Overcoming the curse of dimensionality in the approximative pricing of financial derivatives with default risks

Electronic Journal of Probability

Electron. J. Probab. **25** (2020), article no. 101, 1–73.
ISSN: 1083-6489 https://doi.org/10.1214/20-EJP423

Overcoming the curse of dimensionality in the approximative pricing of financial derivatives with default risks

Martin Hutzenthaler* Arnulf Jentzen†‡
Philippe von Wurstemberger§

Abstract

Parabolic partial differential equations (PDEs) are widely used in the mathematical modeling of natural phenomena and man-made complex systems. In particular, parabolic PDEs are a fundamental tool to approximately determine fair prices of financial derivatives in the financial engineering industry. The PDEs appearing in financial engineering applications are often nonlinear (e.g., in PDE models which take into account the possibility of a defaulting counterparty) and high-dimensional since the dimension typically corresponds to the number of considered financial assets. A major issue in the scientific literature is that most approximation methods for nonlinear PDEs suffer from the so-called curse of dimensionality in the sense that the computational effort to compute an approximation with a prescribed accuracy grows exponentially in the dimension of the PDE or in the reciprocal of the prescribed approximation accuracy and nearly all approximation methods for nonlinear PDEs in the scientific literature have not been shown not to suffer from the curse of dimensionality. Recently, a new class of approximation schemes for semilinear parabolic PDEs, termed *full history recursive multilevel Picard (MLP)* algorithms, were introduced and it was proven that MLP algorithms do overcome the curse of dimensionality for semilinear heat equations. In this paper we extend and generalize those findings to a more general class of semilinear PDEs which includes as special cases the important examples of semilinear Black-Scholes equations used in pricing models for financial derivatives with default risks. In particular, we introduce an MLP algorithm for the approximation of solutions of semilinear Black-Scholes equations and prove, under the assumption that the nonlinearity in the PDE is globally Lipschitz continuous, that the computational effort of the proposed method grows at most polynomially in both the

*Faculty of Mathematics, University of Duisburg-Essen, Germany. E-mail: martin.hutzenthaler@uni-due.de

†Department of Mathematics, ETH Zurich, Switzerland. E-mail: arnulf.jentzen@sam.math.ethz.ch

‡Faculty of Mathematics and Computer Science, University of Münster, Germany. E-mail: ajentzen@uni-muenster.de

§Department of Mathematics, ETH Zurich, Switzerland. E-mail: philippe.vonwurstemberger@math.ethz.ch

dimension and the reciprocal of the prescribed approximation accuracy. We thereby establish, for the first time, that the numerical approximation of solutions of semilinear Black-Scholes equations is a polynomially tractable approximation problem.

Keywords: curse of dimensionality; high-dimensional PDEs; semilinear PDEs; semilinear Kolmogorov PDEs; multilevel Picard method.
MSC2020 subject classifications: 60H35.
Submitted to EJP on July 30, 2019, final version accepted on January 25, 2020.
Supersedes arXiv:1903.05985.

1 Introduction

Parabolic partial differential equations (PDEs) are key mathematical tools to model natural phenomena and man-made complex systems. In particular, parabolic PDEs are used in the financial industry to model fair prices of financial derivatives. The use of PDEs for option pricing originated in the work of Black, Scholes, & Merton (see [12, 80]) which suggested that the price of a financial derivative satisfies a linear parabolic PDE, nowadays known as Black-Scholes equation. The derivation of their theory is based on several assumptions which are not met in the financial practice and consequently various changes and extensions to the original pricing model have been developed. One key modification of the initial Black-Scholes model is to include the possibility of a defaulting counterparty (cf., e.g., Burgard & Kjaer [17], Crepey et al. [28], Duffie et al. [37], and Henry-Labordere [56]). Such extended models suggest that the price process of a financial derivative satisfies a certain semilinear PDE (cf. (1.1) in Theorem 1.1 below and Subsections 4.2–4.3 below). Typically, semilinear parabolic PDEs can not be solved explicitly and it is therefore a very active topic of research to solve such PDEs approximately (see below for a rough overview of the literature).

The PDEs appearing in financial engineering applications are often high-dimensional since the dimension corresponds to the number of financial assets (such as stocks, commodities, interest rates, or exchange rates) in the involved hedging portfolio. A major issue in the scientific literature is that most approximation methods for nonlinear PDEs suffer from the so-called curse of dimensionality (see Bellman [8]) in the sense that the computational effort to compute an approximation with a prescribed accuracy $\varepsilon > 0$ grows exponentially in the dimension $d \in \mathbb{N}$ of the PDE or in the reciprocal ${}^1\!/\!_\varepsilon$ of the prescribed approximation accuracy (cf., e.g., E et al. [40, Section 4] for a discussion of the curse of dimensionality in the PDE approximation literature) and nearly all approximation methods for nonlinear PDEs have not been shown not to suffer from the curse of dimensionality. Recently, a new class of approximation schemes for semilinear parabolic PDEs, termed *full history recursive multilevel Picard (MLP)* algorithms, were introduced in E et al. [39, 40] and it was proven, under restrictive assumptions on the regularity of the solution of the PDE that they overcome the curse of dimensionality for semilinear heat equations. Building on this work, [63] proposed for semilinear heat equations an adaption of the original MLP scheme in [39, 40]. Under the assumption that the nonlinearity in the PDE is globally Lipschitz continuous [63, Theorem 1.1] proves that the proposed scheme does indeed overcome the curse of dimensionality in the sense that the computational effort to compute an approximation with a prescribed accuracy $\varepsilon > 0$ grows at most polynomially in both the dimension $d \in \mathbb{N}$ of the PDE and the reciprocal ${}^1\!/\!_\varepsilon$ of the prescribed approximation accuracy.

In this paper we generalize the MLP algorithm of [63] and the main result of this article, Theorem 3.24 below, proves that the MLP algorithm proposed in this paper overcomes the curse of dimensionality for a more general class of semilinear PDEs which includes as special cases the important examples of semilinear Black-Scholes equations used in pricing models for financial derivatives with default risks. In particular, we show for the first time that the solution of a semilinear Black-Scholes PDE with a globally Lipschitz continuous nonlinearity can be approximated with a computational effort which grows at most polynomially in both the dimension and the reciprocal of the prescribed approximation accuracy. Put differently, we show that the approximation of solutions of such semilinear Black-Scholes equations is a polynomially tractable approximation problem (cf., e.g., Novak & Wozniakowski [85]). To illustrate the main result of this paper, Theorem 3.24 below, we present in the following theorem, Theorem 1.1 below, a special case of Theorem 3.24. Theorem 1.1 demonstrates that the MLP algorithm proposed in this article overcomes the curse of dimensionality for the approximation of solutions of certain semilinear Black-Scholes equations.

Theorem 1.1. *Let $T \in (0,\infty)$, $p,\mathfrak{P},q \in [0,\infty)$, $\alpha,\beta \in \mathbb{R}$, $\Theta = \cup_{n=1}^{\infty}\mathbb{Z}^n$, let $f\colon \mathbb{R} \to \mathbb{R}$ be a Lipschitz continuous function, for every $d \in \mathbb{N}$ let $\|\cdot\|_{\mathbb{R}^d} : \mathbb{R}^d \to [0,\infty)$ be the Euclidean norm on $\mathbb{R}^d$, let $\xi_d \in \mathbb{R}^d$, $d \in \mathbb{N}$, and $g_d \in C^2(\mathbb{R}^d,\mathbb{R})$, $d \in \mathbb{N}$, satisfy that $\sup_{d\in\mathbb{N},x\in\mathbb{R}^d}\big(\frac{|g_d(x)|}{d^{\mathfrak{P}}(1+\|x\|_{\mathbb{R}^d}^p)} + \frac{\|\xi_d\|_{\mathbb{R}^d}}{d^q}\big) < \infty$, let $u_d \in C^{1,2}([0,T]\times\mathbb{R}^d,\mathbb{R})$, $d \in \mathbb{N}$, be polynomially growing functions which satisfy for all $d \in \mathbb{N}$, $t \in (0,T)$, $x = (x_1,x_2,\ldots,x_d) \in \mathbb{R}^d$ that $u_d(T,x) = g_d(x)$ and*

$$\big(\tfrac{\partial u_d}{\partial t}\big)(t,x) + \left[\sum_{i=1}^{d} \tfrac{|\beta|^2|x_i|^2}{2}\big(\tfrac{\partial^2 u_d}{\partial (x_i)^2}\big)(t,x)\right] + \left[\sum_{i=1}^{d} \alpha x_i \big(\tfrac{\partial u_d}{\partial x_i}\big)(t,x)\right] + f(u_d(t,x)) = 0, \quad (1.1)$$

let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $\mathcal{R}^\theta\colon \Omega \to [0,1]$, $\theta \in \Theta$, be independent $\mathcal{U}_{[0,1]}$-distributed random variables, let $R^\theta = (R^\theta_t)_{t\in[0,T]}\colon [0,T]\times\Omega \to [0,T]$, $\theta \in \Theta$, be the stochastic processes which satisfy for all $t \in [0,T]$, $\theta \in \Theta$ that $R^\theta_t = t + (T-t)\mathcal{R}^\theta$, let $W^{d,\theta} = (W^{d,\theta,i})_{i\in\{1,2,\ldots,d\}}\colon [0,T]\times\Omega \to \mathbb{R}^d$, $\theta \in \Theta$, $d \in \mathbb{N}$, be independent standard Brownian motions, assume that $(W^{d,\theta})_{d\in\mathbb{N},\theta\in\Theta}$ and $(\mathcal{R}^\theta)_{\theta\in\Theta}$ are independent, for every $d \in \mathbb{N}$, $\theta \in \Theta$, $t \in [0,T]$, $s \in [t,T]$, $x = (x_1,x_2,\ldots,x_d) \in \mathbb{R}^d$ let $X^{d,\theta,x}_{t,s} = (X^{d,\theta,x,i}_{t,s})_{i\in\{1,2,\ldots,d\}}\colon \Omega \to \mathbb{R}^d$ be the function which satisfies for all $i \in \{1,2,\ldots,d\}$ that

$$X^{d,\theta,x,i}_{t,s} = x_i \exp\!\big(\big(\alpha - \tfrac{\beta^2}{2}\big)(s-t) + \beta\big(W^{d,\theta,i}_s - W^{d,\theta,i}_t\big)\big), \quad (1.2)$$

let $V^{d,\theta}_{M,n}\colon [0,T]\times\mathbb{R}^d\times\Omega \to \mathbb{R}$, $M,n \in \mathbb{Z}$, $\theta \in \Theta$, $d \in \mathbb{N}$, be functions which satisfy for all $d,M,n \in \mathbb{N}$, $\theta \in \Theta$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that $V^{d,\theta}_{M,0}(t,x) = 0$ and

$$\begin{aligned} V^{d,\theta}_{M,n}(t,x) = \sum_{k=0}^{n-1} \frac{(T-t)}{M^{n-k}} \Bigg[\sum_{m=1}^{M^{n-k}} & f\Big(V^{d,(\theta,k,m)}_{M,k}\big(R^{(\theta,k,m)}_t, X^{d,(\theta,k,m),x}_{t,R^{(\theta,k,m)}_t}\big)\Big) \\ & - \mathbb{1}_{\mathbb{N}}(k) f\Big(V^{d,(\theta,k,-m)}_{M,k-1}\big(R^{(\theta,k,m)}_t, X^{d,(\theta,k,m),x}_{t,R^{(\theta,k,m)}_t}\big)\Big)\Bigg] + \left[\sum_{m=1}^{M^n} \frac{g_d(X^{d,(\theta,n,-m),x}_{t,T})}{M^n}\right], \end{aligned} \quad (1.3)$$

and for every $d,n,M \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ let $\mathcal{C}_{d,M,n} \in \mathbb{N}_0$ be the number of realizations of one-dimensional standard normal random variables which are used to compute one realization of $V^{d,0}_{M,n}(t,x)$ (see (4.42) below for a precise definition). Then there exist functions $N = (N_{d,\varepsilon})_{d\in\mathbb{N},\varepsilon\in(0,1]}\colon \mathbb{N}\times(0,1] \to \mathbb{N}$ and $c = (c_\delta)_{\delta\in(0,\infty)}\colon (0,\infty) \to (0,\infty)$ such that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$, $\delta \in (0,\infty)$ it holds that $\mathcal{C}_{d,N_{d,\varepsilon},N_{d,\varepsilon}} \le c_\delta\, d^{1+(\mathfrak{P}+qp)(2+\delta)}\varepsilon^{-(2+\delta)}$ and

$$\big(\mathbb{E}\big[|u_d(0,\xi_d) - V^{d,0}_{N_{d,\varepsilon},N_{d,\varepsilon}}(0,\xi_d)|^2\big]\big)^{1/2} \le \varepsilon. \quad (1.4)$$

Theorem 1.1 is an immediate consequence of Theorem 4.4 below. Theorem 4.4, in turn, is a consequence of Theorem 3.24 below, the main result of this paper. We now provide some explanations for Theorem 1.1. In Theorem 1.1 we present a stochastic approximation scheme (cf. $(V^{d,0}_{M,n})_{M,n,d\in\mathbb{N}}$ in Theorem 1.1 above) which is able to approximate in the strong L^2-sense the initial values $(u_d(0,\xi_d))_{d\in\mathbb{N}}$ of the solutions of uncorrelated semilinear Black-Scholes equations (cf. (1.1) in Theorem 1.1 above) with a computational effort which grows at most polynomially in both the dimension $d\in\mathbb{N}$ and the reciprocal $1/\varepsilon$ of the prescribed approximation accuracy $\varepsilon>0$. The time horizon $T\in(0,\infty)$, the drift parameter $\alpha\in\mathbb{R}$, the diffusion parameter $\beta\in\mathbb{R}$, as well as the Lipschitz continuous nonlinearity $f\colon\mathbb{R}\to\mathbb{R}$ of the semilinear Black-Scholes equations in Theorem 1.1 above (cf. (1.1) in Theorem 1.1 above) are fixed over all dimensions (cf. Theorem 4.3 for a more general result with dimension-dependent drift and diffusion coefficients and dimension-dependent nonlinearities which may additionally depend on the time and the space variable). The approximation points $\xi_d\in\mathbb{R}^d$, $d\in\mathbb{N}$, and the terminal conditions $g_d\colon\mathbb{R}^d\to\mathbb{R}$, $d\in\mathbb{N}$, of the PDEs in (1.1) in Theorem 1.1 above are both allowed to grow in a certain polynomial fashion determined by the constants $p,\mathfrak{P},q\in[0,\infty)$. The idea for the full history multilevel Picard scheme (cf. $(V^{d,\theta}_{M,n})_{M,d\in\mathbb{N},n\in\mathbb{N}_0,\theta\in\Theta}$ in Theorem 1.1 above) is based on a reformulation of the semilinear PDEs in (1.1) as stochastic fixed point equations. More precisely, the Feynman-Kac formula (see Proposition 3.22 below and, e.g., Beck et al. [5, Theorem 1.1, Theorem 3.7, & Corollary 3.9]) and the fact that for all $t\in[0,T]$ it holds that R^θ_t, $\theta\in\Theta$, are independent $\mathcal{U}_{[t,T]}$-distributed random variables ensure that the solutions $u_d\colon[0,T]\times\mathbb{R}^d\to\mathbb{R}$, $d\in\mathbb{N}$, of the PDEs in (1.1) and the solution processes $X^{d,\theta}_{t,\cdot}(x)=(X^{d,\theta}_{t,s}(x))_{s\in[t,T]}\colon[t,T]\times\Omega\to\mathbb{R}^d$, $t\in[0,T]$, $x\in\mathbb{R}^d$, $d\in\mathbb{N}$, $\theta\in\Theta$, of the stochastic differential equations (SDEs) associated to the PDEs in (1.1) satisfy that for all $d\in\mathbb{N}$, $\theta\in\Theta$, $t\in[0,T]$, $x\in\mathbb{R}^d$ it holds that

$$u_d(t,x)=\mathbb{E}\Big[g\big(X^{d,\theta,x}_{t,T}\big)+(T-t)f\big(u_d(R^\theta_t,X^{d,\theta,x}_{t,R^\theta_t})\big)\Big]\,. \tag{1.5}$$

Moreover, note that (1.3) assures that for all $d,M,n\in\mathbb{N}$, $\theta\in\Theta$, $t\in[0,T]$, $x\in\mathbb{R}^d$ it holds that

$$\begin{aligned}\mathbb{E}\big[V^{d,\theta}_{M,n}(t,x)\big]&=\sum_{k=0}^{n-1}(T-t)\mathbb{E}\Big[f\big(V^{d,(\theta,1)}_{M,k}\big(R^{(\theta,1)}_t,X^{d,(\theta,1),x}_{t,R^{(\theta,1)}_t}\big)\big)\\&\qquad-\mathbb{1}_{\mathbb{N}}(k)f\big(V^{d,(\theta,-1)}_{M,k-1}\big(R^{(\theta,1)}_t,X^{d,(\theta,1),x}_{t,R^{(\theta,1)}_t}\big)\big)\Big]+\mathbb{E}\big[g_d(X^{d,\theta,x}_{t,T})\big]\\&=\mathbb{E}\Big[g_d(X^{d,\theta,x}_{t,T})+(T-t)f\big(V^{d,\theta}_{M,n-1}\big(R^\theta_t,X^{d,\theta,x}_{t,R^\theta_t}\big)\big)\Big]\,.\end{aligned} \tag{1.6}$$

Combining this with (1.5) illustrates that, roughly speaking, for every $d,M\in\mathbb{N}$, $\theta\in\Theta$ the sequence of random fields $V^{d,\theta}_{M,n}\colon\Omega\times[0,T]\times\mathbb{R}^d\to\mathbb{R}$, $n\in\mathbb{N}_0$, behave, in expectation, like Picard iterations for the stochastic fixed point equation in (1.5) above. In each iteration in (1.3) the expectation of the Picard iteration for the stochastic fixed point equation in (1.5) is approximated with a multilevel Monte Carlo approach on a telescopic expansion over the full history of the previous iterations. According to the multilevel Monte Carlo paradigm the number of samples in each level is chosen such that computationally inexpensive summands (corresponding to small $k\in\{0,1,2,\dots,n-1\}$ in (1.6)) of the telescope expansion get sampled more often than computationally expensive ones (corresponding to large $k\in\{0,1,2,\dots,n-1\}$ in (1.6)). Roughly speaking, the conclusion of Theorem 1.1 above states that for every $d\in\mathbb{N}$, $\varepsilon\in(0,1]$ there exists a natural number $N\in\mathbb{N}$ such that (i) it holds that $V^{d,0}_{N,N}(0,\xi_d)$ approximates $u_d(0,\xi_d)$ in the strong L^2-sense with accuracy ε and such that (ii) it holds that the computational effort to compute $V^{d,0}_{N,N}(0,\xi_d)$ is essentially of the order $d^{1+2(\mathfrak{P}+pq)}\varepsilon^{-2}$ (cf. (1.4) in Theorem 1.1 above for the precise formulation). Remarkably, this is exactly the computational complexity of standard Monte Carlo approximations of the solutions of the PDEs in (1.1) in the special case where the nonlinearity f vanishes (cf., e.g., Graham & Talay [51]).

We now point out a selection of other approaches from the scientific literature to numerically approximate solutions of nonlinear parabolic PDEs. Deterministic approximation methods, such as finite element and sparse-grid methods, can be found, for example, in [34, 98, 99, 101]. There are also a number of approximation methods for nonlinear parabolic PDEs in the scientific literature whose derivation is based on probabilistic concepts. Several methods leverage a connection between parabolic PDEs and backward stochastic differential equations (BSDEs). For example, [2, 9, 10, 13, 15, 16, 20, 21, 22, 23, 29, 30, 31, 32, 42, 43, 35, 36, 44, 45, 46, 47, 48, 49, 50, 60, 73, 74, 75, 76, 77, 78, 82, 83, 86, 87, 88, 89, 93, 94, 95, 100, 106, 107, 108] use discretizations of the associated first-order BSDEs and [14, 24, 41, 53, 72, 109] use discretizations of the associated second-order BSDEs to conceive approximation methods for parabolic PDEs. Other probabilistic approaches include approximation methods based on branching diffusion processes (cf., e.g., [19, 56, 58, 59, 79, 92, 97, 102, 105]) and approximation methods based on nested Monte Carlo simulations (cf., e.g., [103, 104]). Recently, a further class of methods based on deep learning has turned out to be very successful in approximating solutions of high-dimensional PDEs (cf., e.g., [3, 6, 7, 11, 18, 38, 54, 55, 57, 61, 69, 84, 91, 96]), although theoretical results assuring their convergence are still lacking (cf, e.g., [52, 62] for partial error analysis results of deep learning based methods). Lastly, we refer to, e.g., to [39, 40, 63, 64] for other MLP approximation methods, which are closely related to the method proposed in this paper.

The remainder of this paper is structured as follows. In Section 2 we prove a well-known distributional flow property for the composition of independent solutions fields of a stochastic differential equation (SDE) (see Lemma 2.20 below). This distributional flow property will be a key assumption in the abstract treatment of the stochastic fixed point equations which we study in Section 3. Section 3 also (i) introduces the MLP algorithm in Subsection 3.1, (ii) provides a complexity analysis in the setting of stochastic fixed point equations in Subsections 3.2–3.5, and (iii) carries over this complexity analysis to the case of semilinear Kolmogorov PDEs in Subsection 3.6; see Theorem 3.24 below. In the last section of this article, Section 4 below, we apply the result for general semilinear Kolmogorov PDEs in Theorem 3.24 to semilinear heat PDEs (see Subsection 4.1) and semlinear Black-Scholes PDEs (see Subsection 4.2 and Subsection 4.3). Such PDEs are widely used in derivative pricing models which aim to incorporate default risks of the involved derivative counterparty into the pricing process.

2 On a distributional flow property for stochastic differential equations (SDEs)

In our analysis of the proposed MLP algorithm in Section 3 below we will make use of random fields which satisfy a certain flow-type condition (see (3.3) in Setting 3.1 below). The main purpose of this section is to establish, under suitable assumptions, that solution processes of SDEs enjoy this flow-type property; see Lemma 2.20 in Subsection 2 below for details. In our proof of Lemma 2.20 we employ a series of elementary and well-known results which we establish in Subsections 2.1–2.7 below. Many of these elementary results are not only used in the proof of Lemma 2.20, but are also employed in our error analysis of the proposed MLP algorithm in Section 3.

2.1 Time-discrete Gronwall inequalities

In this subsection we present in Lemma 2.2 and Corollary 2.3 two elementary and well-known time-discrete Gronwall inequalities (cf., e.g., Agarwal [1]).

