kT}{N}}\big\|_{\mathbb{R}^2}^2 + 2\varepsilon\mathbb{1}_{\{\|\mu(Y^N_{\frac{kT}{N}})\|_{\mathbb{R}^2}^2 < \frac{N}{T}\}}\tfrac{(1+\delta+2\gamma)}{2}\big\|Y^N_{\frac{kT}{N}}\big\|_{\mathbb{R}^2}^2 + \varepsilon^2\mathbb{1}_{\{\|\mu(Y^N_{\frac{kT}{N}})\|_{\mathbb{R}^2}^2 < \frac{N}{T}\}}\tfrac{N}{T}\\ &\leq \big(1 + \tfrac{T}{N}\big(1 + \delta + 2\gamma\big)\big)\big\|Y^N_{\frac{kT}{N}}\big\|_{\mathbb{R}^2}^2 + \tfrac{T}{N}.\end{aligned} \tag{59}$$

Inequalities (57), (59), and (55) imply that for all $k \in \{0, \dots, N-1\}$, $\varepsilon \in [0, \frac{T}{N}]$ it holds that

$$\begin{aligned}\mathbb{E}\Big[\exp\Big(\theta_{k+1}\big\|Y^N_{\frac{kT}{N}+\varepsilon}\big\|_{\mathbb{R}^2}^2\Big)\Big] &\leq \exp\big(2\tfrac{T}{N}\beta^2\big)\mathbb{E}\Big[\exp\Big(\theta_{k+1}\Big(1 + 3\beta^2\tfrac{T}{N}\Big)\Big(\big(1 + \tfrac{T}{N}\big(1 + \delta + 2\gamma\big)\big)\big\|Y^N_{\frac{kT}{N}}\big\|_{\mathbb{R}^2}^2 + \tfrac{T}{N}\Big)\Big)\Big]\\ &= \exp\Big(\big(2\beta^2 + \theta_{k+1}(1 + 3\beta^2\tfrac{T}{N})\big)\tfrac{T}{N}\Big)\mathbb{E}\Big[\exp\Big(\theta_k\big\|Y^N_{\frac{kT}{N}}\big\|_{\mathbb{R}^2}^2\Big)\Big]\\ &\leq \exp\Big(\big(2\beta^2 + 1\big)\tfrac{T}{N}\Big)\mathbb{E}\Big[\exp\Big(\theta_k\big\|Y^N_{\frac{kT}{N}}\big\|_{\mathbb{R}^2}^2\Big)\Big].\end{aligned} \tag{60}$$

Next we prove by induction on $k \in \{0, \dots, N-1\}$ that for all $k \in \{0, \dots, N-1\}$, $r \in [\frac{kT}{N}, \frac{(k+1)T}{N}]$ it holds that

$$\mathbb{E}\Big[\exp\Big(\theta_{k+1}\big\|Y^N_r\big\|_{\mathbb{R}^2}^2\Big)\Big] \leq \exp\Big(\big(2\beta^2 + 1\big)\tfrac{(k+1)T}{N}\Big)\exp\big(\|\xi\|_{\mathbb{R}^2}^2\big). \tag{61}$$

Inequality (60) and $\theta_0 \leq 1$ imply the base case. For the induction step $\{0, \dots, N-2\} \ni k \mapsto k+1 \in \{0, \dots, N-1\}$ note that inequality (60) and the induction hypothesis imply for all $r \in [\frac{(k+1)T}{N}, \frac{(k+2)T}{N}]$ that

$$\begin{aligned}\mathbb{E}\Big[\exp\Big(\theta_{k+2}\big\|Y^N_r\big\|_{\mathbb{R}^2}^2\Big)\Big] &\leq \exp\big((2\beta^2 + 1)\tfrac{T}{N}\big)\mathbb{E}\Big[\exp\Big(\theta_{k+1}\Big\|Y^N_{\frac{(k+1)T}{N}}\Big\|_{\mathbb{R}^2}^2\Big)\Big]\\ &\leq \exp\Big(\big(2\beta^2 + 1\big)\tfrac{T}{N}\Big)\exp\Big(\big(2\beta^2 + 1\big)\tfrac{(k+1)T}{N}\Big)\exp\big(\|\xi\|_{\mathbb{R}^2}^2\big)\\ &= \exp\big(2(\beta^2 + 1)\tfrac{(k+2)T}{N}\big)\exp\big(\|\xi\|_{\mathbb{R}^2}^2\big).\end{aligned} \tag{62}$$

This finishes the induction step. Induction thus establishes inequality (61). Finally inequalities (55) and (61) yield that for all $k \in \{0, \ldots, N-1\}$, $r \in [\frac{kT}{N}, \frac{(k+1)T}{N}]$ it holds that

$$\begin{aligned}\mathbb{E}\Big[\exp\Big(c\big\|Y_r^N\big\|_{\mathbb{R}^2}^2\Big)\Big] &\leq \mathbb{E}\Big[\exp\Big(\theta_{k+1}\big\|Y_r^N\big\|_{\mathbb{R}^2}^2\Big)\Big] \leq \exp\big((2\beta^2+1)\tfrac{(k+1)T}{N}\big)\exp\big(\|\xi\|_{\mathbb{R}^2}^2\big) \\ &\leq \exp\Big((2\beta^2+1)T + \|\xi\|_{\mathbb{R}^2}^2\Big).\end{aligned} \tag{63}$$

The proof of Lemma 4.3 is thus completed. □

Lemma 4.4. Assume Setting 4.1, let $p \in [1,\infty)$, $r \in [0,T]$, $q \in (0,\infty)$, and let $Z\colon \Omega \to \mathbb{R}^2$ be an $\mathfrak{S}(\mathcal{N} \cup \mathfrak{S}(W_s\colon s \in [0,r]))/\mathcal{B}(\mathbb{R}^2)$-measurable function. Then it holds that

$$\begin{aligned}&\sup_{i\in\{1,2\}}\sup_{t\in[r,T]}\mathbb{E}\Big[\big\|X_{r,t}^{i,Z}\big\|_{L^{(i)}(\mathbb{R}^2,\mathbb{R}^2)}^p\Big] \\ &\leq \Bigg(\sup_{y\in(0,\infty)}\tfrac{(16\alpha^2Ty)^p}{\exp(q\exp(-(|\delta-1|+2\gamma+4\beta^2q)T)y)} + \exp\Big(6pT\big(1+\tfrac{\alpha^2 6p\exp((|\delta-1|+2\gamma+4\beta^2q)T)}{8\alpha q}+\gamma+\tfrac{\delta}{2}\big)\Big)\Bigg) \\ &\quad\cdot\mathbb{E}\Big[\exp(\tfrac{1}{4}+q\|Z\|_{\mathbb{R}^2}^2)\Big].\end{aligned} \tag{64}$$

Proof of Lemma 4.4. First, equation (49), the dominated convergence theorem together with continuity of the functions $[r,T]\times\mathbb{R}^2 \ni (s,x) \mapsto \frac{\partial}{\partial x}\big(\mu\big(X_{r,s}^x(\omega)\big)\big) \in L(\mathbb{R}^2,\mathbb{R}^2)$, $\omega\in\Omega$, and the chain rule imply that for all $t\in[r,T]$, $x,v\in\mathbb{R}^2$ it holds $\mathbb{P}$-a.s. that

$$\begin{aligned}X_{r,t}^{1,x}v &= \Big(\tfrac{\partial}{\partial x}X_{r,t}^x\Big)v = \tfrac{\partial}{\partial x}\Big(x + \int_r^t \mu(X_{r,s}^x)\,ds + \int_r^t \binom{0}{\beta}\,dW_s\Big)(v) = v + \int_r^t \tfrac{\partial}{\partial x}\Big(\mu(X_{r,s}^x)\Big)v\,ds \\ &= v + \int_r^t \mu'\big(X_{r,s}^x\big)X_{r,s}^{1,x}v\,ds.\end{aligned} \tag{65}$$

This, the fundamental theorem of calculus together with path continuity, and the chain rule imply that for all $t\in[r,T]$, $v\in\mathbb{R}^2$ it holds $\mathbb{P}$-a.s. that

$$\begin{aligned}\big\|X_{r,t}^{1,Z}v\big\|_{\mathbb{R}^2}^2 &= \|v\|_{\mathbb{R}^2}^2 + \int_r^t 2\big\langle X_{r,s}^{1,Z}v, \mu'(X_{r,s}^Z)X_{r,s}^{1,Z}v\big\rangle_{\mathbb{R}^2}\,ds \\ &= \|v\|_{\mathbb{R}^2}^2 + \int_r^t \tfrac{2\langle X_{r,s}^{1,Z}v,\mu'(X_{r,s}^Z)X_{r,s}^{1,Z}v\rangle_{\mathbb{R}^2}}{\|X_{r,s}^{1,Z}v\|_{\mathbb{R}^2}^2}\|X_{r,s}^{1,Z}v\|_{\mathbb{R}^2}^2\,ds.\end{aligned} \tag{66}$$

This and Gronwall's inequality together with path continuity imply that for all $t\in[r,T]$, $v\in\mathbb{R}^2$ it holds that

$$\big\|X_{r,t}^{1,Z}v\big\|_{\mathbb{R}^2}^2 \leq \|v\|_{\mathbb{R}^2}^2\exp\Big(\int_r^t \tfrac{2\langle X_{r,s}^{1,Z}v,\mu'(X_{r,s}^Z)X_{r,s}^{1,Z}v\rangle_{\mathbb{R}^2}}{\|X_{r,s}^{1,Z}v\|_{\mathbb{R}^2}^2}\,ds\Big). \tag{67}$$

For all $(u,v),(x,y)\in\mathbb{R}^2$, $\varepsilon\in(0,\infty)$ with the property that $x^2+y^2=1$ it holds that

$$\begin{aligned}&\big\langle\binom{x}{y},\mu'(u,v)\binom{x}{y}\big\rangle_{\mathbb{R}^2} \\ &= \left\langle\begin{pmatrix}x\\y\end{pmatrix},\begin{pmatrix}0 & 1\\ -2\alpha uv-\delta & \gamma-\alpha u^2\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}\right\rangle_{\mathbb{R}^2} = (1-2\alpha uv-\delta)xy + (\gamma-\alpha u^2)y^2 \\ &\leq \big(\tfrac{1}{2}+\alpha|uv|+\tfrac{\delta}{2}\big)+\gamma \leq \tfrac{1}{2}+\big(\tfrac{\alpha^2}{4\varepsilon}+\varepsilon(uv)^2\big)+\tfrac{\delta}{2}+\gamma = \big(\tfrac{1}{2}+\tfrac{\alpha^2}{4\varepsilon}+\gamma+\tfrac{\delta}{2}\big)+\varepsilon(uv)^2.\end{aligned} \tag{68}$$

Inequalities (67) and (68) imply that for all $t \in [r,T]$, $\varepsilon \in (0,\infty)$, $v \in \mathbb{R}^2$ it holds $\mathbb{P}$-a.s. that

$$\begin{aligned}
\big\|X_{r,t}^{1,Z} v\big\|_{\mathbb{R}^2}^p &\leq \|v\|_{\mathbb{R}^2}^p \exp\left(p \int_r^t \frac{\big\langle X_{r,s}^{1,Z} v, \mu'(X_{r,s}^Z) X_{r,s}^{1,Z} v \big\rangle_{\mathbb{R}^2}}{\|X_{r,s}^{1,Z} v\|_{\mathbb{R}^2}^2} \, ds \right) \\
&\leq \|v\|_{\mathbb{R}^2}^p \exp\Big(p \int_r^t \big(\tfrac{1}{2} + \tfrac{\alpha^2}{4\varepsilon} + \gamma + \tfrac{\delta}{2} \big) + \varepsilon \left(\Big\langle X_{r,s}^Z, e_1^{(2)} \Big\rangle_{\mathbb{R}^2} \Big\langle X_{r,s}^Z, e_2^{(2)} \Big\rangle_{\mathbb{R}^2} \right)^2 ds \Big) \\
&\leq \|v\|_{\mathbb{R}^2}^p \exp\Big(pt\big(\tfrac{1}{2} + \tfrac{\alpha^2}{4\varepsilon} + \gamma + \tfrac{\delta}{2} \big) \Big) \exp\Big(p\varepsilon \int_r^t \left(\Big\langle X_{r,s}^Z, e_1^{(2)} \Big\rangle_{\mathbb{R}^2} \Big\langle X_{r,s}^Z, e_2^{(2)} \Big\rangle_{\mathbb{R}^2} \right)^2 ds \Big).
\end{aligned} \tag{69}$$

Observe that

$$\min_{r \in (0,\infty)} \max\Big\{ \Big[\tfrac{|\delta - 1|}{r} \Big], \big[r|\delta - 1| + 2\gamma + 4\beta^2 q \big] \Big\} \leq |\delta - 1| + 2\gamma + 4\beta^2 q. \tag{70}$$

Inequality (69) (applied with $\varepsilon = \frac{2\alpha q}{p \exp((|\delta - 1| + 2\gamma + 4\beta^2 q)t)}$ in the notation of inequality (69)), inequality (70), and Cox et al. [7, equation (4.4)] imply for all $t \in [r,T]$ that

$$\mathbb{E}\Big[\big\| X_{r,t}^{1,Z} \big\|_{L(\mathbb{R}^2,\mathbb{R}^2)}^p \Big] \leq \exp\Big(pt\big(\tfrac{1}{2} + \tfrac{\alpha^2 p \exp((|\delta - 1| + 2\gamma + 4\beta^2 q)t)}{8\alpha q} + \gamma + \tfrac{\delta}{2} \big) \Big) \mathbb{E}\Big[\exp\big(\tfrac{1}{4} + q \|Z\|_{\mathbb{R}^2}^2 \big) \Big]. \tag{71}$$

Next equation (65), the dominated convergence theorem together with continuity of the functions $[r,T] \times \mathbb{R}^2 \ni (s,x) \mapsto \frac{\partial^2}{\partial x^2} \big(\mu(X_{r,s}^x(\omega)) \big) \in L^{(2)}(\mathbb{R}^2, \mathbb{R}^2)$, $\omega \in \Omega$, and the chain rule imply that for all $t \in [r,T]$, $x, v, w \in \mathbb{R}^2$ it holds $\mathbb{P}$-a.s. that

$$\begin{aligned}
X_{r,t}^{2,x}(v,w) &= \Big(\tfrac{\partial}{\partial x} X_{r,t}^{1,x} v \Big) w = \tfrac{\partial}{\partial x} \Big(v + \int_r^t \mu'(X_{r,s}^x) X_{r,s}^{1,x} v \, ds \Big)(w) \\
&= \int_r^t \tfrac{\partial}{\partial x} \Big(\mu'(X_{r,s}^x) X_{r,s}^{1,x} v \Big)(w) \, ds \\
&= \int_r^t \mu''\big(X_{r,s}^x\big)\big(X_{r,s}^{1,x} v, X_{r,s}^{1,x} w\big) + \mu'\big(X_{r,s}^x\big) X_{r,s}^{2,x}(v,w) \, ds.
\end{aligned} \tag{72}$$