Lemma 2.1. *Let* $K \in \mathbb{N}$*,* $\alpha \in [0, \infty)$*,* $(\beta_k)_{k \in \{0,1,2,\dots,K-1\}} \subseteq [0, \infty)$*,* $(\epsilon_k)_{k \in \{0,1,2,\dots,K\}} \subseteq$

$[0, \infty]$ *satisfy for all* $k \in \{0, 1, 2, \dots, K-1\}$ *that* $\epsilon_k < \infty$ *and*

$$\epsilon_K \le \alpha + \left[\sum_{k=0}^{K-1} \beta_k \epsilon_k \right]. \tag{2.1}$$

Then it holds that $\epsilon_K < \infty$.

Proof of Lemma 2.1. Note that the hypothesis that $\forall\, k \in \{0, 1, 2, \dots, K-1\} \colon \epsilon_k < \infty$ implies that

$$\alpha + \left[\sum_{k=0}^{K-1} \beta_k \epsilon_k \right] < \infty. \tag{2.2}$$

This and (2.1) establishes that $\epsilon_K < \infty$. The proof of Lemma 2.1 is thus completed. □

Lemma 2.2. *Let* $N \in \mathbb{N}$, $\alpha \in [0, \infty)$, $(\beta_n)_{n \in \{0,1,2,\dots,N-1\}} \subseteq [0, \infty)$, $(\epsilon_n)_{n \in \{0,1,2,\dots,N\}} \subseteq [0, \infty]$ *satisfy for all* $n \in \{0, 1, 2, \dots, N\}$ *that*

$$\epsilon_n \le \alpha + \left[\sum_{k=0}^{n-1} \beta_k \epsilon_k \right] \tag{2.3}$$

(cf. Lemma 2.1). Then it holds for all $n \in \{0, 1, 2, \dots, N\}$ *that*

$$\epsilon_n \le \alpha \left[\prod_{k=0}^{n-1} (1 + \beta_k) \right] \le \alpha \exp\!\left(\sum_{k=0}^{n-1} \beta_k \right) < \infty. \tag{2.4}$$

Proof of Lemma 2.2. Throughout this proof let $(u_n)_{n \in \{0,1,2,\dots,N\}} \subseteq [0, \infty]$ be the extended real numbers which satisfy for all $n \in \{0, 1, 2, \dots, N\}$ that

$$u_n = \alpha + \left[\sum_{k=0}^{n-1} \beta_k u_k \right]. \tag{2.5}$$

We claim that for all $n \in \{0, 1, 2, \dots, N\}$ it holds that

$$u_n = \alpha \left[\prod_{k=0}^{n-1} (1 + \beta_k) \right]. \tag{2.6}$$

We now prove (2.6) by induction on $n \in \{0, 1, 2, \dots, N\}$. For the base case $n = 0$ observe that (2.5) ensures that

$$u_0 = \alpha. \tag{2.7}$$

This proves (2.6) in the base case $n = 0$. For the induction step $\{0, 1, 2, \dots, N-1\} \ni (n-1) \to n \in \{1, 2, \dots, N\}$ observe that (2.5) implies that for all $n \in \{1, 2, \dots, N\}$ with $u_{n-1} = \alpha \left[\prod_{k=0}^{n-2} (1 + \beta_k) \right]$ it holds that

$$\begin{aligned} u_n &= \alpha + \left[\sum_{k=0}^{n-1} \beta_k u_k \right] = \alpha + \left[\sum_{k=0}^{n-2} \beta_k u_k \right] + \beta_{n-1} u_{n-1} \\ &= u_{n-1} + \beta_{n-1} u_{n-1} = (1 + \beta_{n-1}) u_{n-1} = \alpha \left[\prod_{k=0}^{n-1} (1 + \beta_k) \right]. \end{aligned} \tag{2.8}$$

Induction thus establishes (2.6). Moreover, note that (2.3), (2.5), and induction prove that for all $n \in \{0, 1, 2, \dots, N\}$ it holds that

$$\epsilon_n \le u_n. \tag{2.9}$$

This and (2.6) establish that for all $n \in \{0, 1, 2, \ldots, N\}$ it holds that

$$\epsilon_n \leq \alpha \left[\prod_{k=0}^{n-1} (1 + \beta_k) \right]. \tag{2.10}$$

The fact that for all $x \in \mathbb{R}$ it holds that $(1 + x) \leq \exp(x)$ therefore ensures that for all $n \in \{0, 1, 2, \ldots, N\}$ it holds that

$$\epsilon_n \leq \alpha \left[\prod_{k=0}^{n-1} (1 + \beta_k) \right] \leq \alpha \left[\prod_{k=0}^{n-1} \exp(\beta_k) \right] = \alpha \exp\left(\sum_{k=0}^{n-1} \beta_k \right). \tag{2.11}$$

The proof of Lemma 2.2 is thus completed. □

Corollary 2.3. *Let $N \in \mathbb{N} \cup \{\infty\}$, $\alpha, \beta \in [0, \infty)$, $(\epsilon_n)_{n \in \mathbb{N}_0 \cap [0,N]} \subseteq [0, \infty]$ satisfy for all $n \in \mathbb{N}_0 \cap [0, N]$ that*

$$\epsilon_n \leq \alpha + \beta \left[\sum_{k=0}^{n-1} \epsilon_k \right] \tag{2.12}$$

(cf. Lemma 2.1). Then it holds for all $n \in \mathbb{N}_0 \cap [0, N]$ that

$$\epsilon_n \leq \alpha (1 + \beta)^n \leq \alpha \, e^{\beta n} < \infty. \tag{2.13}$$

Proof of Corollary 2.3. Note that Lemma 2.2 establishes Corollary 2.3. The proof of Corollary 2.3 is thus completed. □

2.2 A priori moment bounds for solutions of SDEs

In this subsection we establish in the elementary result in Lemma 2.7 below for every $p \in [0, \infty)$ a bound on the p-th absolute moment of the solution of an SDE with a deterministic initial value, a one-sided linear growth condition on the drift coefficient of the SDE, and a linear growth condition on the diffusion coefficient of the SDE (cf. (2.39) in Lemma 2.7 below). Our proof of Lemma 2.7 employs standard Lyapunov-type techniques from the literature to establish the desired a priori moment bound (cf., e.g., Beck et al. [4, Subsection 3.1] and Cox et al. [25, Section 2.2]). These Lyapunov-type techniques are the subject of the elementary and essentially well-known results in Lemma 2.4–Lemma 2.6 below.

Lemma 2.4. *Let $d, m \in \mathbb{N}$, $T, C_1, C_2 \in [0, \infty)$, let $\langle \cdot, \cdot \rangle : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be the Euclidean scalar product on $\mathbb{R}^d$, let $\|\cdot\| : \mathbb{R}^d \to [0, \infty)$ be the Euclidean norm on $\mathbb{R}^d$, let $|||\cdot||| : \mathbb{R}^{d \times m} \to [0, \infty)$ be the Frobenius norm on $\mathbb{R}^{d \times m}$, and let $\mu : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : [0, T] \times \mathbb{R}^d \to \mathbb{R}^{d \times m}$, and $V_p : \mathbb{R}^d \to (0, \infty)$, $p \in [2, \infty)$, be functions which satisfy for all $t \in [0, T]$, $x \in \mathbb{R}^d$, $p \in [2, \infty)$ that*

$$\max\{\langle x, \mu(t, x) \rangle, |||\sigma(t, x)|||^2\} \leq C_1 + C_2 \|x\|^2 \quad \text{and} \quad V_p(x) = (1 + \|x\|^2)^{p/2}. \tag{2.14}$$

Then

(i) it holds for all $p \in [2, \infty)$ that $V_p \in C^\infty(\mathbb{R}^d, (0, \infty))$ and

(ii) it holds for all $t \in [0, T]$, $x \in \mathbb{R}^d$, $p \in [2, \infty)$ that

$$\begin{aligned} &\tfrac{1}{2} \operatorname{Trace}\big(\sigma(t, x)[\sigma(t, x)]^* (\operatorname{Hess} V_p)(x)\big) + \langle \mu(t, x), (\nabla V_p)(x) \rangle \\ &\leq \tfrac{p(p+1)}{2} \big(\tfrac{p-2}{p} + C_2\big) V_p(x) + (p + 1)|C_1|^{p/2}. \end{aligned} \tag{2.15}$$

Proof of Lemma 2.4. Throughout this proof let $\sigma_{i,j}\colon [0,T]\times\mathbb{R}^d\to\mathbb{R}$, $i\in\{1,2,\dots,d\}$, $j\in\{1,2,\dots,m\}$, be the functions which satisfy for all $t\in[0,T]$, $x\in\mathbb{R}^d$ that

$$\sigma(t,x)=\begin{pmatrix}\sigma_{1,1}(t,x) & \sigma_{1,2}(t,x) & \dots & \sigma_{1,m}(t,x)\\ \sigma_{2,1}(t,x) & \sigma_{2,2}(t,x) & \dots & \sigma_{2,m}(t,x)\\ \vdots & \vdots & \ddots & \vdots\\ \sigma_{d,1}(t,x) & \sigma_{d,2}(t,x) & \dots & \sigma_{d,m}(t,x)\end{pmatrix}\in\mathbb{R}^{d\times m}. \tag{2.16}$$

Note that the chain rule, the fact that the function $\mathbb{R}^d\ni x\mapsto 1+\|x\|^2\in(0,\infty)$ is infinitely often differentiable, and the fact that for every $p\in[2,\infty)$ the function $(0,\infty)\ni s\mapsto s^{\frac{p}{2}}\in(0,\infty)$ is infinitely often differentiable establish item (i). It thus remains to prove item (ii). For this, observe that the chain rule ensures that for all $x=(x_1,\dots,x_d)\in\mathbb{R}^d$, $i,j\in\{1,2,\dots,d\}$, $p\in[2,\infty)$ it holds that

$$(\nabla V_p)(x)=\tfrac{p}{2}\left(1+\|x\|^2\right)^{\frac{p}{2}-1}\cdot(2x)=pV_p(x)\left[\tfrac{1}{1+\|x\|^2}\right]x \tag{2.17}$$

and

$$\begin{aligned}(\tfrac{\partial^2 V_p}{\partial x_i\partial x_j})(x)&=\tfrac{\partial}{\partial x_i}\left[p\left(1+\|x\|^2\right)^{\frac{p}{2}-1}x_j\right]\\ &=p\left[\tfrac{\partial}{\partial x_i}\left(1+\|x\|^2\right)^{\frac{p}{2}-1}\right]x_j+p\left(1+\|x\|^2\right)^{\frac{p}{2}-1}\left[\tfrac{\partial}{\partial x_i}x_j\right]\\ &=p(\tfrac{p}{2}-1)\left(1+\|x\|^2\right)^{\frac{p}{2}-2}\cdot(2x_i)x_j+p\left(1+\|x\|^2\right)^{\frac{p}{2}-1}\mathbb{1}_{\{i\}}(j)\\ &=p(p-2)V_p(x)\tfrac{x_ix_j}{(1+\|x\|^2)^2}+pV_p(x)\tfrac{\mathbb{1}_{\{i\}}(j)}{1+\|x\|^2}\\ &=pV_p(x)\left[(p-2)\tfrac{x_ix_j}{(1+\|x\|^2)^2}+\tfrac{\mathbb{1}_{\{i\}}(j)}{1+\|x\|^2}\right].\end{aligned} \tag{2.18}$$

This implies that for all $t\in[0,T]$, $x=(x_1,\dots,x_d)\in\mathbb{R}^d$, $p\in[2,\infty)$ it holds that

$$\begin{aligned}&\tfrac12\operatorname{Trace}\big(\sigma(t,x)[\sigma(t,x)]^*(\operatorname{Hess}V_p)(x)\big)+\langle\mu(t,x),(\nabla V_p)(x)\rangle\\ &=\tfrac12\left[\sum_{k=1}^m\sum_{i,j=1}^d\sigma_{i,k}(t,x)\sigma_{j,k}(t,x)(\tfrac{\partial^2V_p}{\partial x_i\partial x_j})(t,x)\right]+\langle\mu(t,x),(\nabla V_p)(x)\rangle\\ &=\tfrac{pV_p(x)}{2}\left(\left[\sum_{k=1}^m\sum_{i,j=1}^d\sigma_{i,k}(t,x)\sigma_{j,k}(t,x)\left((p-2)\tfrac{x_ix_j}{(1+\|x\|^2)^2}+\tfrac{\mathbb{1}_{\{i\}}(j)}{1+\|x\|^2}\right)\right]+\tfrac{2\langle\mu(t,x),x\rangle}{1+\|x\|^2}\right)\\ &=\tfrac{pV_p(x)}{2}\left(\tfrac{(p-2)}{(1+\|x\|^2)^2}\left[\sum_{k=1}^m\left[\sum_{i=1}^d\sigma_{i,k}(t,x)x_i\right]^2\right]+\tfrac{|||\sigma(t,x)|||^2}{1+\|x\|^2}+\tfrac{2\langle\mu(t,x),x\rangle}{1+\|x\|^2}\right).\end{aligned} \tag{2.19}$$

In addition, note that the Cauchy Schwarz inequality assures that for all $t\in[0,T]$, $x=(x_1,\dots,x_d)\in\mathbb{R}^d$ it holds that

$$\begin{aligned}\sum_{k=1}^m\left[\sum_{i=1}^d\sigma_{i,k}(t,x)x_i\right]^2&\le\sum_{k=1}^m\left[\sum_{i=1}^d|\sigma_{i,k}(t,x)|^2\right]\left[\sum_{i=1}^d|x_i|^2\right]\\ &=|||\sigma(t,x)|||^2\,\|x\|^2\le|||\sigma(t,x)|||^2(1+\|x\|^2).\end{aligned} \tag{2.20}$$

This, (2.14), and (2.19) demonstrate that for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $p \in [2,\infty)$ it holds that

$$\begin{aligned}
&\tfrac{1}{2}\operatorname{Trace}\big(\sigma(t,x)[\sigma(t,x)]^*(\operatorname{Hess} V_p)(x)\big) + \langle \mu(t,x), (\nabla V_p)(x)\rangle \\
&\leq \tfrac{p}{2}\left[\tfrac{(p-2)|||\sigma(t,x)|||^2}{1+\|x\|^2} + \tfrac{|||\sigma(t,x)|||^2}{1+\|x\|^2} + \tfrac{2\langle \mu(t,x),x\rangle}{1+\|x\|^2}\right] V_p(x) \\
&\leq \tfrac{p}{2}(p-2+1+2)\tfrac{(C_1+C_2\|x\|^2)}{1+\|x\|^2} V_p(x) \\
&\leq \tfrac{p(p+1)}{2}\left(C_1\left[\tfrac{V_p(x)}{1+\|x\|^2}\right] + C_2 V_p(x)\right) = \tfrac{p(p+1)}{2}\left(C_1(1+\|x\|^2)^{p/2-1} + C_2 V_p(x)\right).
\end{aligned} \tag{2.21}$$

Young's inequality (with $p = p/2$, $q = p/(p-2) = \frac{p/2}{p/2-1}$ for $p \in (2,\infty)$ in the usual notation of Young's inequality) hence proves that for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $p \in (2,\infty)$ it holds that

$$\begin{aligned}
&\tfrac{1}{2}\operatorname{Trace}\big(\sigma(t,x)[\sigma(t,x)]^*(\operatorname{Hess} V_p)(x)\big) + \langle \mu(t,x), (\nabla V_p)(x)\rangle \\
&\leq \tfrac{p(p+1)}{2}\left(\frac{|C_1|^{p/2}}{p/2} + \frac{\left|(1+\|x\|^2)^{p/2-1}\right|^{p/(p-2)}}{p/(p-2)} + C_2 V_p(x)\right) \\
&= (p+1)|C_1|^{p/2} + \left(\tfrac{p(p+1)}{2}\left(\tfrac{p-2}{p} + C_2\right)\right) V_p(x).
\end{aligned} \tag{2.22}$$

Moreover, note that (2.21) ensures that for all $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that

$$\tfrac{1}{2}\operatorname{Trace}\big(\sigma(t,x)[\sigma(t,x)]^*(\operatorname{Hess} V_2)(x)\big) + \langle \mu(t,x), (\nabla V_2)(x)\rangle \leq 3\,(C_1 + C_2 V_2(x))\,. \tag{2.23}$$

Combining this and (2.22) establishes item (ii). The proof of Lemma 2.4 is thus completed. □

Lemma 2.5. *Let $d,m \in \mathbb{N}$, $T,\rho \in [0,\infty)$, $\xi \in \mathbb{R}^d$, let $\langle\cdot,\cdot\rangle : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be the Euclidean scalar product on $\mathbb{R}^d$, let $\mu \in C([0,T]\times\mathbb{R}^d,\mathbb{R}^d)$, $\sigma \in C([0,T]\times\mathbb{R}^d,\mathbb{R}^{d\times m})$, $V \in C^2(\mathbb{R}^d,(0,\infty))$ satisfy for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that*

$$\tfrac{1}{2}\operatorname{Trace}\big(\sigma(t,x)[\sigma(t,x)]^*(\operatorname{Hess} V)(x)\big) + \langle \mu(t,x), (\nabla V)(x)\rangle \leq \rho, \tag{2.24}$$

let $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_t)_{t\in[0,T]})$ be a filtered probability space which satisfies the usual conditions, let $W\colon [0,T]\times\Omega \to \mathbb{R}^m$ be a standard $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_{t\in[0,T]}))$-Brownian motion, and let $X\colon [0,T]\times\Omega\to\mathbb{R}^d$ be an $(\mathbb{F}_t)_{t\in[0,T]}/\mathcal{B}(\mathbb{R}^d)$-adapted stochastic process with continuous sample paths which satisfies that for all $t \in [0,T]$ it holds $\mathbb{P}$-a.s. that

$$X_t = \xi + \int_0^t \mu(r,X_r)dr + \int_0^t \sigma(r,X_r)dW_r. \tag{2.25}$$

Then it holds for all $t \in [0,T]$ that

$$\mathbb{E}[V(X_t)] \leq V(\xi) + t\rho. \tag{2.26}$$

Proof of Lemma 2.5. Throughout this proof assume w.l.o.g. that $T > 0$ and let $\mathbb{V}\colon [0,T]\times\mathbb{R}^d \to (0,\infty)$ be the function which satisfies for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$\mathbb{V}(t,x) = V(x) - t\rho + T\rho. \tag{2.27}$$

Note that the fact that $V \in C^2(\mathbb{R}^d,(0,\infty))$ ensures that for all $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that

(I) $\mathbb{V} \in C^2([0,T]\times\mathbb{R}^d,(0,\infty))$,

(II) $(\frac{\partial \mathbb{V}}{\partial t})(t,x) = -\rho$,

(III) $(\nabla_x \mathbb{V})(t,x) = (\nabla V)(x)$, and

(IV) $(\operatorname{Hess}_x \mathbb{V})(t,x) = (\operatorname{Hess} V)(x)$.

Observe that items (II)–(IV) and (2.24) show that for all $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned} &(\tfrac{\partial \mathbb{V}}{\partial t})(t,x) + \tfrac{1}{2}\operatorname{Trace}\big(\sigma(t,x)[\sigma(t,x)]^*(\operatorname{Hess}_x \mathbb{V})(t,x)\big) + \langle \mu(t,x), (\nabla_x \mathbb{V})(t,x)\rangle \\ &= -\rho + \tfrac{1}{2}\operatorname{Trace}\big(\sigma(t,x)[\sigma(t,x)]^*(\operatorname{Hess} V)(x)\big) + \langle \mu(t,x), (\nabla V)(x)\rangle \\ &\leq -\rho + \rho = 0. \end{aligned} \tag{2.28}$$

Combining this with Beck et al. [4, Lemma 3.1] (with $d = d$, $m = m$, $T = T$, $\mathcal{O} = \mathbb{R}^d$, $\mu = \mu$, $\sigma = \sigma$, $V = \mathbb{V}$, $\tau = t$, $X = X$ for $t \in [0,T]$ in the notation of Beck et al. [4, Lemma 3.1]) demonstrates that for all $t \in [0,T]$ it holds that

$$\mathbb{E}[\mathbb{V}(t, X_t)] \leq \mathbb{E}[\mathbb{V}(0, X_0)] = V(\xi) + T\rho. \tag{2.29}$$

Therefore, we obtain that for all $t \in [0,T]$ it holds that

$$\begin{aligned} \mathbb{E}[V(X_t)] &= \mathbb{E}[V(X_t) - t\rho + T\rho] + t\rho - T\rho = \mathbb{E}[\,\mathbb{V}(t,X_t)] + t\rho - T\rho \\ &\leq V(\xi) + T\rho + t\rho - T\rho = V(\xi) + t\rho. \end{aligned} \tag{2.30}$$

The proof of Lemma 2.5 is thus completed. □

Lemma 2.6. *Let $d, m \in \mathbb{N}$, $T, \rho_1, \rho_2 \in [0,\infty)$, $\xi \in \mathbb{R}^d$, let $\langle \cdot,\cdot \rangle : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be the Euclidean scalar product on $\mathbb{R}^d$, let $\mu \in C([0,T] \times \mathbb{R}^d, \mathbb{R}^d)$, $\sigma \in C([0,T] \times \mathbb{R}^d, \mathbb{R}^{d\times m})$, $V \in C^2(\mathbb{R}^d, (0,\infty))$ satisfy for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that*

$$\tfrac{1}{2}\operatorname{Trace}\big(\sigma(t,x)[\sigma(t,x)]^*(\operatorname{Hess} V)(x)\big) + \langle \mu(t,x), (\nabla V)(x)\rangle \leq \rho_1 V(x) + \rho_2, \tag{2.31}$$

let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_t)_{t\in[0,T]})$ be a filtered probability space which satisfies the usual conditions, let $W\colon [0,T] \times \Omega \to \mathbb{R}^m$ be a standard $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_{t\in[0,T]}))$-Brownian motion, and let $X\colon [0,T] \times \Omega \to \mathbb{R}^d$ be an $(\mathbb{F}_t)_{t\in[0,T]}$/$\mathcal{B}(\mathbb{R}^d)$-adapted stochastic process with continuous sample paths which satisfies that for all $t \in [0,T]$ it holds $\mathbb{P}$-a.s. that

$$X_t = \xi + \int_0^t \mu(r, X_r)dr + \int_0^t \sigma(r, X_r)dW_r. \tag{2.32}$$

Then it holds for all $t \in [0,T]$ that

$$\mathbb{E}[V(X_t)] \leq e^{\rho_1 t}\left(V(\xi) + t\rho_2\right). \tag{2.33}$$

Proof of Lemma 2.6. Throughout this proof assume w.l.o.g. that $\rho_1 > 0$ (cf. Lemma 2.5) and that $T > 0$ and let $\mathbb{V}\colon [0,T] \times \mathbb{R}^d \to (0,\infty)$ be the function which satisfies for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$\mathbb{V}(t,x) = e^{-\rho_1 t}\big(V(x) + \tfrac{\rho_2}{\rho_1}\big). \tag{2.34}$$

Note that the fact that $V \in C^2(\mathbb{R}^d, (0,\infty))$ ensures that for all $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that

(I) $\mathbb{V} \in C^2([0,T] \times \mathbb{R}^d, (0,\infty))$,

(II) $(\frac{\partial \mathbb{V}}{\partial t})(t,x) = -\rho_1 e^{-\rho_1 t}(V(x) + \frac{\rho_2}{\rho_1})$,

(III) $(\nabla_x \mathbb{V})(t,x) = e^{-\rho_1 t}(\nabla V)(x)$, and

(IV) $(\operatorname{Hess}_x \mathbb{V})(t,x) = e^{-\rho_1 t}(\operatorname{Hess} V)(x)$.