Equation (72), the fundamental theorem of calculus together with path continuity, the chain rule, the Cauchy-Schwarz inequality, Young's inequality, and inequality (68) imply that for all $t \in [r,T]$, $v, w \in \mathbb{R}^2$, $\varepsilon \in (0,\infty)$ it holds $\mathbb{P}$-a.s. that

$$\begin{aligned}
&\big\| X_{r,t}^{2,Z}(v,w) \big\|_{\mathbb{R}^2}^2 \\
&= 2 \int_r^t \big\langle X_{r,s}^{2,Z}(v,w), \mu''\big(X_{r,s}^Z\big)\big(X_{r,s}^{1,Z} v, X_{r,s}^{1,Z} w\big) + \mu'\big(X_{r,s}^Z\big) X_{r,s}^{2,Z}(v,w) \big\rangle_{\mathbb{R}^2} \, ds \\
&\leq \int_r^t 2 \big\| X_{r,s}^{2,Z}(v,w) \big\|_{\mathbb{R}^2} \big\| \mu''(X_{r,s}^Z)(X_{r,s}^{1,Z} v, X_{r,s}^{1,Z} w) \big\|_{\mathbb{R}^2} \, ds + 2 \int_r^t \big\langle X_{r,s}^{2,Z}(v,w), \mu'\big(X_{r,s}^Z\big) X_{r,s}^{2,Z}(v,w) \big\rangle_{\mathbb{R}^2} \, ds \\
&\leq \int_r^t \big\| X_{r,s}^{2,Z}(v,w) \big\|_{\mathbb{R}^2}^2 + \big\| \mu''\big(X_{r,s}^Z\big)\big(X_{r,s}^{1,Z} v, X_{r,s}^{1,Z} w\big) \big\|_{\mathbb{R}^2}^2 \, ds + 2 \int_r^t \big\langle X_{r,s}^{2,Z}(v,w), \mu'\big(X_{r,s}^Z\big) X_{r,s}^{2,Z}(v,w) \big\rangle_{\mathbb{R}^2} \, ds \\
&\leq \int_r^t \big\| \mu''\big(X_{r,s}^Z\big)\big(X_{r,s}^{1,Z} v, X_{r,s}^{1,Z} w\big) \big\|_{\mathbb{R}^2}^2 \, ds \\
&\quad + 2 \int_r^t \Big(\big(\tfrac{1}{2} + \tfrac{1}{2} + \tfrac{\alpha^2}{4\varepsilon} + \gamma + \tfrac{\delta}{2} \big) + \varepsilon \Big(\big\langle X_{r,s}^Z, e_1^{(2)} \big\rangle_{\mathbb{R}^2} \big\langle X_{r,s}^Z, e_2^{(2)} \big\rangle_{\mathbb{R}^2} \Big)^2 \Big) \big\| X_{r,s}^{2,Z}(v,w) \big\|_{\mathbb{R}^2}^2 \, ds.
\end{aligned} \tag{73}$$

This and Gronwall's inequality together with path continuity imply that for all $t \in [r,T]$, $v, w \in \mathbb{R}^2$, $\varepsilon \in (0,\infty)$ it holds that

$$\begin{aligned}
\big\| X_{r,t}^{2,Z}(v,w) \big\|_{\mathbb{R}^2}^2 &\leq \int_r^t \big\| \mu''\big(X_{r,s}^Z\big)\big(X_{r,s}^{1,Z} v, X_{r,s}^{1,Z} w\big) \big\|_{\mathbb{R}^2}^2 \, ds \\
&\quad \cdot \exp\Big(2t\big(1 + \tfrac{\alpha^2}{4\varepsilon} + \gamma + \tfrac{\delta}{2} \big) \Big) \exp\Big(2\varepsilon \int_r^t \Big(\big\langle X_{r,s}^Z, e_1^{(2)} \big\rangle_{\mathbb{R}^2} \big\langle X_{r,s}^Z, e_2^{(2)} \big\rangle_{\mathbb{R}^2} \Big)^2 \, ds \Big).
\end{aligned} \tag{74}$$

For all $(u,v),(x,y),(w,z)\in\mathbb{R}^2$ with the property that $x^2+y^2=1=w^2+z^2$ it holds that

$$\left\|\mu''(u,v)\left(\left(\begin{smallmatrix}x\\y\end{smallmatrix}\right),\left(\begin{smallmatrix}w\\z\end{smallmatrix}\right)\right)\right\|_{\mathbb{R}^2}=\left\|\begin{pmatrix}0&0\\-2\alpha vx-2\alpha uy&-2\alpha ux\end{pmatrix}\begin{pmatrix}w\\z\end{pmatrix}\right\|_{\mathbb{R}^2}=2\alpha|(vx+uy)w+uxz|\leq 4\alpha\left\|\left(\begin{smallmatrix}u\\v\end{smallmatrix}\right)\right\|_{\mathbb{R}^2}. \tag{75}$$

This, inequality (74), and inequality (69) yield that for all $t\in[r,T]$, $\varepsilon\in(0,\infty)$ it holds that

$$\begin{aligned}
&\left\|X^{2,Z}_{r,t}\right\|^p_{L^{(2)}(\mathbb{R}^2,\mathbb{R}^2)}\\
&\leq\left(\int_r^t 16\alpha^2\left\|X^Z_{r,s}\right\|^2_{\mathbb{R}^2}\left\|X^{1,Z}_{r,s}\right\|^4_{L(\mathbb{R}^2,\mathbb{R}^2)}\,ds\right)^{\frac{p}{2}}\exp\Big(pt\big(1+\tfrac{\alpha^2}{4\varepsilon}+\gamma+\tfrac{\delta}{2}\big)\Big)\exp\Big(p\varepsilon\int_r^t\Big(\Big\langle X^Z_{r,s},e^{(2)}_1\Big\rangle_{\mathbb{R}^2}\Big\langle X^Z_{r,s},e^{(2)}_2\Big\rangle_{\mathbb{R}^2}\Big)^2\,ds\Big)\\
&\leq\left(\int_r^T 16\alpha^2\left\|X^Z_{r,s}\right\|^2_{\mathbb{R}^2}\,ds\right)^{\frac{p}{2}}\exp\Big(3pt\big(1+\tfrac{\alpha^2}{4\varepsilon}+\gamma+\tfrac{\delta}{2}\big)\Big)\exp\Big(3p\varepsilon\int_r^t\Big(\Big\langle X^Z_{r,s},e^{(2)}_1\Big\rangle_{\mathbb{R}^2}\Big\langle X^Z_{r,s},e^{(2)}_2\Big\rangle_{\mathbb{R}^2}\Big)^2\,ds\Big).
\end{aligned}\tag{76}$$

The triangle inequality yields that

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\int_r^T 16\alpha^2\left\|X^Z_{r,s}\right\|^2_{\mathbb{R}^2}\,ds\Big)^p\Big]\leq\left(\int_r^T 16\alpha^2\left(\mathbb{E}\Big[\left\|X^Z_{r,s}\right\|^{2p}_{\mathbb{R}^2}\Big]\right)^{\frac{1}{p}}ds\right)^p\leq(16\alpha^2T)^p\sup_{u\in[r,T]}\mathbb{E}\Big[\left\|X^Z_{r,u}\right\|^{2p}_{\mathbb{R}^2}\Big]\\
&\leq(16\alpha^2T)^p\sup_{y\in(0,\infty)}\tfrac{y^p}{\exp(q\exp(-(|\delta-1|+2\gamma+4\beta^2q)T)y)}\sup_{u\in[r,T]}\mathbb{E}\Big[\exp\Big(\tfrac{q}{\exp((|\delta-1|+2\gamma+4\beta^2q)u)}\left\|X^Z_{r,u}\right\|^2_{\mathbb{R}^2}\Big)\Big].
\end{aligned}\tag{77}$$

Inequality (76) (applied with $\varepsilon=\frac{2\alpha q}{6p\exp((|\delta-1|+2\gamma+4\beta^2q)t)}$ for $t\in[r,T]$ in the notation of inequality (76)), the fact that for all $a,b\in\mathbb{R}$ it holds that $ab\leq a^2+b^2$, inequality (77), and Cox et al. [7, equation (4.4)] imply for all $t\in[r,T]$ that

$$\begin{aligned}
&\mathbb{E}\Big[\left\|X^{2,Z}_{r,t}\right\|^p_{L^{(2)}(\mathbb{R}^2,\mathbb{R}^2)}\Big]\\
&\leq\mathbb{E}\Big[\Big(\int_r^T 16\alpha^2\left\|X^Z_{r,s}\right\|^2_{\mathbb{R}^2}\,ds\Big)^p\Big]+\exp\Big(2\cdot 3pt\big(1+\tfrac{\alpha^2 6p\exp((|\delta-1|+2\gamma+4\beta^2q)t)}{8\alpha q}+\gamma+\tfrac{\delta}{2}\big)\Big)\\
&\quad\cdot\mathbb{E}\Big[\exp\Big(2\cdot 3p\tfrac{2\alpha q}{6p\exp((|\delta-1|+2\gamma+4\beta^2q)t)}\int_r^t\Big(\Big\langle X^Z_{r,s},e^{(2)}_1\Big\rangle_{\mathbb{R}^2}\Big\langle X^Z_{r,s},e^{(2)}_2\Big\rangle_{\mathbb{R}^2}\Big)^2\,ds\Big)\Big)\Big]\\
&\leq\Big(\sup_{y\in(0,\infty)}\tfrac{(16\alpha^2Ty)^p}{\exp(q\exp(-(|\delta-1|+2\gamma+4\beta^2q)T)y)}+\exp\Big(6pt\big(1+\tfrac{\alpha^2 6p\exp((|\delta-1|+2\gamma+4\beta^2q)t)}{8\alpha q}+\gamma+\tfrac{\delta}{2}\big)\Big)\Big)\\
&\quad\cdot\sup_{u\in[r,T]}\mathbb{E}\Big[\exp\Big(\tfrac{q}{\exp((|\delta-1|+2\gamma+4\beta^2q)u)}\left\|X^Z_{r,u}\right\|^2_{\mathbb{R}^2}+\int_r^u\tfrac{2\alpha q}{\exp((|\delta-1|+2\gamma+4\beta^2q)s)}\Big(\Big\langle X^Z_{r,s},e^{(2)}_1\Big\rangle_{\mathbb{R}^2}\Big\langle X^Z_{r,s},e^{(2)}_2\Big\rangle_{\mathbb{R}^2}\Big)^2\,ds\Big)\Big]\\
&\leq\Big(\sup_{y\in(0,\infty)}\tfrac{(16\alpha^2Ty)^p}{\exp(q\exp(-(|\delta-1|+2\gamma+4\beta^2q)T)y)}+\exp\Big(6pT\big(1+\tfrac{\alpha^2 6p\exp((|\delta-1|+2\gamma+4\beta^2q)T)}{8\alpha q}+\gamma+\tfrac{\delta}{2}\big)\Big)\Big)\\
&\quad\cdot\mathbb{E}\Big[\exp(\tfrac{1}{4}+q\|Z\|^2_{\mathbb{R}^2})\Big].
\end{aligned}\tag{78}$$

Combining (71) and (78) proves (64). The proof of Lemma 4.4 is thus completed. □

Lemma 4.5. Assume Setting 4.1. Then there exists a constant $\kappa\in(0,\infty)$ such that for all $N\in\mathbb{N}$ it holds that

$$\left(\mathbb{E}\Big[\left\|X^\xi_{0,T}-Y^N_T\right\|^2_{\mathbb{R}^2}\Big]\right)^{\frac{1}{2}}\leq\tfrac{\kappa}{\sqrt{N}}. \tag{79}$$

Proof of Lemma 4.5. For the rest of the proof let $p \in [5 + \exp(T(4 + \delta + 2\gamma)), \infty)$ be a real number with the property that for all $x, y \in \mathbb{R}^2$ it holds that

$$\|\mu(x) - \mu(y)\|_{\mathbb{R}^2} \leq p\|x - y\|_{\mathbb{R}^2}(1 + \|x\|_{\mathbb{R}^2}^p + \|y\|_{\mathbb{R}^2}^p) \tag{80}$$

and

$$\|\mu(x)\|_{\mathbb{R}^2} \leq p(1 + \|x\|_{\mathbb{R}^2}^p). \tag{81}$$

Such a real number exists because μ is a polynomial. Lemma 4.4 (applied with $p = 5$, $r = s$, $q = q\exp(-(|\delta - 1| + 2\gamma + 4\beta^2 q)s)$, $Z = X_{r,s}^{Y_r^N}$ for $r, s \in [0, T]$, $N \in \mathbb{N}$ with $r \leq s$ in the notation of Lemma 4.4), Cox et al. [7, equation (4.4)], and Lemma 4.3 (applied with $c = \exp(-T(1 + 3\beta^2 + \delta + 2\gamma))$ in the notation of Lemma 4.3) imply that there exists a constant $C \in [1, \infty)$ such that for all $N \in \mathbb{N}$, $q \in \{\exp(-T(1 + 3\beta^2 + \delta + 2\gamma))\}$, $r, s, t \in [0, T]$ with the property that $r \leq s \leq t$ and that $N \geq \max\{6\beta^2 T, T\}$ it holds that

$$\begin{aligned}
\max\left\{\mathbb{E}\Big[\Big\|X_{s,t}^{X_{r,s}^{Y_r^N}}\Big\|_{\mathbb{R}^2}^{20p}\Big], \sup_{i\in\{1,2\}} \mathbb{E}\Big[\Big\|X_{s,t}^{i,X_{r,s}^{Y_r^N}}\Big\|_{L^{(i)}(\mathbb{R}^2,\mathbb{R}^2)}^{\frac{4\cdot 20}{20-4}}\Big]\right\} &\leq C\mathbb{E}\Big[\exp\Big(\tfrac{q}{\exp((|\delta-1|+2\gamma+4\beta^2 q)s)}\Big\|X_{r,s}^{Y_r^N}\Big\|_{\mathbb{R}^2}^2\Big)\Big] \\
&\leq Ce^{1/4}\mathbb{E}\Big[\exp\Big(q\big\|Y_r^N\big\|_{\mathbb{R}^2}^2\Big)\Big] \\
&\leq Ce^{1/4}\exp\Big((2\beta^2+1)T + \|\xi\|_{\mathbb{R}^2}^2\Big) < \infty.
\end{aligned} \tag{82}$$

This together with inequality (81) implies that the assumptions of Theorem 1.1 are satisfied. Then the perturbation formula in Theorem 1.1 (applied with $d = 2$, $m = 1$, $p = 20$, $\mu(r, x) = \mu(x)$, $\sigma(r, x) = \binom{0}{\beta}$, $A_r = \mu\big(Y_{\lfloor r\rfloor_{T/N}}^N\big)\mathbb{1}_{\{\|\mu(Y_{\lfloor r\rfloor_{T/N}}^N)\|_{\mathbb{R}^2}^2 < \frac{N}{T}\}}$, $B_r = \binom{0}{\beta}$ for all $r \in [0, T]$, $x \in \mathbb{R}^2$ in the notation of Theorem 1.1), Jensen's inequality, and Hölder's inequality imply for all $N \in \mathbb{N}$ with the property $N \geq \max\{6\beta^2 T, T\}$ that

$$\begin{aligned}
\mathbb{E}\Big[\big\|X_{0,T}^{\xi} - Y_T^N\big\|_{\mathbb{R}^2}^2\Big] &= \mathbb{E}\left[\left\|\int_0^T X_{r,T}^{1,Y_r^N}\Big(\mu(Y_r^N) - \mu\big(Y_{\lfloor r\rfloor_{T/N}}^N\big)\mathbb{1}_{\{\|\mu(Y_{\lfloor r\rfloor_{T/N}}^N)\|_{\mathbb{R}^2}^2 < \frac{N}{T}\}}\Big)\,dr\right\|_{\mathbb{R}^2}^2\right] \\
&\leq T\Big(\int_0^T \mathbb{E}\Big[\Big\|X_{r,T}^{1,Y_r^N}\Big\|_{L(\mathbb{R}^2,\mathbb{R}^2)}^4\Big]\,dr\Big)^{\frac{1}{2}}\Big(\int_0^T \mathbb{E}\Big[\big\|\mu(Y_r^N) - \mu\big(Y_{\lfloor r\rfloor_{T/N}}^N\big)\mathbb{1}_{\{\|\mu(Y_{\lfloor r\rfloor_{T/N}}^N)\|_{\mathbb{R}^2}^2 < \frac{N}{T}\}}\big\|_{\mathbb{R}^2}^4\Big]\,dr\Big)^{\frac{1}{2}}.
\end{aligned} \tag{83}$$