Observe that items (II)–(IV) and (2.31) assure that for all $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}&(\tfrac{\partial \mathbb{V}}{\partial t})(t,x) + \tfrac{1}{2}\operatorname{Trace}\big(\sigma(t,x)[\sigma(t,x)]^*(\operatorname{Hess}_x \mathbb{V})(t,x)\big) + \langle \mu(t,x), (\nabla_x \mathbb{V})(t,x)\rangle \\ &= e^{-\rho_1 t}\Big(-\rho_1\big(V(x) + \tfrac{\rho_2}{\rho_1}\big) + \tfrac{1}{2}\operatorname{Trace}\big(\sigma(t,x)[\sigma(t,x)]^*(\operatorname{Hess} V)(x)\big) + \langle \mu(t,x), (\nabla V)(x)\rangle\Big) \\ &\le e^{-\rho_1 t}\left(-\rho_1 V(x) - \rho_2 + \rho_1 V(x) + \rho_2\right) = 0.\end{aligned} \tag{2.35}$$

Combining this with Beck et al. [4, Lemma 3.1] (with $d = d$, $m = m$, $T = T$, $\mathcal{O} = \mathbb{R}^d$, $\mu = \mu$, $\sigma = \sigma$, $V = \mathbb{V}$, $\tau = t$, $X = X$ for $t \in [0,T]$ in the notation of Beck et al. [4, Lemma 3.1]) demonstrates that for all $t \in [0,T]$ it holds that

$$\mathbb{E}[\mathbb{V}(t, X_t)] \le \mathbb{E}[\mathbb{V}(0, X_0)] = V(\xi) + \tfrac{\rho_2}{\rho_1}. \tag{2.36}$$

Therefore, we obtain that for all $t \in [0,T]$ it holds that

$$\begin{aligned}\mathbb{E}[V(X_t)] &= \mathbb{E}\Big[e^{\rho_1 t}\Big(e^{-\rho_1 t}\big[V(X_t) + \tfrac{\rho_2}{\rho_1}\big]\Big) - \tfrac{\rho_2}{\rho_1}\Big] = e^{\rho_1 t}\,\mathbb{E}[\mathbb{V}(t,X_t)] - \tfrac{\rho_2}{\rho_1} \\ &\le e^{\rho_1 t}\Big[V(\xi) + \tfrac{\rho_2}{\rho_1}\Big] - \tfrac{\rho_2}{\rho_1} = e^{\rho_1 t} V(\xi) + \big(e^{\rho_1 t} - 1\big)\tfrac{\rho_2}{\rho_1}.\end{aligned} \tag{2.37}$$

The fact that for all $a \in \mathbb{R}$ it holds that $e^a - 1 \le a e^a$ hence ensures that for all $t \in [0,T]$ it holds that

$$\mathbb{E}[V(X_t)] \le e^{\rho_1 t} V(\xi) + (\rho_1 t e^{\rho_1 t})\tfrac{\rho_2}{\rho_1} = e^{\rho_1 t}\left(V(\xi) + t\rho_2\right). \tag{2.38}$$

The proof of Lemma 2.6 is thus completed. □

Lemma 2.7. *Let $d, m \in \mathbb{N}$, $T, C_1, C_2 \in [0,\infty)$, $\xi \in \mathbb{R}^d$, let $\langle\cdot,\cdot\rangle : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be the Euclidean scalar product on $\mathbb{R}^d$, let $\|\cdot\| : \mathbb{R}^d \to [0,\infty)$ be the Euclidean norm on $\mathbb{R}^d$, let $|||\cdot||| \colon \mathbb{R}^{d\times m} \to [0,\infty)$ be the Frobenius norm on $\mathbb{R}^{d\times m}$, let $\mu \in C([0,T]\times\mathbb{R}^d, \mathbb{R}^d)$, $\sigma \in C([0,T]\times\mathbb{R}^d, \mathbb{R}^{d\times m})$ satisfy for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that*

$$\max\{\langle x, \mu(t,x)\rangle, |||\sigma(t,x)|||^2\} \le C_1 + C_2\|x\|^2, \tag{2.39}$$

let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_t)_{t\in[0,T]})$ be a filtered probability space which satisfies the usual conditions, let $W \colon [0,T]\times\Omega \to \mathbb{R}^m$ be a standard $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_{t\in[0,T]}))$-Brownian motion, and let $X \colon [0,T]\times\Omega \to \mathbb{R}^d$ be an $(\mathbb{F}_t)_{t\in[0,T]}$/$\mathcal{B}(\mathbb{R}^d)$-adapted stochastic process with continuous sample paths which satisfies that for all $t \in [0,T]$ it holds $\mathbb{P}$-a.s. that

$$X_t = \xi + \int_0^t \mu(r, X_r)dr + \int_0^t \sigma(r, X_r)dW_r. \tag{2.40}$$

Then it holds for all $p \in [0,\infty)$, $t \in [0,T]$ that

$$\begin{aligned}\mathbb{E}[\|X_t\|^p] &\le \Big((1+\|\xi\|^2)^{p/2} + t^{\min\{p/2,1\}}(p+1)|C_1|^{p/2}\Big)\exp\Big(\tfrac{p(p+3)}{2}\big(\mathbb{1}_{(2,\infty)}(p) + C_2\big)t\Big) \\ &\le \max\{T,1\}\Big((1+\|\xi\|^2)^{p/2} + (p+1)|C_1|^{p/2}\Big)\exp\Big(\tfrac{p(p+3)(1+C_2)T}{2}\Big) < \infty.\end{aligned} \tag{2.41}$$

Proof of Lemma 2.7. Throughout this proof let $(\rho_1^{(p)})_{p\in[2,\infty)}, (\rho_2^{(p)})_{p\in[2,\infty)} \subseteq [0,\infty)$ satisfy for all $p \in [2,\infty)$ that

$$\rho_1^{(p)} = \tfrac{p(p+1)}{2}\big(\tfrac{p-2}{p} + C_2\big) \qquad \text{and} \qquad \rho_2^{(p)} = (p+1)|C_1|^{p/2} \tag{2.42}$$

and let $V_p\colon \mathbb{R}^d \to (0,\infty)$, $p \in [2,\infty)$, be the functions which satisfy for all $p \in [2,\infty)$, $x \in \mathbb{R}^d$ that

$$V_p(x) = (1 + \|x\|^2)^{p/2}. \tag{2.43}$$

Observe that Lemma 2.4 and (2.39) assure that for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $p \in [2,\infty)$ it holds that $V_p \in C^\infty(\mathbb{R}^d,(0,\infty))$ and

$$\tfrac{1}{2}\operatorname{Trace}\bigl(\sigma(t,x)[\sigma(t,x)]^*(\operatorname{Hess}V_p)(x)\bigr) + \langle \mu(t,x),(\nabla V_p)(x)\rangle \le \rho_1^{(p)} V_p(x) + \rho_2^{(p)}. \tag{2.44}$$

Lemma 2.6 hence implies that for all $t \in [0,T]$, $p \in [2,\infty)$ it holds that

$$\begin{aligned}
\mathbb{E}[\|X_t\|^p] &\le \mathbb{E}[V_p(X_t)] \le e^{\rho_1^{(p)}t}\left(V_p(\xi) + t\rho_2^{(p)}\right)\\
&= \left((1+\|\xi\|^2)^{p/2} + t(p+1)|C_1|^{p/2}\right)\exp\left(\tfrac{p(p+1)}{2}\bigl(\tfrac{p-2}{p} + C_2\bigr)t\right)\\
&\le \left((1+\|\xi\|^2)^{p/2} + t^{\min\{p/2,1\}}(p+1)|C_1|^{p/2}\right)\exp\left(\tfrac{p(p+3)}{2}\bigl(\mathbb{1}_{(2,\infty)}(p) + C_2\bigr)t\right).
\end{aligned} \tag{2.45}$$

This, Jensen's inequality, and the fact that for all $p \in [0,2]$ it holds that $3^{p/2} \le p+1$ assure that for all $t \in [0,T]$, $p \in [0,2)$ it holds that

$$\begin{aligned}
\mathbb{E}[\|X_t\|^p] &= \mathbb{E}\left[\left(\|X_t\|^2\right)^{p/2}\right] \le \left(\mathbb{E}\left[\|X_t\|^2\right]\right)^{p/2}\\
&\le \left[\left((1+\|\xi\|^2) + t(2+1)|C_1|\right)\exp\left(\tfrac{2(2+1)}{2}C_2 t\right)\right]^{p/2}\\
&\le \left((1+\|\xi\|^2)^{p/2} + t^{p/2}3^{p/2}|C_1|^{p/2}\right)\exp\left(\tfrac{3p}{2}C_2 t\right)\\
&\le \left((1+\|\xi\|^2)^{p/2} + t^{\min\{p/2,1\}}(p+1)|C_1|^{p/2}\right)\exp\left(\tfrac{(p+3)p}{2}\bigl(\mathbb{1}_{(2,\infty)}(p) + C_2\bigr)t\right).
\end{aligned} \tag{2.46}$$

Combining this with (2.45) implies (2.41). The proof of Lemma 2.7 is thus completed. □

2.3 Temporal regularity properties for solutions of SDEs

For the proof of our strong L^2-error estimates for Euler-Maruyama approximations in Subsection 2.4 we need Corollary 2.9 below, which asserts that, under suitable conditions (see Corollary 2.9 below for details), solutions of SDEs have a certain temporal regularity property. To prove Corollary 2.9 we employ (without providing a proof) a well-known temporal regularity property for solutions of SDEs from the literature stated in Lemma 2.8 below (cf., e.g., Da Prato et al. [33, Proposition 3], Cox et al. [26, Corollary 3.8], and Jentzen et al. [67, Proposition 4.1]). Additionally, we offer in Lemma 2.11 below a self-contained proof of an explicit temporal regularity estimate for solutions of SDEs with deterministic initial values which will be used in Subsection 2.8.

Lemma 2.8 (Temporal regularity of solutions of time-homogeneous SDEs)**.** *Let $d,m \in \mathbb{N}$, $T \in (0,\infty)$, let $\|\cdot\| : \mathbb{R}^d \to [0,\infty)$ be the Euclidean norm on $\mathbb{R}^d$, let $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_t)_{t\in[0,T]})$ be a filtered probability space which satisfies the usual conditions, let $W\colon [0,T]\times\Omega \to \mathbb{R}^m$ be a standard $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_t)_{t\in[0,T]})$-Brownian motion, let $\mu\colon \mathbb{R}^d \to \mathbb{R}^d$, $\sigma\colon \mathbb{R}^d \to \mathbb{R}^{d\times m}$ be globally Lipschitz continuous functions, and let $X\colon [0,T]\times\Omega \to \mathbb{R}^d$ be an $(\mathbb{F}_t)_{t\in[0,T]}$/$\mathcal{B}(\mathbb{R}^d)$-adapted stochastic process with continuous sample paths which satisfies that $\mathbb{E}\left[\|X_0\|^2\right] < \infty$ and which satisfies that for all $t \in [0,T]$ it holds $\mathbb{P}$-a.s. that*

$$X_t = X_0 + \int_0^t \mu(X_s)\,ds + \int_0^t \sigma(X_s)\,dW_s. \tag{2.47}$$

Then it holds that

$$\sup\left\{\frac{\left(\mathbb{E}[\|X_t - X_s\|^2]\right)^{1/2}}{|t-s|^{1/2}} \in [0,\infty] \colon t,s \in [0,T], t \neq s\right\} < \infty. \tag{2.48}$$

Corollary 2.9 (Temporal regularity of solutions of time-inhomogeneous SDEs)**.** *Let $d, m \in \mathbb{N}$, $T \in (0,\infty)$, $L \in [0,\infty)$, let $\|\cdot\| : \mathbb{R}^d \to [0,\infty)$ be the Euclidean norm on $\mathbb{R}^d$, let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_t)_{t\in[0,T]})$ be a filtered probability space which satisfies the usual conditions, let $W \colon [0,T] \times \Omega \to \mathbb{R}^m$ be a standard $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_t)_{t\in[0,T]})$-Brownian motion, let $\mu \colon [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma \colon [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ be globally Lipschitz continuous functions, and let $X \colon [0,T] \times \Omega \to \mathbb{R}^d$ be an $(\mathbb{F}_t)_{t\in[0,T]}$/$\mathcal{B}(\mathbb{R}^d)$-adapted stochastic process with continuous sample paths which satisfies that $\mathbb{E}\Big[\|X_0\|^2\Big] < \infty$ and which satisfies that for all $t \in [0,T]$ it holds $\mathbb{P}$-a.s. that*

$$X_t = X_0 + \int_0^t \mu(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dW_s. \tag{2.49}$$

Then it holds that

$$\sup\left\{\frac{\left(\mathbb{E}[\|X_t - X_s\|^2]\right)^{1/2}}{|t-s|^{1/2}} \in [0,\infty] \colon t,s \in [0,T], t \neq s\right\} < \infty. \tag{2.50}$$

Proof of Corollary 2.9. Throughout this proof let $|\!|\!|\cdot|\!|\!| \colon \mathbb{R}^{d+1} \to [0,\infty)$ be the Euclidean norm on $\mathbb{R}^{d+1}$, let $Y \colon [0,T] \times \Omega \to \mathbb{R}^{d+1}$ be the stochastic process which satisfies for all $t \in [0,T]$ that

$$Y_t = \begin{pmatrix} t \\ X_t \end{pmatrix}, \tag{2.51}$$

and let $\tilde{\mu} \colon \mathbb{R}^{d+1} \to \mathbb{R}^{d+1}$ and $\tilde{\sigma} \colon \mathbb{R}^{d+1} \to \mathbb{R}^{(d+1)\times m}$ be the functions which satisfy for all $y = (y_1, y_2, \ldots, y_{d+1}) \in \mathbb{R}^{d+1}$ that

$$\tilde{\mu}(y) = \begin{pmatrix} 1 \\ \mu\big(\min\{\max\{y_1, 0\}, T\}, (y_2, \ldots, y_{d+1})\big) \end{pmatrix} \in \mathbb{R}^{d+1} \qquad \text{and} \tag{2.52}$$

$$\tilde{\sigma}(y) = \begin{pmatrix} 0 \\ \sigma\big(\min\{\max\{y_1, 0\}, T\}, (y_2, \ldots, y_{d+1})\big) \end{pmatrix} \in \mathbb{R}^{(d+1)\times m}. \tag{2.53}$$

Observe that the hypothesis that μ and σ are globally Lipschitz continuous functions and the fact that $\mathbb{R} \ni y \mapsto \min\{\max\{y, 0\}, T\} \in \mathbb{R}$ is a globally Lipschitz continuous function assure that $\tilde{\mu}$ and $\tilde{\sigma}$ are globally Lipschitz continuous functions. Moreover, note that it holds for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$\tilde{\mu}((t,x)) = \begin{pmatrix} 1 \\ \mu(t,x) \end{pmatrix} \qquad \text{and} \qquad \tilde{\sigma}((t,x)) = \begin{pmatrix} 0 \\ \sigma(t,x) \end{pmatrix}. \tag{2.54}$$

This and (2.49) assure that for all $t \in [0,T]$ it holds $\mathbb{P}$-a.s. that

$$\begin{aligned} Y_t = \begin{pmatrix} t \\ X_t \end{pmatrix} &= \begin{pmatrix} \int_0^t 1\, ds \\ X_0 + \int_0^t \mu(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dW_s \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ X_0 \end{pmatrix} + \int_0^t \begin{pmatrix} 1 \\ \mu(s, X_s) \end{pmatrix} ds + \int_0^t \begin{pmatrix} 0 \\ \sigma(s, X_s) \end{pmatrix} dW_s = Y_0 + \int_0^t \tilde{\mu}(Y_s) ds + \int_0^t \tilde{\sigma}(Y_s) dW_s. \end{aligned} \tag{2.55}$$

The fact that $\tilde{\mu}$ and $\tilde{\sigma}$ are globally Lipschitz continuous functions and Lemma 2.8 (with $d = d+1$, $m = m$, $T = T$, $\mu = \tilde{\mu}$, $\sigma = \tilde{\sigma}$, $X = Y$ in the notation of Lemma 2.8) hence prove that

$$\sup_{t,s\in[0,T],t\neq s} \frac{\left(\mathbb{E}[|||Y_t - Y_s|||^2]\right)^{1/2}}{|t-s|^{1/2}} < \infty. \tag{2.56}$$

Hence, we obtain that

$$\begin{aligned} \sup_{t,s\in[0,T],t\neq s} \frac{\left(\mathbb{E}[\|X_t - X_s\|^2]\right)^{1/2}}{|t-s|^{1/2}} &\leq \sup_{t,s\in[0,T],t\neq s} \frac{\left(\mathbb{E}[|t-s|^2 + \|X_t - X_s\|^2]\right)^{1/2}}{|t-s|^{1/2}} \\ &= \sup_{t,s\in[0,T],t\neq s} \frac{\left(\mathbb{E}[|||Y_t - Y_s|||^2]\right)^{1/2}}{|t-s|^{1/2}} < \infty. \end{aligned} \tag{2.57}$$

The proof of Corollary 2.9 is thus completed. □

The following very elementary and well-known result will be helpful in the proof of Lemma 2.11 below and will be repeatedly used throughout this paper.

Lemma 2.10 (A consequence of Hölders inequality)**.** *Let* $(\Omega, \mathcal{F}, \mu)$ *be a measure space and let* $f\colon \Omega \to [0,\infty]$ *be an* $\mathcal{F}/\mathcal{B}([0,\infty])$*-measurable function. Then*

$$\left[\int_\Omega f(\omega)\,\mu(d\omega)\right]^2 \leq \mu(\Omega)\int_\Omega |f(\omega)|^2\,\mu(d\omega). \tag{2.58}$$

Proof of Lemma 2.10. Note that Hölders inequality demonstrates that

$$\begin{aligned} \left[\int_\Omega f(\omega)\,\mu(d\omega)\right]^2 &\leq \left[\left(\int_\Omega 1^2\,\mu(d\omega)\right)^{1/2}\left(\int_\Omega |f(\omega)|^2\,\mu(d\omega)\right)^{1/2}\right]^2 \\ &= \mu(\Omega)\int_\Omega |f(\omega)|^2\,\mu(d\omega). \end{aligned} \tag{2.59}$$

The proof of Lemma 2.10 is thus completed. □

Lemma 2.11 (Explicit temporal regularity for solutions of SDEs with deterministic initial values)**.** *Let* $d, m \in \mathbb{N}$*,* $T \in (0,\infty)$*,* $L \in [0,\infty)$*,* $\xi \in \mathbb{R}^d$*, let* $\|\cdot\|: \mathbb{R}^d \to [0,\infty)$ *be the Euclidean norm on* $\mathbb{R}^d$*, let* $|||\cdot|||\colon \mathbb{R}^{d\times m} \to [0,\infty)$ *be the Frobenius norm on* $\mathbb{R}^{d\times m}$*, let* $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_t)_{t\in[0,T]})$ *be a filtered probability space which satisfies the usual conditions, let* $W\colon [0,T]\times\Omega \to \mathbb{R}^m$ *be a standard* $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_t)_{t\in[0,T]})$*-Brownian motion, 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 functions which satisfy for all* $t, s \in [0,T]$*,* $x, y \in \mathbb{R}^d$ *that*

$$\max\bigl\{\|\mu(t,x) - \mu(s,y)\|, |||\sigma(t,x) - \sigma(s,y)|||\bigr\} \leq L\bigl(|t-s| + \|x-y\|\bigr), \tag{2.60}$$

and let $X\colon [0,T]\times\Omega \to \mathbb{R}^d$ *be an* $(\mathbb{F}_t)_{t\in[0,T]}/\mathcal{B}(\mathbb{R}^d)$*-adapted stochastic process with continuous sample paths which satisfies that for all* $t \in [0,T]$ *it holds* $\mathbb{P}$*-a.s. that*

$$X_t = \xi + \int_0^t \mu(s, X_s)\,ds + \int_0^t \sigma(s, X_s)\,dW_s. \tag{2.61}$$

Then it holds that

$$\begin{aligned} &\sup\left\{\frac{\left(\mathbb{E}[\|X_t - X_s\|^2]\right)^{1/2}}{|t-s|^{1/2}} \in [0,\infty]\colon t, s \in [0,T], t \neq s\right\} \\ &\leq (1 + \|\xi\|)\exp\Bigl(10\bigl(\max\{\|\mu(0,0)\|, |||\sigma(0,0)|||, L, 1\} + LT\bigr)^2(T+1)(L+1)\Bigr) < \infty. \end{aligned} \tag{2.62}$$

Proof of Lemma 2.11. Throughout this proof let $\langle \cdot, \cdot \rangle \colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be the Euclidean scalar product on $\mathbb{R}^d$ and let $C \in (0, \infty)$ be given by

$$C = 2\big(\max\{\|\mu(0,0)\|\,, \|\!|\!|\sigma(0,0)|\!|\!\|, L, 1\} + LT\big)^2. \tag{2.63}$$

Note that (2.60) and the triangle inequality assure that for all $t \in [0, T]$, $x \in \mathbb{R}^d$ it holds that

$$\|\mu(t,x)\| \le \|\mu(0,0)\| + L\big(|t| + \|x\|\big) \le C + L\big(|t| + \|x\|\big) \qquad \text{and} \tag{2.64}$$

$$|\!|\!|\sigma(t,x)|\!|\!| \le |\!|\!|\sigma(0,0)|\!|\!| + L\big(|t| + \|x\|\big) \le C + L\big(|t| + \|x\|\big). \tag{2.65}$$