Moreover, inequality (80), Hölder's inequality, equation (48), the scaling property of Brownian motion, and the fact that $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} x^8 e^{-\frac{x^2}{2}}\,dx = 105$ yields that for all $N \in \mathbb{N}$, $k \in \{0, \ldots, N-1\}$, $r \in [\frac{kT}{N}, \frac{(k+1)T}{N}]$ it holds that

$$\begin{aligned}
&\Big\|\big(\mu(Y_r^N) - \mu(Y_{\frac{kT}{N}}^N)\big)\mathbb{1}_{\{\|\mu(Y_{\frac{kT}{N}}^N)\|_{\mathbb{R}^2}^2 < \frac{N}{T}\}}\Big\|_{L^4(\mathbb{P};\mathbb{R}^2)} \\
&\leq p\Big\|\|Y_r^N - Y_{\frac{kT}{N}}^N\|_{\mathbb{R}^2}\big(1 + \|Y_r^N\|_{\mathbb{R}^2}^p + \|Y_{\frac{kT}{N}}^N\|_{\mathbb{R}^2}^p\big)\mathbb{1}_{\{\|\mu(Y_{\frac{kT}{N}}^N)\|_{\mathbb{R}^2}^2 < \frac{N}{T}\}}\Big\|_{L^4(\mathbb{P};\mathbb{R})} \\
&\leq p\Big\|\big(Y_r^N - Y_{\frac{kT}{N}}^N\big)\mathbb{1}_{\{\|\mu(Y_{\frac{kT}{N}}^N)\|_{\mathbb{R}^2}^2 < \frac{N}{T}\}}\Big\|_{L^8(\mathbb{P};\mathbb{R}^2)}\Big\|\big(1 + \|Y_r^N\|_{\mathbb{R}^2}^p + \|Y_{\frac{kT}{N}}^N\|_{\mathbb{R}^2}^p\big)\Big\|_{L^8(\mathbb{P};\mathbb{R})} \\
&\leq p\bigg(\tfrac{T}{N}\Big\|\mu\big(Y_{\frac{kT}{N}}^N\big)\mathbb{1}_{\{\|\mu(Y_{\frac{kT}{N}}^N)\|_{\mathbb{R}^2}^2 < \frac{N}{T}\}}\Big\|_{L^8(\mathbb{P};\mathbb{R}^2)} + \big\|\big(W_r - W_{\frac{kT}{N}}\big)\binom{0}{\beta}\big\|_{L^8(\mathbb{P};\mathbb{R}^2)}\bigg)\Big(1 + 2\sup_{s\in[0,T]}\big\|Y_s^N\big\|_{L^{8p}(\mathbb{P};\mathbb{R}^2)}^p\Big) \\
&\leq p\bigg(\tfrac{T}{N}\sqrt{\tfrac{N}{T}} + (105)^{\frac{1}{8}}\beta\sqrt{\tfrac{T}{N}}\bigg)\Big(1 + 2\sup_{s\in[0,T]}\big\|Y_s^N\big\|_{L^{20p}(\mathbb{P};\mathbb{R}^2)}^p\Big) < \infty.
\end{aligned} \tag{84}$$

This, inequality (83), Hölder's inequality, and inequality (82) yield that for all $N \in \mathbb{N}$ with $N \geq \max\{6\beta^2 T, T\}$ it holds that

$$\begin{aligned}
&\Big(\mathbb{E}\Big[\big\|X^{\xi}_{0,T} - Y^N_T\big\|^2_{\mathbb{R}^2}\Big]\Big)^{\frac{1}{2}} \\
&\leq \sqrt{T}\Big(\int_0^T \mathbb{E}\Big[\Big\|X^{1,Y^N_r}_{r,T}\Big\|^4_{L(\mathbb{R}^2,\mathbb{R}^2)}\Big]\,dr\Big)^{\frac{1}{4}} \cdot T^{\frac{1}{4}}\sqrt{\tfrac{T}{N}}p(1+2\beta)\Big(1 + 2\sup_{s\in[0,T]}\Big(\mathbb{E}\Big[\big\|Y^N_s\big\|^{20p}_{\mathbb{R}^2}\Big]\Big)^{\frac{1}{20}}\Big) \qquad (85)\\
&\leq \sqrt{\tfrac{T}{N}}T\left(Ce^{1/4}\exp\Big((2\beta^2+1)T + \|\xi\|^2_{\mathbb{R}^2}\Big)\right)^{\frac{1}{5}+\frac{1}{20}} p(1+2\beta)3.
\end{aligned}$$

This together with $\max_{N\in\mathbb{N}\cap[0,\max\{6\beta^2T,T\}+1]} \mathbb{E}\big[\big\|X^{\xi}_{0,T} - Y^N_T\big\|^2_{\mathbb{R}^2}\big] < \infty$ implies (79). The proof of Lemma 4.5 is thus completed. □

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 $W\colon [0,3]\times\Omega \to \mathbb{R}$ the Itô integrals $\int_0^1 \sin(W_s(W_2 - W_1))\,dW_s$ and $\int_1^2 \sin(W_s(W_3 - W_2))\,dW_s$ are well-defined (however with respect to different filtrations) but their sum cannot be written as Itô integral $\int_0^2 \sin(W_s(W_{\lceil s\rceil_1+1} - W_{\lceil s\rceil_1}))\,dW_s$ (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 $W\colon [S,T]\times\Omega \to \mathbb{R}^m$ 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 $\mathfrak{S}(W_t - W_S\colon t\in[S,T])$, let $\mathcal{N} = \{A \in \mathcal{F}\colon \mathbb{P}(A) = 0\}$, let $\mathbb{F}_T \subseteq \mathcal{F}$ be the sigma-algebra which satisfies that $\mathbb{F}_T = \mathfrak{S}(\mathbb{F}_S \cup \mathfrak{S}(W_t - W_S\colon t\in[S,T]) \cup \mathcal{N})$, let $\mathcal{S}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d) \subseteq L^2(\mathbb{P}|_{\mathbb{F}_T}; \mathbb{R}^d)$ be the subset with the property that*

$$\mathcal{S}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d) = \left\{ \begin{array}{c} F \in L^2(\mathbb{P}|_{\mathbb{F}_T}; \mathbb{R}^d)\colon \exists n\in\mathbb{N}, \exists \phi_1,\dots,\phi_n \in \mathcal{L}^2(\lambda_{[S,T]}; \mathbb{R}^m), \\ \exists f \in C_b^{\infty,\mathfrak{S}(\mathbb{F}_S\cup\mathcal{N})}(\mathbb{R}^n\times\Omega, \mathbb{R}), \exists h\in\mathbb{R}^d \text{ such that it holds } \mathbb{P}\text{-a.s. that} \\ F = f\big(\int_S^T \phi_1(r)\,dW_r, \dots, \int_S^T \phi_n(r)\,dW_r\big)h \end{array} \right\}, \qquad (86)$$

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 $\mathbb{F}_{[S,s]\cup[t,T]} = \mathfrak{S}(\mathbb{F}_S \cup \mathfrak{S}(W_r - W_S\colon r\in[S,s]) \cup \mathfrak{S}(W_r - W_t\colon r\in[t,T]) \cup \mathcal{N})$.

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

$$\mathcal{D}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d)\colon \mathbb{D}^{(1,2)}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d) \to L^2(\mathbb{P}|_{\mathbb{F}_T}; L^2(\lambda_{[S,T]}; \mathbb{R}^{d\times m})) \qquad (87)$$

is the closed linear operator with the property that for all $F \in \mathcal{S}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d)$ with the property that $\exists n\in\mathbb{N}$, $\exists \phi_1,\dots,\phi_n \in \mathcal{L}^2(\lambda_{[S,T]}; \mathbb{R}^m)$, $\exists f \in C_b^{\infty,\mathfrak{S}(\mathbb{F}_S\cup\mathcal{N})}(\mathbb{R}^n\times\Omega, \mathbb{R})$, $\exists h\in\mathbb{R}^d$ such that it holds $\mathbb{P}$-a.s. that $F = f\big(\int_S^T \phi_1(r)\,dW_r, \dots, \int_S^T \phi_n(r)\,dW_r\big)h$ it holds $\lambda_{[S,T]}\otimes\mathbb{P}$-a.e. that

$$\mathcal{D}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d)F = \sum_{i=1}^n \frac{\partial f}{\partial x_i}\Big(\int_S^T \phi_1(s)\,dW_s, \dots, \int_S^T \phi_n(s)\,dW_s\Big)\phi_i h \qquad (88)$$

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

$$\|\cdot\|_{\mathbb{D}^{(1,2)}(\mathbb{P},\mathbb{F}_S,W;\mathbb{R}^d)} = \Big(\mathbb{E}\Big[\|\cdot\|^2_{\mathbb{R}^d} + \|\mathcal{D}(\mathbb{P},\mathbb{F}_S,W;\mathbb{R}^d)\cdot\|^2_{L^2(\lambda_{[S,T]};\mathbb{R}^d)}\Big]\Big)^{\frac{1}{2}}. \tag{89}$$

We write $\mathcal{D} = \mathcal{D}(\mathbb{P},\mathfrak{S}(\mathcal{N}),W;\mathbb{R}^d)$ *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 (88) 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 [27] and therefore omitted.

Lemma A.3. Assume Setting A.1. Then the operator

$$\mathcal{D}(\mathbb{P},\mathbb{F}_S,W;\mathbb{R}^d)\colon \mathbb{D}^{(1,2)}(\mathbb{P},\mathbb{F}_S,W;\mathbb{R}^d) \to L^2(\mathbb{P}|_{\mathbb{F}_T};L^2(\lambda_{[S,T]};\mathbb{R}^{d\times m})) \tag{90}$$

is well-defined.

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

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

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*

$$\delta(\mathbb{P},\mathbb{F}_S,W;\mathbb{R}^d)\colon \operatorname{Dom}_\delta(\mathbb{P},\mathbb{F}_S,W;\mathbb{R}^d) \to L^2(\mathbb{P}|_{\mathbb{F}_T};\mathbb{R}^d) \tag{91}$$

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

$$\mathbb{E}[\langle \mathcal{D}(\mathbb{P},\mathbb{F}_S,W;\mathbb{R}^d)F, X\rangle_{L^2(\lambda_{[S,T]};\mathbb{R}^{d\times m})}] \le c\|F\|_{L^2(\mathbb{P};\mathbb{R}^d)} \tag{92}$$

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

$$\mathbb{E}\Big[\Big\langle F, \delta(\mathbb{P},\mathbb{F}_S,W;\mathbb{R}^d)(X)\Big\rangle_{\mathbb{R}^d}\Big] = \mathbb{E}\Big[\Big\langle \mathcal{D}(\mathbb{P},\mathbb{F}_S,W;\mathbb{R}^d)F, X\Big\rangle_{L^2(\lambda_{[S,T]};\mathbb{R}^{d\times m})}\Big]. \tag{93}$$

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

$$\int_S^T X_r\,\delta W_r^{\mathbb{F}_S} = \delta(\mathbb{P},\mathbb{F}_S,W;\mathbb{R}^d)(X). \tag{94}$$

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

$$\int_S^T X_r\,\delta W_r = \int_S^T X_r\,\delta W_r^{\mathfrak{S}(\mathcal{N})} \tag{95}$$

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

$$\mathbb{D}^{(1,2)}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d) \subseteq \mathbb{D}^{(1,2)}(\mathbb{P}, \mathbb{F}_{[S,s]\cup[t,T]}, W|_{[s,t]\times\Omega}; \mathbb{R}^d) \tag{96}$$

and for all $F \in \mathbb{D}^{(1,2)}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d)$ it holds $\lambda_{[s,t]} \otimes \mathbb{P}$-a.e. that

$$\Big(\mathcal{D}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d)F\Big)\Big|_{[s,t]\times\Omega} = \mathcal{D}(\mathbb{P}, \mathbb{F}_{[S,s]\cup[t,T]}, W|_{[s,t]\times\Omega}; \mathbb{R}^d)F. \tag{97}$$

Proof of Lemma A.6. Throughout this proof let $F \in \mathcal{S}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d)$, let $n \in \mathbb{N}$, $\phi_1, \ldots, \phi_n \in \mathcal{L}^2(\lambda_{[S,T]}; \mathbb{R}^m)$, $f \in C_b^{\infty, \mathfrak{S}(\mathbb{F}_S \cup \mathcal{N})}(\mathbb{R}^n \times \Omega, \mathbb{R})$, and $h \in \mathbb{R}^d$ satisfy that it holds $\mathbb{P}$-a.s. that

$$F = f\Big(\int_S^T \phi_1(r)\, dW_r, \ldots, \int_S^T \phi_n(r)\, dW_r\Big)h, \tag{98}$$

and let $g \in C_b^{\infty, \mathbb{F}_{[S,s]\cup[t,T]}}(\mathbb{R}^n \times \Omega, \mathbb{R})$ be a function such that for all $(x_1, \ldots, x_n) \in \mathbb{R}^n$ it holds $\mathbb{P}$-a.s. that

$$g(x_1, \ldots, x_n) = f\Big(x_1 + \int_S^s \phi_1(r)\, dW_r + \int_t^T \phi_1(r)\, dW_r, \ldots, x_n + \int_S^s \phi_n(r)\, dW_r + \int_t^T \phi_n(r)\, dW_r\Big). \tag{99}$$

Then it holds $\mathbb{P}$-a.s. that

$$F = g\Big(\int_s^t \phi_1(r)\, dW_r, \ldots, \int_s^t \phi_n(r)\, dW_r\Big)h. \tag{100}$$

This implies that $F \in \mathcal{S}(\mathbb{P}, \mathbb{F}_{[S,s]\cup[t,T]}, W|_{[s,t]\times\Omega}; \mathbb{R}^d)$. Next for all $i \in \{1, \ldots, n\}$ it holds $\mathbb{P}$-a.s. that

$$\frac{\partial f}{\partial x_i}\Big(\int_S^T \phi_1(r)\, dW_r, \ldots, \int_S^T \phi_n(r)\, dW_r\Big) = \frac{\partial g}{\partial x_i}\Big(\int_s^t \phi_1(r)\, dW_r, \ldots, \int_s^t \phi_n(r)\, dW_r\Big). \tag{101}$$

It follows that it holds $\lambda_{[s,t]} \otimes \mathbb{P}$-a.e. that

$$\begin{aligned}
&\Big(\mathcal{D}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d)F\Big)\Big|_{[s,t]\times\Omega} \\
&= \sum_{i=1}^n \frac{\partial f}{\partial x_i}\Big(\int_S^T \phi_1(r)\, dW_r, \ldots, \int_S^T \phi_n(r)\, dW_r\Big)\big(\phi_i|_{[s,t]}\big)h = \sum_{i=1}^n \frac{\partial g}{\partial x_i}\Big(\int_s^t \phi_1(r)\, dW_r, \ldots, \int_s^t \phi_n(r)\, dW_r\Big)\big(\phi_i|_{[s,t]}\big)h \\
&= \mathcal{D}(\mathbb{P}, \mathbb{F}_{[S,s]\cup[t,T]}, W|_{[s,t]\times\Omega}; \mathbb{R}^d)F.
\end{aligned} \tag{102}$$