This assures that for all $t \in [0, T]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
&\langle x, \mu(t,x) \rangle \\
&\le \|x\|\,\|\mu(t,x)\| \le \|x\|\,(\|\mu(0,0)\| + L(t + \|x\|)) \\
&\le \|x\| \max\{\|\mu(0,0)\| + LT, L\}(1 + \|x\|) \le 2\max\{\|\mu(0,0)\| + LT, L\}(1 + \|x\|^2) \\
&\le C(1 + \|x\|^2).
\end{aligned} \tag{2.66}$$

In addition, note that (2.65) implies that for all $t \in [0, T]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
&|\!|\!|\sigma(t,x)|\!|\!|^2 \\
&\le (|\!|\!|\sigma(0,0)|\!|\!| + L(t + \|x\|))^2 \le (\max\{|\!|\!|\sigma(0,0)|\!|\!| + LT, L\})^2(1 + \|x\|)^2 \\
&\le 2(\max\{|\!|\!|\sigma(0,0)|\!|\!| + LT, L\})^2(1 + \|x\|^2) \\
&\le C(1 + \|x\|^2).
\end{aligned} \tag{2.67}$$

Moreover, observe that (2.61), Lemma 2.10, Tonelli's theorem, and Itô's isometry demonstate that for all $t \in [0, T]$, $s \in [t, T]$ it holds that

$$\begin{aligned}
\big(\mathbb{E}[\|X_t - X_s\|^2]\big)^{1/2} &= \left(\mathbb{E}\left[\left\|\int_t^s \mu(r, X_r)\,dr + \int_t^s \sigma(r, X_r)\,dW_r\right\|^2\right]\right)^{1/2} \\
&\le \left(\mathbb{E}\left[\left\|\int_t^s \mu(r, X_r)\,dr\right\|^2\right]\right)^{1/2} + \left(\mathbb{E}\left[\left\|\int_t^s \sigma(r, X_r)\,dW_r\right\|^2\right]\right)^{1/2} \\
&\le |t - s|^{1/2}\left(\int_t^s \mathbb{E}\Big[\|\mu(r, X_r)\|^2\Big]\,dr\right)^{1/2} + \left(\int_t^s \mathbb{E}\Big[\|\sigma(r, X_r)\|^2\Big]\,dr\right)^{1/2}.
\end{aligned} \tag{2.68}$$

The triangle inequality, (2.64), and (2.65) therefore ensure that for all $t \in [0, T]$, $s \in [t, T]$ it holds that

$$\begin{aligned}
&\big(\mathbb{E}[\|X_t - X_s\|^2]\big)^{1/2} \\
&\le (|t - s|^{1/2} + 1)\left(|t - s|^{1/2} C + L\left(\int_t^s \mathbb{E}\Big[(|r| + \|X_r\|)^2\Big]\,dr\right)^{1/2}\right) \\
&\le (|t - s|^{1/2} + 1)\left(|t - s|^{1/2} C + L\left[\left(\int_t^s r^2 dr\right)^{1/2} + \left(\int_t^s \mathbb{E}\Big[\|X_r\|^2\Big]\,dr\right)^{1/2}\right]\right).
\end{aligned} \tag{2.69}$$

Furthermore, note that (2.66), (2.67), (2.61), and Lemma 2.7 (with $d = d$, $m = m$, $T = T$, $C_1 = C$, $C_2 = C$, $\xi = \xi$, $\mu = \mu$, $\sigma = \sigma$, $X = X$ in the notation of Lemma 2.7) assure that for all $t \in [0, T]$ it holds that

$$\mathbb{E}\Big[\|X_t\|^2\Big] \le \Big((1 + \|\xi\|^2) + t3C\Big)\exp(5Ct) \le \Big((1 + \|\xi\|^2) + 3CT\Big)\exp(5CT). \tag{2.70}$$

This, (2.69), the fact that $C \geq 1$, the fact that for all $x \in [0, \infty)$ it holds that $\max\{x, 1+x\} \leq e^x$, and the fact that for all $x, y \in [0, \infty)$ it holds that $\sqrt{x+y} \leq \sqrt{x} + \sqrt{y}$ demonstrate that for all $t \in [0, T]$, $s \in [t, T]$ it holds that

$$\begin{aligned}
&\big(\mathbb{E}[\|X_t - X_s\|^2]\big)^{1/2} \\
&\leq (T^{1/2} + 1) \left(|t-s|^{1/2} C + L|t-s|^{1/2} \left[T + \left[((1 + \|\xi\|^2) + 3CT) \exp(5CT) \right]^{1/2} \right] \right) \\
&\leq |t-s|^{1/2} \big[1 + \|\xi\|^2 \big]^{1/2} \exp(T^{1/2} + C + L) \big(1 + T + [(1 + 3CT) \exp(5CT)]^{1/2} \big) \\
&\leq |t-s|^{1/2} (1 + \|\xi\|) \exp(T^{1/2} + C + L) 2 \exp(4CT) \\
&\leq |t-s|^{1/2} (1 + \|\xi\|) \exp\Big(C(T^{1/2} + 1 + L + 1 + 4T \Big) \\
&\leq |t-s|^{1/2} (1 + \|\xi\|) \exp\big(5C(T+1)(L+1) \big).
\end{aligned} \tag{2.71}$$

This implies (2.62). The proof of Lemma 2.11 is thus completed. □

2.4 Strong error estimates for Euler-Maruyama approximations

Our proof of the flow-type property of solutions of SDEs in Subsection 2.8 below makes use of Euler-Maruyama approximations of solutions. For that reason we present in this subsection explicit strong L^2-error estimates for Euler-Maruyama approximations in Proposition 2.12 and Corollary 2.13 below. The results in this subsection are essentially well-known (cf., e.g., Kloeden & Platen [71, Chapter 10] and Milstein [81]).

Proposition 2.12 (Strong convergence of the Euler-Maruyama method)**.** *Let $d, m, N \in \mathbb{N}$, $T \in (0, \infty)$, $L \in [0, \infty)$, let $\|\cdot\| : \mathbb{R}^d \to [0, \infty)$ be the Euclidean norm on $\mathbb{R}^d$, let $|\!|\!|\cdot|\!|\!| \colon \mathbb{R}^{d \times m} \to [0, \infty)$ be the Frobenius norm on $\mathbb{R}^{d \times m}$, let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_t)_{t \in [0,T]})$ be a filtered probability space which satisfies the usual conditions, let $W \colon [0, T] \times \Omega \to \mathbb{R}^m$ be a standard $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_t)_{t \in [0,T]})$-Brownian motion, let $\zeta \colon \Omega \to \mathbb{R}^d$ be an $\mathbb{F}_0$/$\mathcal{B}(\mathbb{R}^d)$-measurable function which satisfies that $\mathbb{E}\Big[\|\zeta\|^2\Big] < \infty$, 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 functions which satisfy for all $t, s \in [0, T]$, $x, y \in \mathbb{R}^d$ that*

$$\max\big\{ \|\mu(t,x) - \mu(s,y)\| , |\!|\!|\sigma(t,x) - \sigma(s,y)|\!|\!| \big\} \leq L\big(|t-s| + \|x-y\| \big), \tag{2.72}$$

let $X \colon [0, T] \times \Omega \to \mathbb{R}^d$ be an $(\mathbb{F}_t)_{t \in [0,T]}$/$\mathcal{B}(\mathbb{R}^d)$-adapted stochastic process with continuous sample paths which satisfies that $\mathbb{E}\Big[\|X_0\|^2\Big] < \infty$ and which satisfies that for all $t \in [0, T]$ it holds $\mathbb{P}$-a.s. that

$$X_t = X_0 + \int_0^t \mu(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dW_s, \tag{2.73}$$

let $t_0, t_1, \ldots, t_N \in [0, T]$ satisfy that

$$0 = t_0 \leq t_1 \leq t_2 \leq \ldots \leq t_N = T, \tag{2.74}$$

and let $\mathcal{X} \colon \{0, 1, \ldots, N\} \times \Omega \to \mathbb{R}^d$ be the stochastic process which satisfies for all $n \in \{1, 2, \ldots, N\}$ that

$$\mathcal{X}_0 = \zeta \quad \text{and} \quad \mathcal{X}_n = \mathcal{X}_{n-1} + \mu(t_{n-1}, \mathcal{X}_{n-1})(t_n - t_{n-1}) + \sigma(t_{n-1}, \mathcal{X}_{n-1})(W_{t_n} - W_{t_{n-1}}). \tag{2.75}$$

Then it holds that

$$\begin{aligned}
\Big(\mathbb{E}\Big[\|X_T - \mathcal{X}_N\|^2\Big]\Big)^{1/2} &\leq \left[\Big(\mathbb{E}\Big[\|X_0 - \zeta\|^2\Big] \Big)^{1/2} + \max_{k \in \{1,2,\ldots,N\}} |t_k - t_{k-1}|^{1/2} \right] \\
&\quad \cdot \exp\Big((1+L)^2 (1 + \sqrt{T})^4 \Big) \left(1 + \sup_{s,r \in [0,T], s \neq r} \tfrac{(\mathbb{E}[\|X_s - X_r\|^2])^{1/2}}{|s-r|^{1/2}} \right) < \infty.
\end{aligned} \tag{2.76}$$

Proof of Proposition 2.12. Throughout this proof assume w.l.o.g. that $t_0 < t_1 < t_2 < \ldots < t_N$, let $(h_n)_{n\in\{1,2,\ldots,N\}} \subseteq (0,T]$, $H \in (0,T]$, $K \in [0,\infty]$ satisfy for all $n \in \{1,2,\ldots,N\}$ that $h_n = |t_n - t_{n-1}|$, $H = \max_{k\in\{1,2,\ldots,N\}} |t_k - t_{k-1}|$, and

$$K = \sup_{s,r\in[0,T],s\neq r} \frac{\big(\mathbb{E}[\|X_s - X_r\|^2]\big)^{1/2}}{|s-r|^{1/2}}, \tag{2.77}$$

let $\mathfrak{t}\colon [0,T] \to \{t_0, t_1, t_2, \ldots, t_N\}$ be the function which satisfies for all $s \in [0,T]$ that

$$\mathfrak{t}(s) = \max\left(\{t_0, t_1, \ldots, t_N\} \cap [0,s]\right), \tag{2.78}$$

and let $\mathfrak{n}\colon [0,T] \to \{0,1,2,\ldots,N\}$ be the function which satisfies for all $s \in [0,T]$ that

$$\mathfrak{n}(s) = \max\left(\{n \in \{0,1,2,\ldots,N\}\colon t_n \leq s\}\right). \tag{2.79}$$

Note that the hypothesis that $\mathbb{E}\Big[\|X_0\|^2\Big] < \infty$, the fact that μ and σ are globally Lipschitz continuous functions, (2.73), and Corollary 2.9 imply that $K < \infty$. Next observe that (2.75) and induction assure that for all $n \in \{0,1,2,\ldots,N\}$ it holds $\mathbb{P}$-a.s. that

$$\begin{aligned} \mathcal{X}_n &= \mathcal{X}_0 + \left[\sum_{k=1}^{n} \mu(t_{k-1}, \mathcal{X}_{k-1})(t_k - t_{k-1})\right] + \left[\sum_{k=1}^{n} \sigma(t_{k-1}, \mathcal{X}_{k-1})(W_{t_k} - W_{t_{k-1}})\right] \\ &= \zeta + \int_0^{t_n} \mu\big(\mathfrak{t}(s), \mathcal{X}_{\mathfrak{n}(s)}\big)\, ds + \int_0^{t_n} \sigma\big(\mathfrak{t}(s), \mathcal{X}_{\mathfrak{n}(s)}\big)\, dW_s. \end{aligned} \tag{2.80}$$

This and (2.73) imply that for all $n \in \{0,1,2,\ldots,N\}$ it holds $\mathbb{P}$-a.s. that

$$\begin{aligned} X_{t_n} - \mathcal{X}_n &= X_0 - \zeta + \int_0^{t_n} \mu\big(s, X_s\big) - \mu\big(\mathfrak{t}(s), \mathcal{X}_{\mathfrak{n}(s)}\big)\, ds + \int_0^{t_n} \sigma\big(s, X_s\big) - \sigma\big(\mathfrak{t}(s), \mathcal{X}_{\mathfrak{n}(s)}\big)\, dW_s \\ &= X_0 - \zeta + \int_0^{t_n} \mu\big(s, X_s\big) - \mu\big(\mathfrak{t}(s), X_{\mathfrak{t}(s)}\big)\, ds + \int_0^{t_n} \sigma\big(s, X_s\big) - \sigma\big(\mathfrak{t}(s), X_{\mathfrak{t}(s)}\big)\, dW_s \\ &\quad + \int_0^{t_n} \mu\big(\mathfrak{t}(s), X_{\mathfrak{t}(s)}\big) - \mu\big(\mathfrak{t}(s), \mathcal{X}_{\mathfrak{n}(s)}\big)\, ds + \int_0^{t_n} \sigma\big(\mathfrak{t}(s), X_{\mathfrak{t}(s)}\big) - \sigma\big(\mathfrak{t}(s), \mathcal{X}_{\mathfrak{n}(s)}\big)\, dW_s. \end{aligned} \tag{2.81}$$

The triangle inequality hence proves that for all $n \in \{0,1,2,\ldots,N\}$ it holds that

$$\begin{aligned} &\Big(\mathbb{E}\Big[\|X_{t_n} - \mathcal{X}_n\|^2\Big]\Big)^{1/2} \\ &\leq \Big(\mathbb{E}\Big[\|X_0 - \zeta\|^2\Big]\Big)^{1/2} + \left(\mathbb{E}\left[\left\|\int_0^{t_n} \mu\big(s, X_s\big) - \mu\big(\mathfrak{t}(s), X_{\mathfrak{t}(s)}\big)\, ds\right\|^2\right]\right)^{1/2} \\ &+ \left(\mathbb{E}\left[\left\|\int_0^{t_n} \sigma\big(s, X_s\big) - \sigma\big(\mathfrak{t}(s), X_{\mathfrak{t}(s)}\big)\, dW_s\right\|^2\right]\right)^{1/2} \\ &+ \left(\mathbb{E}\left[\left\|\int_0^{t_n} \mu\big(\mathfrak{t}(s), X_{\mathfrak{t}(s)}\big) - \mu\big(\mathfrak{t}(s), \mathcal{X}_{\mathfrak{n}(s)}\big)\, ds\right\|^2\right]\right)^{1/2} \\ &+ \left(\mathbb{E}\left[\left\|\int_0^{t_n} \sigma\big(\mathfrak{t}(s), X_{\mathfrak{t}(s)}\big) - \sigma\big(\mathfrak{t}(s), \mathcal{X}_{\mathfrak{n}(s)}\big)\, dW_s\right\|^2\right]\right)^{1/2}. \end{aligned} \tag{2.82}$$

Lemma 2.10, Tonelli's Theorem, and Itô's isometry therefore imply that for all $n \in$

$\{0, 1, 2, \ldots, N\}$ it holds that

$$\begin{aligned}
&\Big(\mathbb{E}\Big[\|X_{t_n} - \mathcal{X}_n\|^2\Big]\Big)^{1/2} \\
&\le \Big(\mathbb{E}\Big[\|X_0 - \zeta\|^2\Big]\Big)^{1/2} + \bigg(T \int_0^{t_n} \mathbb{E}\Big[\big\|\mu\big(s, X_s\big) - \mu\big(\mathfrak{t}(s), X_{\mathfrak{t}(s)}\big)\big\|^2\Big]\, ds\bigg)^{1/2} \\
&+ \bigg(\int_0^{t_n} \mathbb{E}\Big[\big|\!\big|\!\big|\sigma\big(s, X_s\big) - \sigma\big(\mathfrak{t}(s), X_{\mathfrak{t}(s)}\big)\big|\!\big|\!\big|^2\Big]\, ds\bigg)^{1/2} \\
&+ \bigg(T \int_0^{t_n} \mathbb{E}\Big[\big\|\mu\big(\mathfrak{t}(s), X_{\mathfrak{t}(s)}\big) - \mu\big(\mathfrak{t}(s), \mathcal{X}_{\mathfrak{n}(s)}\big)\big\|^2\Big]\, ds\bigg)^{1/2} \\
&+ \bigg(\int_0^{t_n} \mathbb{E}\Big[\big|\!\big|\!\big|\sigma\big(\mathfrak{t}(s), X_{\mathfrak{t}(s)}\big) - \sigma\big(\mathfrak{t}(s), \mathcal{X}_{\mathfrak{n}(s)}\big)\big|\!\big|\!\big|^2\Big]\, ds\bigg)^{1/2}.
\end{aligned} \tag{2.83}$$

This and (2.72) show that for all $n \in \{0, 1, 2, \ldots, N\}$ it holds that

$$\begin{aligned}
&\Big(\mathbb{E}\Big[\|X_{t_n} - \mathcal{X}_n\|^2\Big]\Big)^{1/2} \\
&\le \Big(\mathbb{E}\Big[\|X_0 - \zeta\|^2\Big]\Big)^{1/2} + L\sqrt{T}\bigg(\int_0^{T} \mathbb{E}\Big[\big(|s - \mathfrak{t}(s)| + \big\|X_s - X_{\mathfrak{t}(s)}\big\|\big)^2\Big]\, ds\bigg)^{1/2} \\
&+ L\bigg(\int_0^{T} \mathbb{E}\Big[\big(|s - \mathfrak{t}(s)| + \big\|X_s - X_{\mathfrak{t}(s)}\big\|\big)^2\Big]\, ds\bigg)^{1/2} \\
&+ L\sqrt{T}\bigg(\int_0^{t_n} \mathbb{E}\Big[\big\|X_{\mathfrak{t}(s)} - \mathcal{X}_{\mathfrak{n}(s)}\big\|^2\Big]\, ds\bigg)^{1/2} \\
&+ L\bigg(\int_0^{t_n} \mathbb{E}\Big[\big\|X_{\mathfrak{t}(s)} - \mathcal{X}_{\mathfrak{n}(s)}\big\|^2\Big]\, ds\bigg)^{1/2}.
\end{aligned} \tag{2.84}$$

This, the triangle inequality, and the fact that for all $s \in [0, T]$ it holds that $|s - \mathfrak{t}(s)| \le H$ imply that for all $n \in \{0, 1, 2, \ldots, N\}$ it holds that

$$\begin{aligned}
&\Big(\mathbb{E}\Big[\|X_{t_n} - \mathcal{X}_n\|^2\Big]\Big)^{1/2} \\
&\le \Big(\mathbb{E}\Big[\|X_0 - \zeta\|^2\Big]\Big)^{1/2} + L(1 + \sqrt{T})\left[\sqrt{T}H + \bigg(\int_0^{T} \mathbb{E}\Big[\big\|X_s - X_{\mathfrak{t}(s)}\big\|^2\Big]\, ds\bigg)^{1/2}\right] \\
&+ L(1 + \sqrt{T})\left(\sum_{k=1}^{n} h_k\, \mathbb{E}\Big[\big\|X_{t_{k-1}} - \mathcal{X}_{k-1}\big\|^2\Big]\right)^{1/2}.
\end{aligned} \tag{2.85}$$

The fact that for all $x, y \in [0, \infty)$ it holds that $(x + y)^2 \le 2x^2 + 2y^2$ hence proves that for all $n \in \{0, 1, 2, \ldots, N\}$ it holds that

$$\begin{aligned}
&\mathbb{E}\Big[\|X_{t_n} - \mathcal{X}_n\|^2\Big] \\
&\le 2\left(\Big(\mathbb{E}\Big[\|X_0 - \zeta\|^2\Big]\Big)^{1/2} + L(1 + \sqrt{T})\left[\sqrt{T}H + \bigg(\int_0^{T} \mathbb{E}\Big[\big\|X_s - X_{\mathfrak{t}(s)}\big\|^2\Big]\, ds\bigg)^{1/2}\right]\right)^2 \\
&+ 2L^2(1 + \sqrt{T})^2\left(\sum_{k=1}^{n} h_k\, \mathbb{E}\Big[\big\|X_{t_{k-1}} - \mathcal{X}_{k-1}\big\|^2\Big]\right).
\end{aligned} \tag{2.86}$$

The discrete Gronwall-type inequality in Lemma 2.2 (with $N = N$, $\alpha = 2\big(\big(\mathbb{E}\big[\|X_0 - \zeta\|^2\big]\big)^{1/2} + L(1 + \sqrt{T})[\sqrt{T}H + (\int_0^T \mathbb{E}\big[\|X_s - X_{\mathfrak{t}(s)}\|^2\big]\, ds)^{1/2}]\big)^2$, $(\beta_n)_{n \in \{0,1,2,\ldots,N-1\}} =$

$(2L^2(1+\sqrt{T})^2 h_{n+1})_{n\in\{0,1,2,\dots,N-1\}}$, $(\epsilon_n)_{n\in\{0,1,2,\dots,N\}} = (\mathbb{E}\Big[\|X_{t_n} - \mathcal{X}_n\|^2\Big])_{n\in\{0,1,2,\dots,N\}}$ in the notation of Lemma 2.2) and the fact that $\sum_{k=1}^N h_k = T$ therefore show that

$$\begin{aligned}
&\mathbb{E}\Big[\|X_{t_N} - \mathcal{X}_N\|^2\Big] \\
&\leq 2\left(\left(\mathbb{E}\Big[\|X_0-\zeta\|^2\Big]\right)^{1/2} + L(1+\sqrt{T})\left[\sqrt{T}H + \left(\int_0^T \mathbb{E}\Big[\|X_s - X_{\mathfrak{t}(s)}\|^2\Big]\,ds\right)^{1/2}\right]\right)^2 \\
&\quad\cdot \exp\left(2L^2(1+\sqrt{T})^2 T\right).
\end{aligned} \tag{2.87}$$

This and the fact that for all $s\in[0,T]$ it holds that $|s-\mathfrak{t}(s)|\leq H$ imply that

$$\begin{aligned}
&\left(\mathbb{E}\Big[\|X_T - \mathcal{X}_N\|^2\Big]\right)^{1/2} \\
&\leq \sqrt{2}\left(\left(\mathbb{E}\Big[\|X_0-\zeta\|^2\Big]\right)^{1/2} + L(1+\sqrt{T})\left[\sqrt{T}H + \left(TK^2H\right)^{1/2}\right]\right)\exp\left(L^2(1+\sqrt{T})^2T\right).
\end{aligned} \tag{2.88}$$