Equation (102) implies that

$$\begin{aligned}
&\|F\|^2_{\mathbb{D}^{(1,2)}(\mathbb{P}, \mathbb{F}_{[S,s]\cup[t,T]}, W|_{[s,t]\times\Omega}; \mathbb{R}^d)} \\
&= \mathbb{E}\Big[\|F\|^2_{\mathbb{R}^d} + \|\mathcal{D}(\mathbb{P}, \mathbb{F}_{[S,s]\cup[t,T]}, W|_{[s,t]\times\Omega}; \mathbb{R}^d)F\|^2_{L^2(\lambda_{[s,t]}; \mathbb{R}^d)}\Big] \\
&= \mathbb{E}\Big[\|F\|^2_{\mathbb{R}^d} + \Big\|\Big(\mathcal{D}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d)F\Big)\Big|_{[s,t]\times\Omega}\Big\|^2_{L^2(\lambda_{[s,t]}; \mathbb{R}^d)}\Big] \\
&\leq \mathbb{E}\big[\|F\|^2_{\mathbb{R}^d} + \|\mathcal{D}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d)F\|^2_{L^2(\lambda_{[S,T]}; \mathbb{R}^d)}\big] = \|F\|^2_{\mathbb{D}^{(1,2)}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d)}.
\end{aligned} \tag{103}$$

Since $F \in \mathcal{S}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d)$ was chosen arbitrarily it follows that

$$\operatorname{span}(\mathcal{S}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d)) \subseteq \operatorname{span}(\mathcal{S}(\mathbb{P}, \mathbb{F}_{[S,s]\cup[t,T]}, W|_{[s,t]\times\Omega}; \mathbb{R}^d)). \tag{104}$$

This and inequality (103) yield the inclusion (96), and equation (102) implies equation (97). The proof of Lemma A.6 is thus completed. □

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

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

$$\int_s^t X_r \, \delta W_r^{\mathbb{F}_{[S,s]\cup[t,T]}} = \int_S^T \mathbb{1}_{[s,t]}(r) X_r \, \delta W_r^{\mathbb{F}_S}. \tag{105}$$

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 $\mathbb{1}_{[s,t]}X \in L^2(\mathbb{P}|_{\mathbb{F}_T}; L^2(\lambda_{[S,T]}; \mathbb{R}^{d\times m}))$. Lemma A.6, the definition of the Skorohod integral, and the Cauchy-Schwarz inequality imply for all $F \in \mathbb{D}^{(1,2)}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d)$ that

$$\begin{aligned}
&\mathbb{E}\Big[\Big\langle \mathcal{D}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d)F, \mathbb{1}_{[s,t]}X \Big\rangle_{L^2(\lambda_{[S,T]};\mathbb{R}^{d\times m})}\Big] \\
&= \mathbb{E}\Big[\Big\langle (\mathcal{D}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d)F)|_{[s,t]\times\Omega}, X|_{[s,t]\times\Omega} \Big\rangle_{L^2(\lambda_{[s,t]};\mathbb{R}^{d\times m})}\Big] \\
&= \mathbb{E}\Big[\Big\langle \mathcal{D}(\mathbb{P}, \mathbb{F}_{[S,s]\cup[t,T]}, W|_{[s,t]\times\Omega}; \mathbb{R}^d)F, X|_{[s,t]\times\Omega} \Big\rangle_{L^2(\lambda_{[s,t]};\mathbb{R}^{d\times m})}\Big] \\
&= \mathbb{E}\Big[\Big\langle F, \int_s^t X_r \, \delta W_r^{\mathbb{F}_{[S,s]\cup[t,T]}} \Big\rangle_{\mathbb{R}^d}\Big] \\
&\leq \Big\| \int_s^t X_r \, \delta W_r^{\mathbb{F}_{[S,s]\cup[t,T]}} \Big\|_{L^2(\mathbb{P};\mathbb{R}^d)} \cdot \|F\|_{L^2(\mathbb{P};\mathbb{R}^d)} < \infty.
\end{aligned} \tag{106}$$

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

'(ii) implies (i)': Assume, that $\mathbb{1}_{[s,t]}X$ is $(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d)$-Skorohod integrable. This implies that it holds that $X|_{[s,t]\times\Omega} \in L^2(\mathbb{P}|_{\mathbb{F}_T}; L^2(\lambda_{[s,t]}; \mathbb{R}^{d\times m}))$. Lemma A.6 and the definition of the Skorohod integral yield for all $F \in \mathbb{D}^{(1,2)}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d)$ that $F \in \mathbb{D}^{(1,2)}(\mathbb{P}, \mathbb{F}_{[S,s]\cup[t,T]}, W|_{[s,t]\times\Omega}; \mathbb{R}^d)$ and that

$$\begin{aligned}
&\mathbb{E}\Big[\Big\langle \big(\mathcal{D}(\mathbb{P}, \mathbb{F}_{[S,s]\cup[t,T]}, W|_{[s,t]\times\Omega}; \mathbb{R}^d)F\big), X\big|_{[s,t]\times\Omega} \Big\rangle_{L^2(\lambda_{[s,t]};\mathbb{R}^{d\times m})}\Big] \\
&= \mathbb{E}\Big[\Big\langle \big(\mathcal{D}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d)F\big)\big|_{[s,t]\times\Omega}, X\big|_{[s,t]\times\Omega} \Big\rangle_{L^2(\lambda_{[s,t]};\mathbb{R}^{d\times m})}\Big] \\
&= \mathbb{E}\big[\big\langle \mathcal{D}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d)F, \mathbb{1}_{[s,t]}X \big\rangle_{L^2(\lambda_{[S,T]};\mathbb{R}^{d\times m})}\big] \\
&= \mathbb{E}\big[\big\langle F, \int_S^T \mathbb{1}_{[s,t]}(r) X_r \, \delta W_r^{\mathbb{F}_S} \big\rangle_{\mathbb{R}^d}\big] \\
&\leq \Big\| \int_S^T \mathbb{1}_{[s,t]}(r) X_r \, \delta W_r^{\mathbb{F}_S} \Big\|_{L^2(\mathbb{P};\mathbb{R}^d)} \cdot \|F\|_{L^2(\mathbb{P};\mathbb{R}^d)} < \infty.
\end{aligned} \tag{107}$$

Lemma A.4 shows that $\operatorname{span}(\mathcal{S}(\mathbb{P}, \mathbb{F}_S, W; \mathbb{R}^d))$ is dense in $L^2(\mathbb{P}|_{\mathbb{F}_T}; \mathbb{R}^d)$. This, (106) (107), 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 $\mathbb{P}$-a.s. that

$$\int_s^t X_r \, \delta W_r^{\mathbb{F}_{[S,s]\cup[t,T]}} = \int_S^T \mathbb{1}_{[s,t]}(r) X_r \, \delta W_r^{\mathbb{F}_S}. \tag{108}$$

The proof of Proposition A.7 is thus completed. □

It is well-known (e.g., Nualart [36, 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.8 is analogous to the proof of Nualart [36, 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 $\tilde{\mathbb{F}} = (\tilde{\mathbb{F}}_r)_{r\in[s,t]}$ be a filtration with the property that for all $r \in [s,t]$ it holds that $\tilde{\mathbb{F}}_r = \mathfrak{S}(\mathfrak{S}(W_u - W_s \colon u \in [s,r]) \cup \mathbb{F}_{[S,s]\cup[t,T]})$ and let $X \in \mathcal{L}^2(\mathbb{P};\mathcal{L}^2(\lambda_{[s,t]};\mathbb{R}^{d\times m}))$ be $\tilde{\mathbb{F}}$-predictable. Then X is $(\mathbb{P},\mathbb{F}_{[S,s]\cup[t,T]},W|_{[s,t]\times\Omega};\mathbb{R}^d)$-Skorohod integrable and it holds $\mathbb{P}$-a.s. that*

$$\int_s^t X_r\,\delta W_r^{\mathbb{F}_{[S,s]\cup[t,T]}} = \int_s^t X_r\,dW_r. \tag{109}$$

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

Acknowledgement

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

References

[1] Alekseev, V. *An estimate for the perturbations of the solution of ordinary differential equations (Russian).* Vestn. Mosk.Univ., Ser.I, Math. Meh, 2, 1961.