The fact that $H\leq \sqrt{T}\sqrt{H}$ hence assures that

$$\begin{aligned}
&\left(\mathbb{E}\Big[\|X_T - \mathcal{X}_N\|^2\Big]\right)^{1/2} \\
&\leq \sqrt{2}\left(\left(\mathbb{E}\Big[\|X_0-\zeta\|^2\Big]\right)^{1/2} + L(1+\sqrt{T})\sqrt{T}(\sqrt{T}+1)\sqrt{H}(1+K)\right)\exp\Big(L^2(1+\sqrt{T})^2T\Big) \\
&\leq \sqrt{2}\left(\left(\mathbb{E}\Big[\|X_0-\zeta\|^2\Big]\right)^{1/2} + \sqrt{H}\exp\Big(L(1+\sqrt{T})^2\sqrt{T}\Big)\right)\exp\Big(L^2(1+\sqrt{T})^2T\Big)(1+K) \\
&\leq \left(\left(\mathbb{E}\Big[\|X_0-\zeta\|^2\Big]\right)^{1/2} + \sqrt{H}\right)\exp\Big((1+L)^2(1+\sqrt{T})^4\Big)(1+K).
\end{aligned} \tag{2.89}$$

This implies (2.76). The proof of Proposition 2.12 is thus completed. □

Corollary 2.13. *Let $d,m,N\in\mathbb{N}$, $T\in(0,\infty)$, $t\in[0,T]$, $s\in[t,T]$, $L\in[0,\infty)$, let $\|\cdot\|:\mathbb{R}^d\to[0,\infty)$ be the Euclidean norm on $\mathbb{R}^d$, let $|||\cdot|||:\mathbb{R}^{d\times m}\to[0,\infty)$ be the Frobenius norm on $\mathbb{R}^{d\times m}$, let $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_r)_{r\in[0,T]})$ be a filtered probability space which satisfies the usual conditions, let $W\colon[0,T]\times\Omega\to\mathbb{R}^m$ be a standard $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_r)_{r\in[0,T]})$-Brownian motion, let $\zeta\colon\Omega\to\mathbb{R}^d$ be an $\mathbb{F}_t$/$\mathcal{B}(\mathbb{R}^d)$-measurable function with $\mathbb{E}\Big[\|\zeta\|^2\Big]<\infty$, 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 functions which satisfy for all $r,h\in[0,T]$, $x,y\in\mathbb{R}^d$ that*

$$\max\big\{\|\mu(r,x)-\mu(h,y)\|, |||\sigma(r,x)-\sigma(h,y)|||\big\}\leq L\big(|r-h|+\|x-y\|\big), \tag{2.90}$$

let $X\colon[t,s]\times\Omega\to\mathbb{R}^d$ be an $(\mathbb{F}_r)_{r\in[t,s]}$/$\mathcal{B}(\mathbb{R}^d)$-adapted stochastic process with continuous sample paths which satisfies that $\mathbb{E}\Big[\|X_t\|^2\Big]<\infty$ and which satisfies that for all $r\in[t,s]$ it holds $\mathbb{P}$-a.s. that

$$X_r = X_t + \int_t^r \mu(h,X_h)\,dh + \int_t^r \sigma(h,X_h)\,dW_h, \tag{2.91}$$

let $r_0,r_1,\dots,r_N\in[0,T]$ satisfy that

$$t = r_0\leq r_1\leq r_2\leq\ldots\leq r_N = s, \tag{2.92}$$

and let $\mathcal{X}\colon \{0,1,\dots,N\}\times\Omega\to\mathbb{R}^d$ *be the stochastic process which satisfies for all* $n\in\{1,2,\dots,N\}$ *that*

$$\mathcal{X}_0=\zeta \qquad \textit{and} \qquad \mathcal{X}_n=\mathcal{X}_{n-1}+\mu(r_{n-1},\mathcal{X}_{n-1})(r_n-r_{n-1})+\sigma(r_{n-1},\mathcal{X}_{n-1})(W_{r_n}-W_{r_{n-1}}). \tag{2.93}$$

Then it holds that

$$\begin{aligned}&\Big(\mathbb{E}\Big[\|X_s-\mathcal{X}_N\|^2\Big]\Big)^{1/2}\\ &\leq\bigg[\Big(\mathbb{E}\Big[\|X_t-\zeta\|^2\Big]\Big)^{1/2}+\max_{k\in\{1,2,\dots,N\}}|r_k-r_{k-1}|^{1/2}\bigg]\exp\Big((1+L)^2(1+\sqrt{T})^4\Big)\\ &\quad\cdot\bigg(1+\sup\Big(\Big\{\tfrac{(\mathbb{E}[\|X_r-X_h\|^2])^{1/2}}{|r-h|^{1/2}}\in[0,\infty]\colon (r,h\in[t,s],r\neq h)\Big\}\cup\{0\}\Big)\bigg)<\infty.\end{aligned} \tag{2.94}$$

Proof of Corollary 2.13. Throughout this proof assume w.l.o.g. that $s>t$. Observe that Proposition 2.12 (with $d=d$, $m=m$, $N=N$, $T=s-t$, $L=L$, $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_r)_{r\in[0,T]})=(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_{t+r})_{r\in[0,s-t]})$, $(W_r)_{r\in[0,T]}=(W_{t+r}-W_t)_{r\in[0,s-t]}$, $\zeta=\zeta$, $(\mu(r,x))_{r\in[0,T],x\in\mathbb{R}^d}=(\mu(t+r,x))_{r\in[0,s-t],x\in\mathbb{R}^d}$, $(\sigma(r,x))_{r\in[0,T],x\in\mathbb{R}^d}=(\sigma(t+r,x))_{r\in[0,s-t],x\in\mathbb{R}^d}$, $(X_r)_{r\in[0,T]}=(X_{t+r})_{r\in[0,s-t]}$, $(t_n)_{n\in\{0,1,\dots,N\}}=(r_n-t)_{n\in\{0,1,\dots,N\}}$, $(\mathcal{X}_n)_{n\in\{0,1,\dots,N\}}=(\mathcal{X}_n)_{n\in\{0,1,\dots,N\}}$ in the notation of Proposition 2.12) establishes that

$$\begin{aligned}&\Big(\mathbb{E}\Big[\big\|X_{t+(s-t)}-\mathcal{X}_N\big\|^2\Big]\Big)^{1/2}\\ &\leq\bigg[\Big(\mathbb{E}\Big[\|X_{t+0}-\zeta\|^2\Big]\Big)^{1/2}+\max_{k\in\{1,2,\dots,N\}}|r_k-t-(r_{k-1}-t)|^{1/2}\bigg]\\ &\quad\cdot\exp\Big((1+L)^2(1+\sqrt{|s-t|})^4\Big)\left(1+\sup_{r,h\in[0,s-t],r\neq h}\frac{(\mathbb{E}[\|X_{t+r}-X_{t+h}\|^2])^{1/2}}{|r-h|^{1/2}}\right)<\infty.\end{aligned} \tag{2.95}$$

This implies (2.94). The proof of Corollary 2.13 is thus completed. □

2.5 On identically distributed random variables

The next elementary and well-known result, Lemma 2.14 below, provides a sufficient condition for two random variables to have the same distribution.

Lemma 2.14. *Let* $(\Omega,\mathcal{F},\mathbb{P})$ *be a probability space, let* (E,d) *be a metric space, let* $X,Y\colon\Omega\to E$ *be random variables which satisfy that for all globally bounded and Lipschitz continuous functions* $g\colon E\to\mathbb{R}$ *it holds that*

$$\mathbb{E}[g(X)]=\mathbb{E}[g(Y)]\,. \tag{2.96}$$

Then it holds that X *and* Y *are identically distributed random variables.*

Proof of Lemma 2.14. Throughout this proof for every $n\in\mathbb{N}$ let $h_n\colon[0,\infty)\to[0,1]$ be the function which satisfies for all $r\in[0,\infty)$ that

$$h_n(r)=\max\{1-nr,0\}, \tag{2.97}$$

for every closed and non-empty set $A\subseteq E$ let $D_A\colon E\to[0,\infty)$ be the function which satisfies for all $e\in E$ that

$$D_A(e)=\inf_{a\in A}d(e,a), \tag{2.98}$$

and for every $n\in\mathbb{N}$ and every closed and non-empty set $A\subseteq E$ let $f_{A,n}\colon E\to[0,1]$ be the function which satisfies for all $e\in E$ that

$$f_{A,n}(e)=h_n(D_A(e)). \tag{2.99}$$

Note that the triangle inequality assures that for all closed and non-empty sets $A \subseteq E$ and all $e_1, e_2 \in E$, $a \in A$, $\varepsilon \in (0,\infty)$ with $D_A(e_1) \geq D_A(e_2)$ and $d(e_2, a) \leq D_A(e_2) + \varepsilon$ it holds that

$$\begin{aligned} |D_A(e_1) - D_A(e_2)| &= D_A(e_1) - D_A(e_2) \leq d(e_1, a) - d(e_2, a) + \varepsilon \\ &\leq d(e_1, e_2) + d(e_2, a) - d(e_2, a) + \varepsilon = d(e_1, e_2) + \varepsilon. \end{aligned} \tag{2.100}$$

The fact that for all closed and non-empty sets $A \subseteq E$ and all $e \in E$, $\varepsilon \in (0,\infty)$ there exists $a \in A$ such that $d(e, a) \leq D_A(e) + \varepsilon$ hence assures that for all closed and non-empty sets $A \subseteq E$ and all $e_1, e_2 \in E$ it holds that

$$|D_A(e_1) - D_A(e_2)| \leq d(e_1, e_2). \tag{2.101}$$

Moreover note that for all $n \in \mathbb{N}$, $r_1, r_2 \in [0,\infty)$ with $r_1 \leq r_2$ it holds that

$$\begin{aligned} |h_n(r_1) - h_n(r_2)| &= |h_n(r_2) - h_n(r_1)| = h_n(r_1) - h_n(r_2) \\ &= \max\{1 - nr_1, 0\} - \max\{1 - nr_2, 0\} \\ &= \max\big\{1 - nr_1 - \max\{1 - nr_2, 0\}, -\max\{1 - nr_2, 0\}\big\} \\ &\leq \max\{1 - nr_1 - (1 - nr_2), 0\} = \max\{n(r_2 - r_1), 0\} = n|r_1 - r_2|. \end{aligned} \tag{2.102}$$

Combining this with (2.101) establishes that for all closed and non-empty sets $A \subseteq E$ and all $n \in \mathbb{N}$, $e_1, e_2 \in E$ it holds that

$$|f_{A,n}(e_1) - f_{A,n}(e_2)| = |h_n(D_A(e_1)) - h_n(D_A(e_2))| \leq n|D_A(e_1) - D_A(e_2)| \leq nd(e_1, e_2). \tag{2.103}$$

This demonstrates that for every closed and non-empty set $A \subseteq E$ and every $n \in \mathbb{N}$ it holds that $f_{A,n} \colon E \to [0,1]$ is a globally bounded and Lipschitz continuous function. Next observe that the fact that for all closed and non-empty sets $A \subseteq E$ and all $e \in A$ it holds that $D_A(e) = 0$ assures that for all closed and non-empty sets $A \subseteq E$ and all $n \in \mathbb{N}$, $e \in A$ it holds that

$$f_{A,n}(e) = h_n(D_A(e)) = h_n(0) = 1. \tag{2.104}$$

Moreover, note the fact that for all closed and non-empty sets $A \subseteq E$ and all $e \in E \setminus A$ there exists $n \in \mathbb{N}$ such that $D_A(e) > \frac{1}{n}$ and the fact that for all $n \in \mathbb{N}$ it holds that h_n is a non-increasing function assure that for all closed and non-empty sets $A \subseteq E$ and all $e \in E \setminus A$ there exist $n \in \mathbb{N}$ such that for all $m \in \{n, n+1, \ldots\}$ it holds that

$$f_{A,m}(e) = h_m(D_A(e)) \leq h_m(\tfrac{1}{n}) = \max\{1 - \tfrac{m}{n}, 0\} = 0. \tag{2.105}$$

Combining this and (2.104) establishes that for all closed and non-empty sets $A \subseteq E$ and all $e \in E$ it holds that

$$\lim_{n\to\infty} f_{A,n}(e) = \mathbb{1}_A(e). \tag{2.106}$$

The theorem of dominated convergence, the fact that for all closed and non-empty sets $A \subseteq E$ and all $n \in \mathbb{N}$ it holds that $f_{A,n} \colon E \to [0,1]$ is a globally bounded and Lipschitz continuous function, and (2.96) therefore imply that for all closed and non-empty sets $A \subseteq E$ it holds that

$$\mathbb{P}(X \in A) = \mathbb{E}[\mathbb{1}_A(X)] = \lim_{n\to\infty} \mathbb{E}[f_{A,n}(X)] = \lim_{n\to\infty} \mathbb{E}[f_{A,n}(Y)] = \mathbb{E}[\mathbb{1}_A(Y)] = \mathbb{P}(Y \in A). \tag{2.107}$$

The fact that $\mathcal{B}(E) = \mathfrak{S}(\{A \subseteq E \colon A \text{ is closed}\})$, the fact that $\{A \subseteq E \colon A \text{ is closed}\}$ is closed under intersections, and the uniqueness theorem for measures (see, e.g., Klenke [70, Lemma 1.42]) hence assure that for all $B \in \mathcal{B}(E)$ it holds that

$$\mathbb{P}(X \in B) = \mathbb{P}(Y \in B). \tag{2.108}$$

The proof of Lemma 2.14 is thus completed. □

2.6 On random evaluations of random fields

This subsection collects elementary and well-known results about random variables originating from evaluations of random fields at random indices.

Lemma 2.15. *Let $(\Omega,\mathcal{F})$, $(S,\mathcal{S})$, $(E,\mathcal{E})$ be measurable spaces, let $U=(U(s))_{s\in S}=(U(s,\omega))_{s\in S,\omega\in\Omega}\colon S\times\Omega\to E$ be an $(\mathcal{S}\otimes\mathcal{F})/\mathcal{E}$-measurable function, and let $X\colon\Omega\to S$ be an $\mathcal{F}/\mathcal{S}$-measurable function. Then it holds that the function $U(X)=(U(X(\omega),\omega))_{\omega\in\Omega}\colon\Omega\to E$ is $\mathcal{F}/\mathcal{E}$-measurable.*

Proof of Lemma 2.15. Throughout this proof let $\mathcal{X}\colon\Omega\to S\times\Omega$ be the function which satisfies for all $\omega\in\Omega$ that

$$\mathcal{X}(\omega)=(X(\omega),\omega). \tag{2.109}$$

Observe that the hypothesis that $X\colon\Omega\to S$ is an $\mathcal{F}/\mathcal{S}$-measurable function assures that $\mathcal{X}\colon\Omega\to S\times\Omega$ is an $\mathcal{F}/(\mathcal{S}\otimes\mathcal{F})$-measurable function. Combining this with the fact that $U\colon S\times\Omega\to E$ is an $(\mathcal{S}\otimes\mathcal{F})/\mathcal{E}$-measurable function demonstrates that

$$U(X)=U\circ\mathcal{X} \tag{2.110}$$

is an $\mathcal{F}/\mathcal{E}$-measurable function. The proof of Lemma 2.15 is thus completed. □

A proof for the next two elementary and well-known results (see Lemma 2.16 and Lemma 2.17 below) can, e.g., be found in [63, Lemma 2.3 and Lemma 2.4].

Lemma 2.16. *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let (S,δ) be a separable metric space, let $U=(U(s))_{s\in S}\colon S\times\Omega\to[0,\infty)$ be a continuous random field, let $X\colon\Omega\to S$ be a random variable, and assume that U and X are independent. Then it holds that*

$$\mathbb{E}[U(X)]=\int_S \mathbb{E}[U(s)]\,(X(\mathbb{P})_{\mathcal{B}(S)})(ds). \tag{2.111}$$

Lemma 2.17. *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let (S,δ) be a separable metric space, let $U=(U(s))_{s\in S}\colon S\times\Omega\to\mathbb{R}$ be a continuous random field, let $X\colon\Omega\to S$ be a random variable, assume that U and X are independent, and assume that $\int_S\mathbb{E}[|U(s)|]\,(X(\mathbb{P})_{\mathcal{B}(S)})(ds)<\infty$. Then it holds that $(X(\mathbb{P})_{\mathcal{B}(S)})(\{s\in S\colon\mathbb{E}[|U(s)|]=\infty\})=0$, $\mathbb{E}[|U(X)|]<\infty$, and*

$$\mathbb{E}[U(X)]=\int_S \mathbb{E}[U(s)]\,(X(\mathbb{P})_{\mathcal{B}(S)})(ds). \tag{2.112}$$

2.7 Brownian motions and right-continuous filtrations

The next result, Lemma 2.18 below, states that a Brownian motion with respect to a filtration is also a Brownian motion with respect to the smallest right-continuous filtration containing the original filtration (cf. (2.113)). Lemma 2.18 and its proof are very similar to Prévôt & Röckner [90, Proposition 2.1.13].

Lemma 2.18. *Let $m\in\mathbb{N}$, $T\in(0,\infty)$, let $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_t)_{t\in[0,T]})$ be a filtered probability space, let $W\colon[0,T]\times\Omega\to\mathbb{R}^m$ be a standard $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_t)_{t\in[0,T]})$-Brownian motion, and let $\mathbb{H}_t\subseteq\mathcal{F}$, $t\in[0,T]$, satisfy for all $t\in[0,T]$ that*

$$\mathbb{H}_t=\begin{cases}\cap_{s\in(t,T]}\mathbb{F}_s & : t<T\\ \mathbb{F}_T & : t=T.\end{cases} \tag{2.113}$$

Then it holds that W is a standard $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{H}_t)_{t\in[0,T]})$-Brownian motion.

Proof of Lemma 2.18. Throughout this proof let $\|\cdot\| : \mathbb{R}^d \to [0,\infty)$ be the Euclidean norm on $\mathbb{R}^d$, for every $n \in \mathbb{N}$ let $h_n \colon [0,\infty) \to [0,1]$ be the function which satisfies for all $r \in [0,\infty)$ that

$$h_n(r) = \max\{1 - nr, 0\}, \tag{2.114}$$

for every closed and non-empty set $A \subseteq \mathbb{R}^d$ let $D_A \colon \mathbb{R}^d \to [0,\infty)$ be the function which satisfies for all $x \in \mathbb{R}^d$ that

$$D_A(x) = \inf_{a \in A} \|x - a\|, \tag{2.115}$$

and for every $n \in \mathbb{N}$ and every closed and non-empty set $A \subseteq \mathbb{R}^d$ let $f_{A,n} \colon \mathbb{R}^d \to [0,1]$ be the function which satisfies for all $x \in \mathbb{R}^d$ that

$$f_{A,n}(x) = h_n(D_A(x)). \tag{2.116}$$

Observe that the fact that W has continuous sample paths, the fact that for all $t \in [0,T)$, $s \in (t,T]$, $k \in \mathbb{N}$ it holds that $W_s - W_{\min\{t+1/k,s\}}$ and $\mathbb{H}_t$ are independent, Klenke [70, Theorem 5.4], and the theorem of dominated convergence assure that for all $t \in [0,T)$, $s \in (t,T]$, $B \in \mathbb{H}_t$ and all globally bounded and continuous functions $g \colon \mathbb{R}^d \to \mathbb{R}$ it holds that

$$\begin{aligned}
\mathbb{E}[g(W_s - W_t)\mathbb{1}_B] &= \mathbb{E}\left[\left(\lim_{k\to\infty} g(W_s - W_{\min\{t+1/k,s\}})\right)\mathbb{1}_B\right] \\
&= \lim_{k\to\infty} \mathbb{E}\big[g(W_s - W_{\min\{t+1/k,s\}})\mathbb{1}_B\big] \\
&= \lim_{k\to\infty} \mathbb{E}\big[g(W_s - W_{\min\{t+1/k,s\}})\big]\,\mathbb{E}[\mathbb{1}_B] \\
&= \mathbb{E}\left[\lim_{k\to\infty} g(W_s - W_{\min\{t+1/k,s\}})\right]\mathbb{P}(B) = \mathbb{E}[g(W_s - W_t)]\,\mathbb{P}(B).
\end{aligned} \tag{2.117}$$

Next note that the fact that closed and non-empty sets $A \subseteq \mathbb{R}^d$ and all $x \in \mathbb{R}^d$ it holds that $D_A(x) = 0 \Leftrightarrow x \in A$ assures that for all closed and non-empty sets $A \subseteq \mathbb{R}^d$ and all $x \in \mathbb{R}^d$ it holds that

$$\lim_{n\to\infty} f_{A,n}(x) = \mathbb{1}_A(x). \tag{2.118}$$

Moreover, note that the fact that for every $n \in \mathbb{N}$ it holds that $h_n \colon [0,\infty) \to [0,1]$ is a continuous function and the fact that for every closed and non-empty set $A \subseteq \mathbb{R}^d$ it holds that $D_A \colon \mathbb{R}^d \to [0,\infty)$ is a continuous function assure that for every $n \in \mathbb{N}$ and every closed and non-empty set $A \subseteq \mathbb{R}^d$ it holds that $f_{A,n} \colon \mathbb{R}^d \to [0,1]$ is a continuous function. Combining this, (2.117), (2.118), and the theorem of dominated convergence shows that for all $t \in [0,T)$, $s \in (t,T]$, $B \in \mathbb{H}_t$ and all closed and non-empty sets $A \subseteq \mathbb{R}^d$ it holds that

$$\begin{aligned}
\mathbb{P}(\{W_s - W_t \in A\} \cap B) &= \mathbb{E}[\,\mathbb{1}_A(W_s - W_t)\mathbb{1}_B] = \lim_{n\to\infty} \mathbb{E}[\,f_{A,n}(W_s - W_t)\mathbb{1}_B] \\
&= \lim_{n\to\infty}\Big(\mathbb{E}[\,f_{A,n}(W_s - W_t)]\,\mathbb{P}(B)\Big) = \mathbb{E}[\,\mathbb{1}_A(W_s - W_t)]\,\mathbb{P}(B) \\
&= \mathbb{P}(\{W_s - W_t \in A\})\,\mathbb{P}(B).
\end{aligned} \tag{2.119}$$

This proves that for all $t \in [0,T)$, $s \in (t,T]$, $B \in \mathbb{H}_t$ it holds that $(\mathbb{1}_B)^{-1}(\{\}, \{0\}, \{1\}, \{0,1\})$ and $(W_s - W_t)^{-1}(\{A \subseteq \mathbb{R}^d \colon A \text{ is a closed set}\})$ are independent. The fact that $\{A \subseteq \mathbb{R}^d \colon A \text{ is a closed set}\}$ is closed under intersections, the fact that $\mathfrak{S}(\{A \subseteq \mathbb{R}^d \colon A \text{ is a closed set}\}) = \mathcal{B}(\mathbb{R}^d)$, and Klenke [70, Theorem 2.16] hence assure that for all $t \in [0,T)$, $s \in (t,T]$, $B \in \mathbb{H}_t$ it holds that $W_s - W_t$ and B are independent. This implies that for all $t \in [0,T]$, $s \in [t,T]$ it holds that $W_s - W_t$ and $\mathbb{H}_t$ are independent. Combining this with the hypothesis that W is a Brownian motion, and the fact that $W\colon [0,T]\times\Omega\to\mathbb{R}^m$ is an $(\mathbb{H}_t)_{t\in[0,T]}$/$\mathcal{B}(\mathbb{R}^m)$-adapted stochastic process establishes that $W\colon [0,T]\times\Omega\to\mathbb{R}^m$ is a standard $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{H}_t)_{t\in[0,T]})$-Brownian motion. The proof of Lemma 2.18 is thus completed. □

2.8 On a distributional flow property for solutions of SDEs

In this subsection we prove a distributional flow property for solutions of SDEs in Lemma 2.20 below. The idea for the proof of Lemma 2.20 is based on the observation that if we replace solution processes of SDEs by Euler-Maruyama approximations the flow-type condition trivially holds (cf. the argument below (2.146) in the proof of Lemma 2.20 below). To prove Lemma 2.20 below we also need, besides several auxiliary results of the previous subsections, the following well-known statement (see Lemma 2.19 below).