[2] Alòs, E., and Nualart, D. An extension of Itô's formula for anticipating processes. *Journal of Theoretical Probability 11*, 2 (1998), 493–514.

[3] Becker, S., Gess, B., Jentzen, A., and Kloeden, P. E. Strong convergence rates for explicit space-time discrete numerical approximations of stochastic Allen–Cahn equations. *arXiv preprint arXiv:1711.02423* (2017).

[4] Becker, S., and Jentzen, A. Strong convergence rates for nonlinearity-truncated Euler-type approximations of stochastic Ginzburg–Landau equations. *Stochastic Processes and their Applications 129*, 1 (2019), 28–69.

[5] Bréhier, C.-E., Cui, J., and Hong, J. Strong convergence rates of semi-discrete splitting approximations for stochastic Allen–Cahn equation. *arXiv preprint arXiv:1802.06372* (2018).

[6] Bréhier, C.-E., and Goudenège, L. Analysis of some splitting schemes for the stochastic Allen–Cahn equation. *arXiv preprint arXiv:1801.06455* (2018).

[7] Cox, S. G., Hutzenthaler, M., and Jentzen, A. Local Lipschitz continuity in the initial value and strong completeness for nonlinear stochastic differential equations. *arXiv:1309.5595v1* (2013), 1–54.

[8] Debrabant, K., and Rössler, A. Continuous weak approximation for stochastic differential equations. *Journal of computational and applied mathematics 214*, 1 (2008), 259–273.

[9] Friz, P. K., and Hairer, M. *A course on rough paths: with an introduction to regularity structures*. Springer, 2014.

[10] Graham, C., and Talay, D. *Stochastic simulation and Monte Carlo methods: mathematical foundations of stochastic simulation*, vol. 68. Springer Science & Business Media, 2013.

[11] Gröbner, W. *Die Lie-Reihen und ihre Anwendungen*. VEB Deutscher Verlag der Wiss., Berlin, 1960.

[12] Gyöngy, I., and Millet, A. Rate of convergence of implicit approximations for stochastic evolution equations. In *Stochastic differential equations: theory and applications*, vol. 2 of *Interdiscip. Math. Sci.* World Sci. Publ., Hackensack, NJ, 2007, pp. 281–310.

[13] Hairer, E., Nørsett, S. P., and Wanner, G. *Solving Ordinary Differential Equations I*, second edition ed. Springer-Verlag, Berlin Heidelberg New York, 1993.

[14] Hairer, M., Hutzenthaler, M., and Jentzen, A. Loss of regularity for Kolmogorov equations. *Ann. Probab. 43*, 2 (2015), 468–527.

[15] Hutzenthaler, M., and Jentzen, A. On a perturbation theory and on strong convergence rates for stochastic ordinary and partial differential equations with non-globally monotone coefficients. *arXiv:1401.0295* (2014), 1–41.

[16] Hutzenthaler, M., and Jentzen, A. Numerical approximations of stochastic differential equations with non-globally Lipschitz continuous coefficients. *Mem. Amer. Math. Soc. 4* (2015), 1–112.

[17] Hutzenthaler, M., Jentzen, A., and Kloeden, P. E. Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 467* (2011), 1563–1576.

[18] Hutzenthaler, M., Jentzen, A., and Kloeden, P. E. Strong convergence of an explicit numerical method for SDEs with non-globally Lipschitz continuous coefficients. *Ann. Appl. Probab. 22*, 4 (2012), 1611–1641.

[19] Hutzenthaler, M., Jentzen, A., and Kloeden, P. E. Divergence of the multilevel Monte Carlo Euler method for nonlinear stochastic differential equations. *Ann. Appl. Probab. 23*, 5 (2013), 1913–1966.

[20] Hutzenthaler, M., Jentzen, A., and Wang, X. Exponential integrability properties of numerical approximation processes for nonlinear stochastic differential equations. *Math. Comp. 87*, 311 (2018), 1353–1413.

[21] Iserles, A. *A first course in the numerical analysis of differential equations*, second ed. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2009.

[22] Iserles, A., and Söderlind, G. Global bounds on numerical error for ordinary differential equations. *Journal of Complexity 9*, 1 (1993), 97–112.

[23] Jentzen, A., and Pušnik, P. Strong convergence rates for an explicit numerical approximation method for stochastic evolution equations with non-globally Lipschitz continuous nonlinearities. *arXiv:1504.03523* (2015), 1–38.

[24] Kato, T. *Perturbation theory for linear operators*, vol. 132. Springer Science & Business Media, 1980.

[25] Klenke, A. *Probability theory*. Universitext. Springer-Verlag London Ltd., London, 2008. A comprehensive course, Translated from the 2006 German original.

[26] Kovács, M., Larsson, S., and Lindgren, F. On the backward Euler approximation of the stochastic Allen–Cahn equation. *J. Appl. Probab. 52*, 2 (2015), 323–338.

[27] Kruse, R. *Strong and weak approximation of semilinear stochastic evolution equations*, vol. 2093 of *Lecture Notes in Mathematics*. Springer, Cham, 2014.

[28] 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.

[29] Li, X.-M. Strong p-completeness of stochastic differential equations and the existence of smooth flows on noncompact manifolds. *Probab. Theory Related Fields 100*, 4 (1994), 485–511.

[30] Li, X.-M., and Scheutzow, M. Lack of strong completeness for stochastic flows. *Ann. Probab. 39*, 4 (2011), 1407–1421.

[31] Lie, I., and Nørsett, S. P. Superconvergence for multistep collocation. *Mathematics of computation 52*, 185 (1989), 65–79.

[32] Liu, Z., and Qiao, Z. Strong approximation of monotone stochastic partial different equations driven by multiplicative noise. *arXiv preprint arXiv:1811.05392* (2018).

[33] Liu, Z., and Qiao, Z. Strong approximation of stochastic Allen-Cahn equation with white noise. *arXiv preprint arXiv:1801.09348* (2018).

[34] Majee, A. K., and Prohl, A. Optimal strong rates of convergence for a space-time discretization of the stochastic Allen–Cahn equation with multiplicative noise. *Computational Methods in Applied Mathematics 18*, 2 (2018), 297–311.

[35] Nørsett, S. P., and Wanner, G. The real-pole sandwich for rational approximations and oscillation equations. *BIT Numerical Mathematics 19*, 1 (1979), 79–94.

[36] Nualart, D. *The Malliavin calculus and related topics*, vol. 1995. Springer, 2006.

[37] Revuz, D., and Yor, M. *Continuous martingales and Brownian motion*, second ed., vol. 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1994.

[38] Talay, D., and Tubaro, L. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic analysis and applications 8*, 4 (1990), 483–509.

[39] Tambue, A., and Ngnotchouye, J. M. T. Weak convergence for a stochastic exponential integrator and finite element discretization of stochastic partial differential equation with multiplicative & additive noise. *Applied Numerical Mathematics 108* (2016), 57–86.

[40] Wang, X. An efficient explicit full discrete scheme for strong approximation of stochastic Allen–Cahn equation. *arXiv preprint arXiv:1802.09413* (2018).

[41] Zhang, X. Stochastic flows and Bismut formulas for stochastic Hamiltonian systems. *Stochastic Process. Appl. 120*, 10 (2010), 1929–1949.