Lemma 2.19. *Let $d,m\in\mathbb{N}$, $T\in(0,\infty)$, $t\in[0,T]$, $s\in[t,T]$, let $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_r)_{r\in[0,T]})$ be a filtered probability space which satisfies the usual conditions, let $W\colon[0,T]\times\Omega\to\mathbb{R}^m$ be a standard $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_r)_{r\in[0,T]})$-Brownian motion, let $\mu\colon[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$ and $\sigma\colon[0,T]\times\mathbb{R}^d\to\mathbb{R}^{d\times m}$ be globally Lipschitz continuous functions, let $X=(X_r(x))_{r\in[t,s],x\in\mathbb{R}^d}\colon[t,s]\times\mathbb{R}^d\times\Omega\to\mathbb{R}^d$ be a continuous random field which satisfies for every $x\in\mathbb{R}^d$ that $(X_r(x))_{r\in[t,s]}\colon[t,s]\times\Omega\to\mathbb{R}^d$ is an $(\mathbb{F}_r)_{r\in[t,s]}$/$\mathcal{B}(\mathbb{R}^d)$-adapted stochastic process and which satisfies that for all $r\in[t,s]$, $x\in\mathbb{R}^d$ it holds $\mathbb{P}$-a.s. that*

$$X_r(x)=x+\int_t^r\mu(h,X_h(x))\,dh+\int_t^r\sigma(h,X_h(x))\,dW_h, \tag{2.120}$$

and let $\xi\colon\Omega\to\mathbb{R}^d$ be an $\mathbb{F}_t$/$\mathcal{B}(\mathbb{R}^d)$-measurable function with $\mathbb{E}\Big[\|\xi\|^2\Big]<\infty$. Then for all $r\in[t,s]$ it holds $\mathbb{P}$-a.s. that

$$X_r(\xi)=\xi+\int_t^r\mu\big(h,X_h(\xi)\big)\,dh+\int_t^r\sigma\big(h,X_h(\xi)\big)\,dW_h. \tag{2.121}$$

Proof of Lemma 2.19. Throughout this proof assume w.l.o.g. that $s>t$, let $(u_n^{N,r})_{n\in\{0,1,2,\dots,N\},N\in\mathbb{N},r\in(t,s]}\subseteq[t,s]$ satisfy for all $N\in\mathbb{N}$, $n\in\{0,1,2,\dots,N\}$, $r\in(t,s]$ that $u_n^{N,r}=t+\frac{n(r-t)}{N}$, for every $N\in\mathbb{N}$, $r\in(t,s]$ let $\mathcal{X}^{N,r}=(\mathcal{X}_n^{N,r}(x))_{n\in\{0,1,2,\dots,N\},x\in\mathbb{R}^d}\colon\{0,1,2,\dots,N\}\times\mathbb{R}^d\times\Omega\to\mathbb{R}^d$ be the continuous random field which satisfies for all $n\in\{1,2,\dots,N\}$, $x\in\mathbb{R}^d$ that $\mathcal{X}_0^{N,r}(x)=x$ and

$$\mathcal{X}_n^{N,r}(x)=\mathcal{X}_{n-1}^{N,r}(x)+\mu\big(u_{n-1}^{N,r},\mathcal{X}_{n-1}^{N,r}(x)\big)\tfrac{(r-t)}{N}+\sigma\big(u_{n-1}^{N,r},\mathcal{X}_{n-1}^{N,r}(x)\big)(W_{u_n^{N,r}}-W_{u_{n-1}^{N,r}}), \tag{2.122}$$

let $\|\cdot\|\colon\mathbb{R}^d\to[0,\infty)$ be the Euclidean norm on $\mathbb{R}^d$, let $|||\cdot|||\colon\mathbb{R}^{d\times m}\to[0,\infty)$ be the Frobenius norm on $\mathbb{R}^{d\times m}$, and let $L\in[0,\infty)$ satisfy for all $r,h\in[0,T]$, $x,y\in\mathbb{R}^d$ that

$$\max\big\{\|\mu(r,x)-\mu(h,y)\|,|||\sigma(r,x)-\sigma(h,y)|||\big\}\le L\big(|r-h|+\|x-y\|\big). \tag{2.123}$$

Note that (2.120), (2.122), (2.123), Corollary 2.13 (with $d=d$, $m=m$, $N=N$, $T=T$, $t=t$, $s=r$, $L=L$, $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_h)_{h\in[0,T]})=(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_h)_{h\in[0,T]})$, $W=W$, $\zeta=x$, $\mu=\mu$, $\sigma=\sigma$, $(X_h)_{h\in[t,s]}=(X_h)_{h\in[t,r]}$, $(r_n)_{n\in\{0,1,\dots,N\}}=(u_n^{N,r})_{n\in\{0,1,\dots,N\}}$, $(\mathcal{X}_n)_{n\in\{0,1,\dots,N\}}=(\mathcal{X}_n^{N,r}(x))_{n\in\{0,1,\dots,N\}}$ for $N\in\mathbb{N}$, $x\in\mathbb{R}^d$, $r\in(t,s]$ in the notation of Corollary 2.13), and Lemma 2.11 (with $d=d$, $m=m$, $T=r-t$, $\xi=x$, $L=L$, $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_h)_{h\in[0,T]})=(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_{t+r})_{r\in[0,r-t]})$, $(W_h)_{h\in[0,T]}=(W_{t+h}-W_t)_{h\in[0,r-t]}$, $(\mu(h,x))_{h\in[0,T],x\in\mathbb{R}^d}=(\mu(t+h,x))_{h\in[0,r-t],x\in\mathbb{R}^d}$, $(\sigma(h,x))_{h\in[0,T],x\in\mathbb{R}^d}=(\sigma(t+h,x))_{h\in[0,r-t],x\in\mathbb{R}^d}$, $(X_h)_{h\in[0,T]}=(X_{t+h})_{h\in[0,r-t]}$ for $x\in\mathbb{R}^d$, $r\in(t,s]$ in the notation of Lemma 2.11) assure that for all $x \in \mathbb{R}^d$, $N \in \mathbb{N}$, $r \in (t, s]$ it holds that

$$
\begin{aligned}
&\Big(\mathbb{E}\Big[\big\|X_r(x) - \mathcal{X}_N^{N,r}(x)\big\|^2\Big]\Big)^{1/2} \\
&\leq \bigg[\Big(\mathbb{E}\Big[\|X_t(x) - x\|^2\Big]\Big)^{1/2} + \max_{k \in \{1,2,\dots,N\}} |u_k^{N,r} - u_{k-1}^{N,r}|^{1/2}\bigg] \\
&\quad \cdot \exp\Big((1+L)^2(1+\sqrt{T})^4\Big)\left(1 + \sup_{h,l \in [t,r], h \neq l} \tfrac{(\mathbb{E}[\|X_h(x) - X_l(x)\|^2])^{1/2}}{|h-l|^{1/2}}\right) \\
&\leq \frac{\sqrt{|r-t|}}{\sqrt{N}} \exp\Big((1+L)^2(1+\sqrt{T})^4\Big) \\
&\quad \cdot \Big(1 + (1 + \|x\|) \exp\Big(10\big(\max\{\|\mu(t,0)\|, |||\sigma(t,0)|||, L, 1\} + LT\big)^2 (T+1)(L+1)\Big)\Big) \\
&\leq \frac{(1+\|x\|)}{\sqrt{N}} \exp\Big(12\big(\max\{\|\mu(t,0)\|, |||\sigma(t,0)|||, L, 1\} + LT\big)^2 (1+L)^2(1+\sqrt{T})^4\Big).
\end{aligned}
\tag{2.124}
$$

This ensures that for all $r \in [t, s]$, $x \in \mathbb{R}^d$ it holds that $\limsup_{N \to \infty} \mathbb{E}[\|X_r(x) - \mathcal{X}_N^{N,r}(x)\|^2] = 0$. This and the fact that for all $r \in [t, s]$, $x \in \mathbb{R}^d$, $N \in \mathbb{N}$ it holds that $\mathcal{X}_N^{N,r}(x)\colon \Omega \to \mathbb{R}^d$ is $\mathfrak{S}(W_h - W_t\colon h \in [t, r])$/$\mathcal{B}(\mathbb{R}^d)$-measurable imply that for all $r \in [t, s]$, $x \in \mathbb{R}^d$ it holds that $X_r(x)\colon \Omega \to \mathbb{R}^d$ is $\mathfrak{S}(\mathfrak{S}(W_h - W_t\colon h \in [t, r]) \cup \{A \in \mathcal{F}\colon \mathbb{P}(A) = 0\})$/$\mathcal{B}(\mathbb{R}^d)$-measurable. Combining this with the fact that $\xi\colon \Omega \to \mathbb{R}^d$ is an $\mathbb{F}_t$/$\mathcal{B}(\mathbb{R}^d)$-measurable function and the fact that $W\colon [0, T] \times \Omega \to \mathbb{R}^m$ is a standard $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_r)_{r \in [0,T]})$-Brownian motion demonstrates for all $r \in [t, s]$, $N \in \mathbb{N}$ it holds that $(X_r(x) - \mathcal{X}_N^{N,r}(x))_{x \in \mathbb{R}^d}$ and ξ are independent. Lemma 2.16 and (2.124) hence assure that for all $N \in \mathbb{N}$, $r \in (t, s]$ it holds that

$$
\begin{aligned}
&\mathbb{E}\Big[\big\|X_r(\xi) - \mathcal{X}_N^{N,r}(\xi)\big\|^2\Big] \\
&= \int_{\mathbb{R}^d} \mathbb{E}\Big[\big\|X_r(x) - \mathcal{X}_N^{N,r}(x)\big\|^2\Big] \, (\xi(\mathbb{P})_{\mathcal{B}(\mathbb{R}^d)})(dx) \\
&\leq \int_{\mathbb{R}^d} \left[\tfrac{\exp\left(12(\max\{\|\mu(0,0)\|, |||\sigma(0,0)|||, L, 1\} + LT)^2(1+L)^2(1+\sqrt{T})^4\right)}{\sqrt{N}} (1 + \|x\|)\right]^2 (\xi(\mathbb{P})_{\mathcal{B}(\mathbb{R}^d)})(dx) \\
&\leq \left[\tfrac{\exp\left(24(\max\{\|\mu(t,0)\|, |||\sigma(t,0)|||, L, 1\} + LT)^2(1+L)^2(1+\sqrt{T})^4\right)}{N}\right] 2\left(1 + \mathbb{E}\Big[\|\xi\|^2\Big]\right).
\end{aligned}
\tag{2.125}
$$

Next observe that the hypothesis that μ and σ are globally Lipschitz continuous functions, the hypothesis that $\mathbb{E}\Big[\|\xi\|^2\Big] < \infty$, and the existence theorem for the solutions of SDEs (see, e.g., Karatzas & Shreve [68, Proposition 5.2.9]) prove that there exists an $(\mathbb{F}_r)_{r \in [t,s]}$/$\mathcal{B}(\mathbb{R}^d)$-adapted stochastic process $Y\colon [t, s] \times \Omega \to \mathbb{R}^d$ with continuous sample paths which satisfies that for all $r \in [t, s]$ it holds $\mathbb{P}$-a.s. that

$$
Y_r = \xi + \int_t^r \mu(h, Y_h)\, dr + \int_t^r \sigma(h, Y_h)\, dW_h. \tag{2.126}
$$

Moreover, observe that (2.122) ensures that for all $N \in \mathbb{N}$, $n \in \{1, 2, \dots, N\}$, $r \in (t, s]$ and all functions $\zeta\colon \Omega \to \mathbb{R}^d$ it holds that $\mathcal{X}_0^{N,r}(\zeta) = \zeta$ and

$$
\mathcal{X}_n^{N,r}(\zeta) = \mathcal{X}_{n-1}^{N,r}(\zeta) + \mu\big(u_{n-1}^{N,r}, \mathcal{X}_{n-1}^{N,r}(\zeta)\big)\tfrac{(r-t)}{N} + \sigma\big(u_{n-1}^{N,r}, \mathcal{X}_{n-1}^{N,r}(\zeta)\big)(W_{u_n^{N,r}} - W_{u_{n-1}^{N,r}}). \tag{2.127}
$$

Combining this, (2.123), the fact that $\mathbb{E}\Big[\|Y_t\|^2\Big] = \mathbb{E}\Big[\|\xi\|^2\Big] < \infty$, and (2.126) with Corollary 2.13 (with $d = d$, $m = m$, $N = N$, $T = T$, $t = t$, $s = r$, $L = L$, $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_h)_{h \in [0,T]}) =$

$(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_h)_{h\in[0,T]})$, $W = W$, $\zeta = \xi$, $\mu = \mu$, $\sigma = \sigma$, $(X_h)_{h\in[t,s]} = (Y_h)_{h\in[t,r]}$, $(r_n)_{n\in\{0,1,\dots,N\}} = (u_n^{N,r})_{n\in\{0,1,\dots,N\}}$, $(\mathcal{X}_n)_{n\in\{0,1,\dots,N\}} = (\mathcal{X}_n^{N,r}(\xi))_{n\in\{0,1,\dots,N\}}$ for $N \in \mathbb{N}$, $r \in (t,s]$ in the notation of Corollary 2.13) demonstrates that for all $N \in \mathbb{N}$, $r \in (t,s]$ it holds that

$$\begin{aligned}
&\Big(\mathbb{E}\Big[\big\|Y_r - \mathcal{X}_N^{N,r}(\xi)\big\|^2\Big]\Big)^{1/2} \\
&\leq \bigg[\Big(\mathbb{E}\Big[\|Y_t - \xi\|^2\Big]\Big)^{1/2} + \max_{k\in\{1,2,\dots,N\}} |u_k^{N,r} - u_{k-1}^{N,r}|^{1/2}\bigg] \\
&\quad \cdot \exp\Big((1+L)^2(1+\sqrt{T})^4\Big)\left(1 + \sup_{h,l\in[t,r],h\neq l} \tfrac{(\mathbb{E}[\|Y_h - Y_l\|^2])^{1/2}}{|h-l|^{1/2}}\right) \\
&= \tfrac{\sqrt{|r-t|}\exp\left((1+L)^2(1+\sqrt{T})^4\right)}{\sqrt{N}}\left(1 + \sup_{h,l\in[t,r],h\neq l} \tfrac{(\mathbb{E}[\|Y_h - Y_l\|^2])^{1/2}}{|h-l|^{1/2}}\right) < \infty.
\end{aligned} \tag{2.128}$$

The triangle inequality and (2.125) hence show that for all $r \in (t,s]$ it holds that

$$\begin{aligned}
&\Big(\mathbb{E}\Big[\big\|X_r(\xi) - Y_r\big\|^2\Big]\Big)^{1/2} \\
&\leq \limsup_{N\to\infty}\bigg[\Big(\mathbb{E}\Big[\big\|X_r(\xi) - \mathcal{X}_N^{N,r}(\xi)\big\|^2\Big]\Big)^{1/2} + \Big(\mathbb{E}\Big[\big\|\mathcal{X}_N^{N,r}(\xi) - Y_r\big\|^2\Big]\Big)^{1/2}\bigg] \\
&\leq \bigg[\limsup_{N\to\infty}\Big(\mathbb{E}\Big[\big\|X_r(\xi) - \mathcal{X}_N^{N,r}(\xi)\big\|^2\Big]\Big)^{1/2}\bigg] + \bigg[\limsup_{N\to\infty}\Big(\mathbb{E}\Big[\big\|\mathcal{X}_N^{N,r}(\xi) - Y_r\big\|^2\Big]\Big)^{1/2}\bigg] = 0.
\end{aligned} \tag{2.129}$$

Combining this with the fact that $(X_r(\xi))_{r\in[t,s]}$ and $(Y_r)_{r\in[t,s]}$ are continuous random fields demonstrates that

$$\mathbb{P}\big(\forall\, r \in [t,s]\colon X_r(\xi) = Y_r\big) = \mathbb{P}\big(\forall\, r \in (t,s]\cap\mathbb{Q}\colon X_r(\xi) = Y_r\big) = 1. \tag{2.130}$$

This and (2.126) prove that for all $r \in [t,s]$ it holds $\mathbb{P}$-a.s. that

$$X_r(\xi) = \xi + \int_t^r \mu\big(h, X_h(\xi)\big)\,dh + \int_t^r \sigma\big(h, X_h(\xi)\big)\,dW_h. \tag{2.131}$$

The proof of Lemma 2.19 is thus completed. □

Lemma 2.20. *Let $d, m \in \mathbb{N}$, $T \in (0,\infty)$, let $\mu\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ and $\sigma\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}^{d\times m}$ be globally Lipschitz continuous functions, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, let $(\mathbb{F}_t^1)_{t\in[0,T]}$ and $(\mathbb{F}_t^2)_{t\in[0,T]}$ be filtrations on $(\Omega, \mathcal{F}, \mathbb{P})$ which satisfy the usual conditions, assume that $\mathbb{F}_T^1$ and $\mathbb{F}_T^2$ are independent, for every $i \in \{1,2\}$ let $W^i\colon [0,T]\times\Omega \to \mathbb{R}^m$ be a standard $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_t^i)_{t\in[0,T]})$-Brownian motion, and for every $i \in \{1,2\}$ let $X^i = (X_{t,s}^i(x))_{s\in[t,T],t\in[0,T],x\in\mathbb{R}^d}\colon \{(t,s)\in[0,T]^2\colon t\leq s\}\times\mathbb{R}^d\times\Omega \to \mathbb{R}^d$ be a continuous random field which satisfies for every $t \in [0,T]$, $x \in \mathbb{R}^d$ that $(X_{t,s}^i(x))_{s\in[t,T]}\colon [t,T]\times\Omega \to \mathbb{R}^d$ is an $(\mathbb{F}_s^i)_{s\in[t,T]}$/$\mathcal{B}(\mathbb{R}^d)$-adapted stochastic process and which satisfies that for all $t \in [0,T]$, $s \in [t,T]$, $x \in \mathbb{R}^d$ it holds $\mathbb{P}$-a.s. that*

$$X_{t,s}^i(x) = x + \int_t^s \mu\big(r, X_{t,r}^i(x)\big)\,dr + \int_t^s \sigma\big(r, X_{t,r}^i(x)\big)\,dW_r^i. \tag{2.132}$$

Then it holds for all $r,s,t \in [0,T]$, $x \in \mathbb{R}^d$, $B \in \mathcal{B}(\mathbb{R}^d)$ with $t \leq s \leq r$ that $\mathbb{P}(X_{t,t}^1(x) = x) = 1$ and

$$\mathbb{P}\big(X_{s,r}^1(X_{t,s}^2(x)) \in B\big) = \mathbb{P}\big(X_{t,r}^1(x) \in B\big). \tag{2.133}$$

Proof of Lemma 2.20. Throughout this proof let $r,s,t \in [0,T]$, $x \in \mathbb{R}^d$ satisfy that $t \leq s \leq r$, let $(u_n^N)_{n\in\{0,1,2,\dots,N\},N\in\mathbb{N}} \subseteq [t,s]$, $(v_n^N)_{n\in\{0,1,2,\dots,N\},N\in\mathbb{N}} \subseteq [s,r]$ satisfy for all $N \in \mathbb{N}$,

$n \in \{0, 1, 2, \dots, N\}$ that $u_n^N = t + \frac{n(s-t)}{N}$ and $v_n^N = s + \frac{n(r-s)}{N}$, for every $N \in \mathbb{N}$ let $\mathcal{X}^N \colon \{0, 1, 2, \dots, 2N\} \times \Omega \to \mathbb{R}^d$ and $\mathcal{Y}^N, \mathcal{Z}^N \colon \{0, 1, 2, \dots, N\} \times \Omega \to \mathbb{R}^d$ be the stochastic processes which satisfy for all $n \in \{1, 2, \dots, N\}$ that

$$\mathcal{X}_0^N = x, \qquad \mathcal{X}_n^N = \mathcal{X}_{n-1}^N + \mu(u_{n-1}^N, \mathcal{X}_{n-1}^N)\tfrac{(s-t)}{N} + \sigma(u_{n-1}^N, \mathcal{X}_{n-1}^N)(W^1_{u_n^N} - W^1_{u_{n-1}^N}), \tag{2.134}$$

$$\mathcal{X}_{N+n}^N = \mathcal{X}_{N+n-1}^N + \mu(v_{n-1}^N, \mathcal{X}_{n-1}^N)\tfrac{(r-s)}{N} + \sigma(v_{n-1}^N, \mathcal{X}_{N+n-1}^N)(W^1_{v_n^N} - W^1_{v_{n-1}^N}), \tag{2.135}$$

$$\mathcal{Y}_0^N = x, \qquad \mathcal{Y}_n^N = \mathcal{Y}_{n-1}^N + \mu(u_{n-1}^N, \mathcal{Y}_{n-1}^N)\tfrac{(s-t)}{N} + \sigma(u_{n-1}^N, \mathcal{Y}_{n-1}^N)(W^2_{u_n^N} - W^2_{u_{n-1}^N}), \tag{2.136}$$

$$\mathcal{Z}_0^N = \mathcal{Y}_N^N, \qquad \text{and} \qquad \mathcal{Z}_n^N = \mathcal{Z}_{n-1}^N + \mu(v_{n-1}^N, \mathcal{Z}_{n-1}^N)\tfrac{(r-s)}{N} + \sigma(v_{n-1}^N, \mathcal{Z}_{n-1}^N)(W^1_{v_n^N} - W^1_{v_{n-1}^N}), \tag{2.137}$$

let $\mathbb{G}_h \subseteq \mathcal{F}$, $h \in [0,T]$, and $\mathbb{H}_h \subseteq \mathcal{F}$, $h \in [0,T]$, be the sigma-algebras which satisfy for all $h \in [0,T]$ that

$$\mathbb{G}_h = \mathfrak{S}(\mathbb{F}_h^1 \cup \mathbb{F}_h^2) \qquad \text{and} \qquad \mathbb{H}_h = \begin{cases} \cap_{l \in (h,T]} \mathbb{G}_l & : h < T \\ \mathbb{G}_T & : h = T, \end{cases} \tag{2.138}$$

let $\langle \cdot, \cdot \rangle : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be the Euclidean scalar product on $\mathbb{R}^d$, let $\|\cdot\| : \mathbb{R}^d \to [0,\infty)$ be the Euclidean norm on $\mathbb{R}^d$, and let $|||\cdot||| \colon \mathbb{R}^{d\times m} \to [0,\infty)$ be the Frobenius norm on $\mathbb{R}^{d\times m}$. Note that the hypothesis that $(\mathbb{F}_t^1)_{t\in[0,T]}$ and $(\mathbb{F}_t^2)_{t\in[0,T]}$ are filtrations on $(\Omega, \mathcal{F}, \mathbb{P})$ which satisfy the usual conditions and (2.138) imply that $(\mathbb{H}_t)_{t\in[0,T]}$ is a filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ which satisfies the usual conditions. Moreover, observe that (2.132) assures that

$$\mathbb{P}(X_{t,t}^1(x) = x) = 1. \tag{2.139}$$

Furthermore, note that the hypothesis that μ and σ are globally Lipschitz continuous, (2.132), (2.134), (2.135), (2.136), and Corollary 2.13 demonstrate that there exists a real number $C \in (0,\infty)$ which satisfies that for all $N \in \mathbb{N}$ it holds that

$$\Big(\mathbb{E}\Big[\big\|X_{t,r}^1(x) - \mathcal{X}_{2N}^N\big\|^2\Big]\Big)^{1/2} \le \frac{C}{\sqrt{N}} \qquad \text{and} \qquad \Big(\mathbb{E}\Big[\big\|X_{t,s}^2(x) - \mathcal{Y}_N^N\big\|^2\Big]\Big)^{1/2} \le \frac{C}{\sqrt{N}}. \tag{2.140}$$

This implies that

$$\limsup_{N\to\infty} \Big(\mathbb{E}\Big[\big\|X_{t,r}^1(x) - \mathcal{X}_{2N}^N\big\|^2\Big]\Big)^{1/2} \le \limsup_{N\to\infty} \frac{C}{\sqrt{N}} = 0. \tag{2.141}$$

Moreover, observe that the hypothesis that μ and σ are globally Lipschitz continuous implies that

$$\sup_{h\in[0,T], y\in\mathbb{R}^d} \frac{\langle y, \mu(h,y)\rangle + |||\sigma(h,y)|||^2}{1 + \|y\|^2} < \infty. \tag{2.142}$$

Lemma 2.7 therefore demonstrates that

$$\mathbb{E}\Big[\big\|X_{t,s}^2(x)\big\|^2\Big] < \infty. \tag{2.143}$$

Next note that the fact that for all $h \in [0,T]$, $l \in [h,T]$ it holds that $W_l^1 - W_h^1$, $\mathbb{F}_h^1$, and $\mathbb{F}_h^2$ are independent assures that for all $h \in [0,T]$, $l \in [h,T]$ it holds that $W_l^1 - W_h^1$ and $\mathbb{G}_h$ are independent. This, the fact that $W^1 \colon [0,T] \times \Omega \to \mathbb{R}^d$ is a Brownian motion, and the fact that $W^1 \colon [0,T] \times \Omega \to \mathbb{R}^d$ is an $(\mathbb{G}_h)_{h\in[0,T]}$/$\mathcal{B}(\mathbb{R}^d)$-adapted stochastic process imply that $W^1 \colon [0,T] \times \Omega \to \mathbb{R}^d$ is a standard $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{G}_h)_{h\in[0,T]})$-Brownian motion. Lemma 2.18 and (2.138) hence ensure that $W^1 \colon [0,T] \times \Omega \to \mathbb{R}^d$ is a standard $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{H}_h)_{h\in[0,T]})$-Brownian motion. Combining this, the fact that $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{H}_h)_{h\in[0,T]})$ is a filtered probability space which satisfies the usual conditions, the fact that for all $y \in \mathbb{R}^d$ it holds that $(X_{s,h}^1(y))_{h\in[s,r]} \colon [s,r] \times \Omega \to \mathbb{R}^d$ is an $(\mathbb{H}_h)_{h\in[s,r]}$/$\mathcal{B}(\mathbb{R}^d)$-adapted stochastic process, (2.132), the fact that $X^2_{t,s}(x)\colon \Omega \to \mathbb{R}^d$ is $\mathbb{H}_s/\mathcal{B}(\mathbb{R}^d)$-measurable, and (2.143) with Lemma 2.19 (with $d = d$, $m = m$, $T = T$, $t = s$, $s = r$, $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_h)_{h\in[0,T]}) = (\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{H}_h)_{h\in[0,T]})$, $W = W^1$, $\mu = \mu$, $\sigma = \sigma$, $(X_h(y))_{h\in[t,s],y\in\mathbb{R}^d} = (X^1_{s,h}(y))_{h\in[s,r],y\in\mathbb{R}^d}$, $\xi = X^2_{t,s}(x)$ in the notation of Lemma 2.19) proves that for all $h \in [s, r]$ it holds $\mathbb{P}$-a.s. that

$$X^1_{s,h}(X^2_{t,s}(x)) = X^2_{t,s}(x) + \int_s^h \mu\big(l, X^1_{s,l}(X^2_{t,s}(x))\big)dl + \int_s^h \sigma\big(l, X^1_{s,l}(X^2_{t,s}(x))\big)dW^1_l. \tag{2.144}$$

The fact that $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{H}_h)_{h\in[0,T]})$ is a filtered probability space which satisfies the usual conditions, the fact that $W^1\colon [0,T]\times\Omega \to \mathbb{R}^d$ is a standard $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{H}_h)_{h\in[0,T]})$-Brownian motion, the fact that $\mathcal{Y}^N_N\colon \Omega \to \mathbb{R}^d$ is $\mathbb{H}_s/\mathcal{B}(\mathbb{R}^d)$-measurable, the hypothesis that μ and σ are globally Lipschitz continuous functions, the fact that $(X^1_{s,h}(X^2_{t,s}(x)))_{h\in[s,r]}\colon [s,r]\times\Omega \to \mathbb{R}^d$ is an $(\mathbb{H}_h)_{h\in[s,r]}/\mathcal{B}(\mathbb{R}^d)$-adapted stochastic process with continuous sample paths, (2.143), (2.137), and Corollary 2.13 (with $d = d$, $m = m$, $N = N$, $T = T$, $t = s$, $s = r$, $L = \sup_{h,l\in[0,T],y,z\in\mathbb{R}^d\colon (h,y)\neq(l,z)} \frac{\|\mu(h,y)-\mu(l,z)\|+|||\sigma(h,y)-\sigma(l,z)|||}{|h-l|+\|y-z\|}$, $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_h)_{h\in[0,T]}) = (\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{H}_h)_{h\in[0,T]})$, $W = W^1$, $\zeta = \mathcal{Y}^N_N$, $\mu = \mu$, $\sigma = \sigma$, $(X_h)_{h\in[t,s]} = (X^1_{s,h}(X^2_{t,s}(x)))_{h\in[s,r]}$, $(r_n)_{n\in\{0,1,\ldots,N\}} = (v^N_n)_{n\in\{0,1,\ldots,N\}}$, $(\mathcal{X}_n)_{n\in\{0,1,\ldots,N\}} = (\mathcal{Z}^N_n)_{n\in\{0,1,\ldots,N\}}$ for $N \in \mathbb{N}$ in the notation of Corollary 2.13) hence demonstrate that there exists a real number $K \in (0, \infty)$ which satisfies that for all $N \in \mathbb{N}$ it holds that

$$\Big(\mathbb{E}\Big[\big\|X^1_{s,r}(X^2_{t,s}(x)) - \mathcal{Z}^N_N\big\|^2\Big]\Big)^{1/2} \leq K\left[\Big(\mathbb{E}\Big[\big\|X^2_{t,s}(x) - \mathcal{Y}^N_N\big\|^2\Big]\Big)^{1/2} + \frac{1}{\sqrt{N}}\right]. \tag{2.145}$$

This and (2.140) imply that

$$\limsup_{N\to\infty}\Big(\mathbb{E}\Big[\big\|X^1_{s,r}(X^2_{t,s}(x)) - \mathcal{Z}^N_N\big\|^2\Big]\Big)^{1/2} \leq \limsup_{N\to\infty} K\left[\frac{C}{\sqrt{N}} + \frac{1}{\sqrt{N}}\right] = 0. \tag{2.146}$$

Furthermore, observe that (2.134)–(2.137) assure that for all $N \in \mathbb{N}$ it holds that $\mathcal{X}^N_{2N}$ and $\mathcal{Z}^N_N$ have the same distribution. This, (2.141), and (2.146) imply that for all globally bounded and Lipschitz continuous functions $g\colon \mathbb{R}^d \to \mathbb{R}$ it holds that

$$\mathbb{E}\big[g(X^1_{s,r}(X^2_{t,s}(x)))\big] = \lim_{N\to\infty}\mathbb{E}\big[g(\mathcal{Z}^N_N)\big] = \lim_{N\to\infty}\mathbb{E}\big[g(\mathcal{X}^N_{2N})\big] = \mathbb{E}\big[g(X^1_{t,r}(x))\big]. \tag{2.147}$$

Lemma 2.14 hence assures that $X^1_{s,r}(X^2_{t,s}(x))$ and $X^1_{t,r}(x)$ are identically distributed. Combining this with (2.139) completes the proof of Lemma 2.20. □

3 Full history recursive multilevel Picard (MLP) approximation algorithms

In this section we present the proposed MLP scheme and perform a rigorous complexity analysis. First, we introduce our MLP scheme (cf. (3.5) in Subsection 3.1 below) as an approximation algorithm for a solution (cf. u in Setting 3.1 in Subsection 3.1 below) of a certain type of stochastic fixed point equation (cf. (3.4) in Subsection 3.1 below) in Subsection 3.1. Subsequently, the goal of Subsections 3.2–3.4 is to obtain an estimate for the L^2-error between the MLP scheme and the solution of the stochastic fixed point equation (see Proposition 3.15 and Corollary 3.16 in Subsection 3.4 below). In Subsection 3.5 we estimate the computational effort needed to simulate realizations of the MLP scheme and combine this with the L^2-error estimate in Corollary 3.16 to obtain a computational complexity analysis for the MLP algorithm in Proposition 3.18. Finally, in Subsection 3.6, we exploit a connection between stochastic fixed point equations and viscosity solutions of semilinear Kolmogorov PDEs (see Proposition 3.22 in Subsection 3.6.1 below) to carry over the complexity analysis of Subsection 3.5 to semilinear Kolmogorov PDEs. This culminates in Theorem 3.24, the main result of this paper, which demonstrates that our proposed MLP algorithm overcomes the curse of dimensionality in the approximation of certain semilinear Kolmogorov PDEs.

3.1 Stochastic fixed point equations and MLP approximations

Setting 3.1. Let $d \in \mathbb{N}$, $T \in (0,\infty)$, $L \in [0,\infty)$, $\Theta = \cup_{n=1}^{\infty} \mathbb{Z}^n$, $u \in C([0,T] \times \mathbb{R}^d, \mathbb{R})$, $g \in C(\mathbb{R}^d, \mathbb{R})$, $f \in C([0,T] \times \mathbb{R}^d \times \mathbb{R}, \mathbb{R})$ satisfy for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $v, w \in \mathbb{R}$ that

$$|f(t,x,v) - f(t,x,w)| \leq L|v - w|, \tag{3.1}$$

let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\mathcal{R}^\theta \colon \Omega \to [0,1]$, $\theta \in \Theta$, be independent $\mathcal{U}_{[0,1]}$-distributed random variables, let $R^\theta = (R^\theta_t)_{t \in [0,T]} \colon [0,T] \times \Omega \to [0,T]$, $\theta \in \Theta$, be the stochastic processes which satisfy for all $t \in [0,T]$, $\theta \in \Theta$ that

$$R^\theta_t = t + (T-t)\mathcal{R}^\theta, \tag{3.2}$$

let $X^\theta = (X^\theta_{t,s}(x))_{s \in [t,T], t \in [0,T], x \in \mathbb{R}^d} \colon \{(t,s) \in [0,T]^2 \colon t \leq s\} \times \mathbb{R}^d \times \Omega \to \mathbb{R}^d$, $\theta \in \Theta$, be independent continuous random fields which satisfy for all $r, s, t \in [0,T]$, $x \in \mathbb{R}^d$, $\theta, \vartheta \in \Theta$, $B \in \mathcal{B}(\mathbb{R}^d)$ with $t \leq s \leq r$ and $\theta \neq \vartheta$ that $\mathbb{P}(X^\theta_{t,t}(x) = x) = 1$ and

$$\mathbb{P}\big(X^\theta_{s,r}(X^\vartheta_{t,s}(x)) \in B\big) = \mathbb{P}\big(X^\theta_{t,r}(x) \in B\big), \tag{3.3}$$

assume that $(X^\theta)_{\theta \in \Theta}$ and $(\mathcal{R}^\theta)_{\theta \in \Theta}$ are independent, assume for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that $\mathbb{E}\big[|g(X^0_{t,T}(x))| + \int_t^T |f(r, X^0_{t,r}(x), u(r, X^0_{t,r}(x)))|\, dr\big] < \infty$ and

$$u(t,x) = \mathbb{E}\left[g\big(X^0_{t,T}(x)\big) + \int_t^T f\big(r, X^0_{t,r}(x), u(r, X^0_{t,r}(x))\big)\, dr\right], \tag{3.4}$$

and let $V^\theta_{M,n} \colon [0,T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}$, $M, n \in \mathbb{Z}$, $\theta \in \Theta$, be functions which satisfy for all $M, n \in \mathbb{N}$, $\theta \in \Theta$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that $V^\theta_{M,0}(t,x) = 0$ and

$$\begin{aligned} V^\theta_{M,n}(t,x) = {} & \frac{1}{M^n}\left[\sum_{m=1}^{M^n} g\big(X^{(\theta,n,-m)}_{t,T}(x)\big)\right] \\ & + \sum_{k=0}^{n-1} \frac{(T-t)}{M^{n-k}} \left[\sum_{m=1}^{M^{n-k}} f\Big(R^{(\theta,k,m)}_t, X^{(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x), V^{(\theta,k,m)}_{M,k}\big(R^{(\theta,k,m)}_t, X^{(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x)\big)\Big)\right. \\ & \left. - \mathbb{1}_{\mathbb{N}}(k) f\Big(R^{(\theta,k,m)}_t, X^{(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x), V^{(\theta,k,-m)}_{M,k-1}\big(R^{(\theta,k,m)}_t, X^{(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x)\big)\Big)\right]. \end{aligned} \tag{3.5}$$

3.2 A priori bounds for solutions of stochastic fixed point equations

In our L^2-error analysis (see Subsection 3.4 below) of the MLP scheme introduced in Setting 3.1 we need to estimate expectations involving the solution of the stochastic fixed point equation. This estimate is carried out in Lemma 3.3 below. In order to prove Lemma 3.3 we need the elementary and well-known time-reversed Gronwall inequality in Lemma 3.2.

Lemma 3.2 (Time-reversed time-continuous Gronwall inequality)**.** *Let $T, \alpha, \beta \in [0,\infty)$ and let $\epsilon \colon [0,T] \to [0,\infty]$ be a $\mathcal{B}([0,T])/\mathcal{B}([0,\infty])$-measurable function which satisfies for all $t \in [0,T]$ that $\int_0^T \epsilon(r)\, dr < \infty$ and $\epsilon(t) \leq \alpha + \beta \int_t^T \epsilon(r)\, dr$. Then*

(i) it holds for all $t \in [0,T]$ that $\epsilon(t) \leq \alpha \exp(\beta(T-t))$ and

(ii) it holds that $\sup_{t \in [0,T]} \epsilon(t) \leq \alpha \exp(\beta T) < \infty$.

Proof of Lemma 3.2. Throughout this proof let $\Phi \colon [0,T] \to [0,T]$ and $\varepsilon \colon [0,T] \to [0,\infty]$ be the functions which satisfy for all $t \in [0,T]$ that

$$\Phi(t) = T - t \qquad \text{and} \qquad \varepsilon(t) = \epsilon(\Phi(t)) = \epsilon(T-t). \tag{3.6}$$

Observe that the integral transformation theorem (see, e.g., Klenke [70, Theorem 4.10]) implies that for all $t \in [0,T]$ it holds that

$$\begin{aligned} \int_0^t \varepsilon(r)\,dr &= \int_{[0,t]} \epsilon(\Phi(r))\ \mathrm{Borel}_{[0,t]}(dr) = \int_{\Phi([0,t])} \epsilon(s)\,\Phi(\mathrm{Borel}_{[0,t]})_{\mathcal{B}(\Phi([0,t]))}(ds) \\ &= \int_{[T-t,T]} \epsilon(s)\ \mathrm{Borel}_{[T-t,T]}(ds) = \int_{T-t}^T \epsilon(s)\,ds. \end{aligned} \tag{3.7}$$

Hence, we obtain that

$$\int_0^T \varepsilon(r)\,dr = \int_0^T \epsilon(r)\,dr < \infty. \tag{3.8}$$

Moreover, observe that (3.6), (3.7), and the hypothesis that for all $t \in [0,T]$ it holds that $\epsilon(t) \leq \alpha + \beta \int_t^T \epsilon(r)\,dr$ assure that for all $t \in [0,T]$ it holds that

$$\varepsilon(t) = \epsilon(T-t) \leq \alpha + \beta \int_{T-t}^T \epsilon(r)\,dr = \alpha + \beta \int_0^t \varepsilon(r)\,dr. \tag{3.9}$$

Combining this and (3.8) with Gronwall's integral inequality (cf, e.g., Grohs et al. [52, Lemma 2.11]) demonstrates that for all $t \in [0,T]$ it holds that

$$\varepsilon(t) \leq \alpha \exp(\beta t). \tag{3.10}$$

Hence, we obtain that for all $t \in [0,T]$ it holds that

$$\epsilon(t) = \epsilon(T-(T-t)) = \varepsilon(T-t) \leq \alpha \exp(\beta(T-t)) \leq \alpha \exp(\beta T). \tag{3.11}$$

This establishes items (i)–(ii). The proof of Lemma 3.2 is thus completed. □

Lemma 3.3. *Assume Setting 3.1, let $\xi \in \mathbb{R}^d$, $C \in [0,\infty]$ satisfy that*

$$C = \left(\mathbb{E}\big[|g(X^0_{0,T}(\xi))|^2\big]\right)^{1/2} + \sqrt{T}\left(\int_0^T \mathbb{E}\big[|f(t, X^0_{0,t}(\xi), 0)|^2\big]\ dt\right)^{1/2}, \tag{3.12}$$

and assume that $\int_0^T \left(\mathbb{E}\big[|u(t, X^0_{0,t}(\xi))|^2\big]\right)^{1/2} dt < \infty$. Then

(i) it holds for all $t \in [0,T]$ that $\left(\mathbb{E}\big[|u(t, X^0_{0,t}(\xi))|^2\big]\right)^{1/2} \leq C \exp(L(T-t))$ and

(ii) it holds that $\sup_{t\in[0,T]} \left(\mathbb{E}\big[|u(t, X^0_{0,t}(\xi))|^2\big]\right)^{1/2} \leq C \exp(LT)$.

Proof of Lemma 3.3. Throughout this proof assume w.l.o.g. that $C < \infty$ and let $\mu_t \colon \mathcal{B}(\mathbb{R}^d) \to [0,1]$, $t \in [0,T]$, be the probability measures which satisfy for all $t \in [0,T]$, $B \in \mathcal{B}(\mathbb{R}^d)$ that

$$\mu_t(B) = \mathbb{P}(X^0_{0,t}(\xi) \in B) = \mathbb{P}(X^1_{0,t}(\xi) \in B) = \big((X^1_{0,t}(\xi))(\mathbb{P})_{\mathcal{B}(\mathbb{R}^d)}\big)(B) \tag{3.13}$$

(cf. (3.3) or item (iv) in Lemma 3.6). Note that (3.4) and the triangle inequality ensure that for all $t \in [0,T]$ it holds that

$$\begin{aligned} &\left(\mathbb{E}\big[|u(t, X^0_{0,t}(\xi))|^2\big]\right)^{1/2} = \left(\int_{\mathbb{R}^d} |u(t,z)|^2\,\mu_t(dz)\right)^{1/2} \\ &= \left(\int_{\mathbb{R}^d} \left|\mathbb{E}\Big[g\big(X^0_{t,T}(z)\big) + \textstyle\int_t^T f\big(r, X^0_{t,r}(z), u(r, X^0_{t,r}(z))\big)\,dr\Big]\right|^2 \mu_t(dz)\right)^{1/2} \\ &\leq \left(\int_{\mathbb{R}^d} \left|\mathbb{E}\big[g\big(X^0_{t,T}(z)\big)\big]\right|^2 \mu_t(dz)\right)^{1/2} \\ &\quad + \left(\int_{\mathbb{R}^d} \left|\mathbb{E}\Big[\textstyle\int_t^T f\big(r, X^0_{t,r}(z), u(r, X^0_{t,r}(z))\big)\,dr\Big]\right|^2 \mu_t(dz)\right)^{1/2}. \end{aligned} \tag{3.14}$$

Jensen's inequality hence assures that for all $t \in [0,T]$ it holds that

$$\begin{aligned}\big(\mathbb{E}\big[|u(t,X^0_{0,t}(\xi))|^2\big]\big)^{1/2} &\leq \left(\int_{\mathbb{R}^d} \mathbb{E}\Big[\big|g\big(X^0_{t,T}(z)\big)\big|^2\Big]\; \mu_t(dz)\right)^{1/2} \\ &\quad + \left(\int_{\mathbb{R}^d} \mathbb{E}\bigg[\Big(\textstyle\int_t^T \big|f\big(r,X^0_{t,r}(z),u(r,X^0_{t,r}(z))\big)\big|\,dr\Big)^2\bigg]\; \mu_t(dz)\right)^{1/2}.\end{aligned} \tag{3.15}$$

Furthermore, observe that (3.13), the fact that X^0 and X^1 are independent and continuous random fields, (3.3), and Lemma 2.16 demonstrate that for all $t \in [0,T]$ it holds that

$$\left(\int_{\mathbb{R}^d} \mathbb{E}\Big[\big|g\big(X^0_{t,T}(z)\big)\big|^2\Big]\; \mu_t(dz)\right)^{1/2} = \Big(\mathbb{E}\Big[\big|g\big(X^0_{t,T}(X^1_{0,t}(\xi))\big)\big|^2\Big]\Big)^{1/2} = \Big(\mathbb{E}\Big[\big|g\big(X^0_{0,T}(\xi)\big)\big|^2\Big]\Big)^{1/2}. \tag{3.16}$$

In addition, note that Minkowski's integral inequality (cf., e.g., Jentzen & Kloeden [65, Proposition 8 in Appendix A.1]), (3.13), the fact that X^0 and X^1 are independent and continuous random fields, (3.3), and Lemma 2.16 imply that for all $t \in [0,T]$ it holds that

$$\begin{aligned}&\left(\int_{\mathbb{R}^d} \mathbb{E}\bigg[\Big(\textstyle\int_t^T \big|f\big(r,X^0_{t,r}(z),u(r,X^0_{t,r}(z))\big)\big|\,dr\Big)^2\bigg]\; \mu_t(dz)\right)^{1/2} \\ &\leq \int_t^T \left(\int_{\mathbb{R}^d} \mathbb{E}\Big[\big|f\big(r,X^0_{t,r}(z),u(r,X^0_{t,r}(z))\big)\big|^2\Big]\; \mu_t(dz)\right)^{1/2} dr \\ &= \int_t^T \Big(\mathbb{E}\Big[\big|f\big(r,X^0_{t,r}(X^1_{0,t}(\xi)),u(r,X^0_{t,r}(X^1_{0,t}(\xi)))\big)\big|^2\Big]\Big)^{1/2} dr \\ &= \int_t^T \Big(\mathbb{E}\Big[\big|f\big(r,X^0_{0,r}(\xi),u(r,X^0_{0,r}(\xi))\big)\big|^2\Big]\Big)^{1/2} dr.\end{aligned} \tag{3.17}$$

Moreover, observe that (3.1) ensures that for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $v \in \mathbb{R}$ it holds that

$$|f(t,x,v)| \leq |f(t,x,0)| + |f(t,x,v)-f(t,x,0)| \leq |f(t,x,0)| + L|v|. \tag{3.18}$$

This, (3.17), and the triangle inequality imply that for all $t \in [0,T]$ it holds that

$$\begin{aligned}&\left(\int_{\mathbb{R}^d} \mathbb{E}\bigg[\Big(\textstyle\int_t^T \big|f\big(r,X^0_{t,r}(z),u(r,X^0_{t,r}(z))\big)\big|\,dr\Big)^2\bigg]\; \mu_t(dz)\right)^{1/2} \\ &\leq \int_t^T \Big(\mathbb{E}\Big[\big|f\big(r,X^0_{0,r}(\xi),0\big)\big|^2\Big]\Big)^{1/2} dr + L\int_t^T \big(\mathbb{E}\big[|u(r,X^0_{0,r}(\xi))|^2\big]\big)^{1/2} dr.\end{aligned} \tag{3.19}$$

Furthermore, note that Lemma 2.10 assures that for all $t \in [0,T]$ it holds that

$$\begin{aligned}\int_t^T \Big(\mathbb{E}\Big[\big|f\big(r,X^0_{0,r}(\xi),0\big)\big|^2\Big]\Big)^{1/2} dr &= \left(\left[\int_t^T \Big(\mathbb{E}\Big[\big|f\big(r,X^0_{0,r}(\xi),0\big)\big|^2\Big]\Big)^{1/2} dr\right]^2\right)^{1/2} \\ &\leq \left((T-t)\int_t^T \mathbb{E}\Big[\big|f\big(r,X^0_{0,r}(\xi),0\big)\big|^2\Big]\, dr\right)^{1/2} \\ &\leq \sqrt{T}\left(\int_0^T \mathbb{E}\Big[\big|f\big(r,X^0_{0,r}(\xi),0\big)\big|^2\Big]\, dr\right)^{1/2}.\end{aligned} \tag{3.20}$$

Combining this with (3.12), (3.15), (3.16), and (3.19) implies that for all $t \in [0,T]$ it holds that

$$\begin{aligned}
&\big(\mathbb{E}\big[|u(t, X^0_{0,t}(\xi))|^2\big]\big)^{1/2} \\
&\leq \Big(\mathbb{E}\Big[\big|g\big(X^0_{0,T}(\xi)\big)\big|^2\Big]\Big)^{1/2} + \sqrt{T}\left(\int_0^T \mathbb{E}\Big[\big|f\big(r, X^0_{0,r}(\xi), 0\big)\big|^2\Big]\, dr\right)^{1/2} \\
&\quad + L\int_t^T \big(\mathbb{E}\big[|u(r, X^0_{0,r}(\xi))|^2\big]\big)^{1/2}\, dr. \\
&= C + L\int_t^T \big(\mathbb{E}\big[|u(r, X^0_{0,r}(\xi))|^2\big]\big)^{1/2}\, dr.
\end{aligned} \tag{3.21}$$

The hypothesis that $\int_0^T \big(\mathbb{E}\big[|u(t, X^0_{0,t}(\xi))|^2\big]\big)^{1/2}\, dt < \infty$ and Lemma 3.2 (with $T = T$, $\alpha = C$, $\beta = L$, $(\epsilon(t))_{t\in[0,T]} = \big((\mathbb{E}[|u(t, X^0_{0,t}(\xi))|^2])^{1/2}\big)_{t\in[0,T]}$ in the notation of Lemma 3.2) hence establish items (i)–(ii). The proof of Lemma 3.3 is thus completed. □

3.3 Properties of MLP approximations

In this subsection we establish in Lemma 3.6 below some elementary properties of the MLP approximations (cf. (3.5) in Setting 3.1 above) introduced in Setting 3.1 above. For this we need two elementary and well-known results on identically distributed random variables (see Lemma 3.4 and Lemma 3.5 below).

Lemma 3.4. *Let $d, N \in \mathbb{N}$, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $X_k \colon \Omega \to \mathbb{R}^d$, $k \in \{1, 2, \dots, N\}$, be independent random variables, let $Y_k \colon \Omega \to \mathbb{R}^d$, $k \in \{1, 2, \dots, N\}$, be independent random variables, and assume for every $k \in \{1, 2, \dots, N\}$ that X_k and Y_k are identically distributed. Then it holds that $\big(\sum_{k=1}^N X_k\big) \colon \Omega \to \mathbb{R}^d$ and $\big(\sum_{k=1}^N Y_k\big) \colon \Omega \to \mathbb{R}^d$ are identically distributed random variables.*

Proof of Lemma 3.4. Throughout this proof let $\mathfrak{X}, \mathfrak{Y} \colon \Omega \to \mathbb{R}^{Nd}$ be the random variables which satisfy that

$$\mathfrak{X} = (X_1, \dots, X_N) \qquad \text{and} \qquad \mathfrak{Y} = (Y_1, \dots, Y_N) \tag{3.22}$$

and let $f \in C(\mathbb{R}^{Nd}, \mathbb{R}^d)$ be the function which satisfies for all $v_1, v_2, \dots, v_N \in \mathbb{R}^d$ that $f(v_1, v_2, \dots, v_N) = \sum_{k=1}^N v_k$. Observe that the hypothesis that $(X_k)_{k\in\{1,2,\dots,N\}}$ are independent, the hypothesis that $(Y_k)_{k\in\{1,2,\dots,N\}}$ are independent, and the hypothesis that for every $k \in \{1, 2, \dots, N\}$ it holds that X_k and Y_k are identically distributed random variables assure that for all $(B_k)_{k\in\{1,2,\dots,N\}} \subseteq \mathcal{B}(\mathbb{R}^d)$ it holds that

$$\begin{aligned}
\mathbb{P}\big(\mathfrak{X} \in (B_1 \times B_2 \times \ldots \times B_N)\big) &= \mathbb{P}(\forall\, k \in \{1, 2, \dots, N\} \colon X_k \in B_k) \\
&= \prod_{k=1}^N \mathbb{P}(X_k \in B_k) = \prod_{k=1}^N \mathbb{P}(Y_k \in B_k) \\
&= \mathbb{P}(\forall\, k \in \{1, 2, \dots, N\} \colon Y_k \in B_k) \\
&= \mathbb{P}\big(\mathfrak{Y} \in (B_1 \times B_2 \times \ldots \times B_N)\big).
\end{aligned} \tag{3.23}$$

This, the fact that

$$\mathcal{B}(\mathbb{R}^{Nd}) = \mathfrak{S}\Big((B_1 \times B_2 \times \ldots \times B_N) \in \mathcal{P}(\mathbb{R}^{LNd}) \colon \big(\forall\, k \in \{1, 2, \dots, N\} \colon B_k \in \mathcal{B}(\mathbb{R}^d)\big)\Big), \tag{3.24}$$

and the uniqueness theorem for measures (see, e.g., Klenke [70, Lemma 1.42]) imply that it holds for all $B \in \mathcal{B}(\mathbb{R}^{Nd})$ that

$$\mathbb{P}\big(\mathfrak{X} \in B\big) = \mathbb{P}\big(\mathfrak{Y} \in B\big). \tag{3.25}$$

Hence, we obtain that for all $B \in \mathcal{B}(\mathbb{R}^d)$ it holds that

$$\begin{aligned} \mathbb{P}\left(\textstyle\sum_{k=1}^{N} X_k \in B\right) &= \mathbb{P}\left(f(\mathfrak{X}) \in B\right) = \mathbb{P}\left(\mathfrak{X} \in f^{-1}(B)\right) \\ &= \mathbb{P}\left(\mathfrak{Y} \in f^{-1}(B)\right) = \mathbb{P}\left(f(\mathfrak{Y}) \in B\right) = \mathbb{P}\left(\textstyle\sum_{k=1}^{N} Y_k \in B\right). \end{aligned} \tag{3.26}$$

This shows that $\left(\sum_{k=1}^{N} X_k\right) : \Omega \to \mathbb{R}^d$ and $\left(\sum_{k=1}^{N} Y_k\right) : \Omega \to \mathbb{R}^d$ are identically distributed random variables. The proof of Lemma 3.4 is thus completed. □

Lemma 3.5. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let (S, δ) be a separable metric space, let (E, δ) be a metric space, let $U, V : S \times \Omega \to E$ be continuous random fields, let $X, Y : \Omega \to S$ be random variables, assume that U and X are independent, assume that V and Y are independent, assume for all $s \in S$ that $U(s)$ and $V(s)$ are identically distributed, and assume that X and Y are identically distributed. Then it holds that $U(X) = (U(X(\omega), \omega))_{\omega \in \Omega} : \Omega \to E$ and $V(Y) = (V(Y(\omega), \omega))_{\omega \in \Omega} : \Omega \to E$ are identically distributed random variables.*

Proof of Lemma 3.5. First, note that Grohs et al. [3, Lemma 2.4], the fact that U and V are continuous random fields, and Lemma 2.15 ensure that $U(X)$ and $V(Y)$ are random variables. Next observe the hypothesis that U and X are independent, the hypothesis that V and Y are independent, the hypothesis that for all $s \in S$ it holds that $U(s)$ and $V(s)$ are identically distributed, the hypothesis that X and Y are identically distributed and Lemma 2.17 demonstrate that for all globally bounded and Lipschitz continuous functions $g : E \to \mathbb{R}$ it holds that

$$\begin{aligned} \mathbb{E}[g(U(X))] &= \int_S \mathbb{E}[g(U(s))] \, (X(\mathbb{P})_{\mathcal{B}(S)})(ds) \\ &= \int_S \mathbb{E}[g(V(s))] \, (Y(\mathbb{P})_{\mathcal{B}(S)})(ds) = \mathbb{E}[g(V(Y))]. \end{aligned} \tag{3.27}$$

Combining this with Lemma 2.14 assures that $U(X)$ and $V(Y)$ are identically distributed. The proof of Lemma 3.5 is thus completed. □

Lemma 3.6 (Properties of MLP approximations)**.** *Assume Setting 3.1 and let $M \in \mathbb{N}$. Then*

(i) for all $\theta \in \Theta$, $n \in \mathbb{N}_0$ it holds that $V^{\theta}_{M,n} : [0, T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}$ is a continuous random field,

(ii) for all $\theta \in \Theta$, $n \in \mathbb{N}_0$ it holds that $V^{\theta}_{M,n} : [0, T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}$ is $(\mathcal{B}([0, T] \times \mathbb{R}^d) \otimes \mathfrak{S}((\mathcal{R}^{(\theta,\vartheta)})_{\vartheta \in \Theta}, (X^{(\theta,\vartheta)})_{\vartheta \in \Theta}))/\mathcal{B}(\mathbb{R})$-measurable,

(iii) for all $\theta \in \Theta$, $n \in \mathbb{N}_0$, $t \in [0, T]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned} &\big((\mathbb{N} \cap [0, n]) \times \mathbb{N}\big) \ni (k, m) \mapsto \\ &\begin{cases} g(X^{(\theta,n,-m)}_{t,T}(x)) & : k = n \\ \Big[f\big(R^{(\theta,k,m)}_t, X^{(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x), V^{(\theta,k,m)}_{M,k}\big(R^{(\theta,k,m)}_t, X^{(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x)\big)\big) & \\ \quad - \mathbb{1}_{\mathbb{N}}(k) f\big(R^{(\theta,k,m)}_t, X^{(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x), V^{(\theta,k,-m)}_{M,k-1}\big(R^{(\theta,k,m)}_t, X^{(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x)\big)\big)\Big] & : k < n \end{cases} \end{aligned} \tag{3.28}$$

is an independent family of random variables,

(iv) for all $t \in [0, T]$, $s \in [t, T]$, $x \in \mathbb{R}^d$ it holds that $X^{\theta}_{t,s}(x) : \Omega \to \mathbb{R}^d$, $\theta \in \Theta$, are identically distributed random variables, and

(v) *for all* $n \in \mathbb{N}_0$*,* $t \in [0,T]$*,* $x \in \mathbb{R}^d$ *it holds that* $V^\theta_{M,n}(t,x)\colon \Omega \to \mathbb{R}^d$*,* $\theta \in \Theta$*, are identically distributed random variables.*

Proof of Lemma 3.6. We first prove item (i) by induction on $n \in \mathbb{N}_0$. For the base case $n = 0$ observe that the hypothesis that for all $\theta \in \Theta$ it holds that $V^\theta_{M,0} = 0$ demonstrates that for all $\theta \in \Theta$ it holds that $V^\theta_{M,0}\colon [0,T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}^d$ is a continuous random field. This establishes item (i) in the base case $n = 0$. For the induction step $\mathbb{N}_0 \ni (n-1) \to n \in \mathbb{N}$ let $n \in \mathbb{N}$ and assume that for every $k \in \mathbb{N}_0 \cap [0,n)$, $\theta \in \Theta$ it holds that $V^\theta_{M,k}\colon [0,T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}^d$ is a continuous random field. Combining this, the hypothesis that g and f are continuous functions, and the fact that for all $\theta \in \Theta$ it holds that $R^\theta\colon [0,T] \times \Omega \to [0,T]$ and $X^\theta\colon \{(t,s) \in [0,T]^2\colon t \leq s\} \times \mathbb{R}^d \times \Omega \to \mathbb{R}^d$ are continuous random fields with (3.5), Grohs et al. [3, Lemma 2.4], and Lemma 2.15 proves that for all $\theta \in \Theta$ it holds that $V^\theta_{M,n}\colon [0,T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}^d$ is a continuous random field. Induction thus establishes item (i). Next we prove item (ii) by induction on $n \in \mathbb{N}_0$. For the base case $n = 0$ observe that the hypothesis that for all $\theta \in \Theta$ it holds that $V^\theta_{M,0} = 0$ demonstrates that for all $\theta \in \Theta$ it holds that $V^\theta_{M,0}\colon [0,T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}$ is $\big(\mathcal{B}([0,T] \times \mathbb{R}^d) \otimes \mathfrak{S}((\mathcal{R}^{(\theta,\vartheta)})_{\vartheta\in\Theta}, (X^{(\theta,\vartheta)})_{\vartheta\in\Theta})\big)\,/\mathcal{B}(\mathbb{R})$-measurable. This implies item (ii) in the base case $n = 0$. For the induction step $\mathbb{N}_0 \ni (n-1) \to n \in \mathbb{N}$ let $n \in \mathbb{N}$ and assume that for all $k \in \mathbb{N}_0 \cap [0,n)$, $\theta \in \Theta$ it holds that $V^\theta_{M,k}$ is $\big(\mathcal{B}([0,T] \times \mathbb{R}^d) \otimes \mathfrak{S}((\mathcal{R}^{(\theta,\vartheta)})_{\vartheta\in\Theta}, (X^{(\theta,\vartheta)})_{\vartheta\in\Theta})\big)\,/\mathcal{B}(\mathbb{R})$-measurable. Combining this, the fact that f and g are Borel measurable, and the fact that for all $\theta \in \Theta$ it holds that $X^\theta\colon \{(t,s) \in [0,T]^2\colon t \leq s\} \times \mathbb{R}^d \times \Omega \to \mathbb{R}^d$ is a continuous random field with (3.5) and Lemma 2.15 proves that for all $\theta \in \Theta$, $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that

$$
\begin{aligned}
&\mathfrak{S}(V^\theta_{M,n}(t,x)) \\
&\subseteq \mathfrak{S}\Big((X^{(\theta,n,-m)}_{t,T}(x))_{m\in\{1,2,\ldots,M^n\}}, (R^{(\theta,k,m)}_t)_{m\in\{1,2,\ldots,M^{n-k}\},k\in\mathbb{N}_0\cap[0,n)}, \\
&\quad (X^{(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x))_{m\in\{1,2,\ldots,M^{n-k}\},k\in\mathbb{N}_0\cap[0,n)}, (X^{(\theta,k,m,\vartheta)})_{m\in\{1,2,\ldots,M^{n-k}\},k\in\mathbb{N}_0\cap[0,n),\vartheta\in\Theta}, \\
&\quad (\mathcal{R}^{(\theta,k,m,\vartheta)})_{m\in\{1,2,\ldots,M^{n-k}\},k\in\mathbb{N}_0\cap[0,n),\vartheta\in\Theta}, (X^{(\theta,k,-m,\vartheta)})_{m\in\{1,2,\ldots,M^{n-k}\},k\in\mathbb{N}\cap[0,n),\vartheta\in\Theta}, \\
&\quad (\mathcal{R}^{(\theta,k,-m,\vartheta)})_{m\in\{1,2,\ldots,M^{n-k}\},k\in\mathbb{N}\cap[0,n),\vartheta\in\Theta}\Big) \\
&\subseteq \mathfrak{S}\left((\mathcal{R}^{(\theta,\vartheta)})_{\vartheta\in\Theta}, (X^{(\theta,\vartheta)})_{\vartheta\in\Theta}\right).
\end{aligned}
\tag{3.29}
$$

Moreover, observe that item (i) and Grohs et al. [3, Lemma 2.4] ensure that for all $\theta \in \Theta$ it holds that $V^\theta_{M,n}$ is $\big(\mathcal{B}([0,T] \times \mathbb{R}^d) \otimes \mathfrak{S}(V^\theta_{M,n})\big)/\mathcal{B}(\mathbb{R})$-measurable. Combining this with (3.29) demonstrates that for all $\theta \in \Theta$ it holds that $V^\theta_{M,n}$ is $(\mathcal{B}([0,T] \times \mathbb{R}^d) \otimes \mathfrak{S}((\mathcal{R}^{(\theta,\vartheta)})_{\vartheta\in\Theta}, (X^{(\theta,\vartheta)})_{\vartheta\in\Theta}))/\mathcal{B}(\mathbb{R})$-measurable. Induction thus establishes item (ii). Furthermore, observe that item (ii), the hypothesis that $(X^\theta)_{\theta\in\Theta}$ are independent, the hypothesis that $(\mathcal{R}^\theta)_{\theta\in\Theta}$ are independent, the hypothesis that $(X^\theta)_{\theta\in\Theta}$ and $(\mathcal{R}^\theta)_{\theta\in\Theta}$ are independent, and Lemma 2.15 prove item (iii). Next observe that (3.3), the hypothesis that $(X^\theta)_{\theta\in\Theta}$ are independent, Lemma 2.17 (with $S = \mathbb{R}^d$, $U = g(X^\theta_{s,s}(\cdot))$, $X = X^\vartheta_{t,s}(x)$ for $g \in C(\mathbb{R}^d, \mathbb{R})$, $t \in [0,T]$, $s \in [t,T]$, $x \in \mathbb{R}^d$, $\theta, \vartheta \in \Theta$ in the notation of Lemma 2.17), and the fact that for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $\theta \in \Theta$ it holds that $\mathbb{P}(X^\theta_{t,t}(x) = x) = 1$ assure that for all $t \in [0,T]$, $s \in [t,T]$, $x \in \mathbb{R}^d$, $\theta, \vartheta \in \Theta$ with $\theta \neq \vartheta$ and all globally bounded and continuous functions $g\colon \mathbb{R}^d \to \mathbb{R}$ it holds that

$$
\begin{aligned}
\mathbb{E}\big[g(X^\theta_{t,s}(x))\big] &= \mathbb{E}\big[g(X^\theta_{s,s}(X^\vartheta_{t,s}(x)))\big] = \int_{\mathbb{R}^d} \mathbb{E}\big[g(X^\theta_{s,s}(z))\big]\,((X^\vartheta_{t,s}(x))(\mathbb{P})_{\mathcal{B}(\mathbb{R}^d)})(dz) \\
&= \int_{\mathbb{R}^d} g(z)((X^\vartheta_{t,s}(x))(\mathbb{P})_{\mathcal{B}(\mathbb{R}^d)})(dz) = \mathbb{E}\big[g(X^\vartheta_{t,s}(x))\big]\,.
\end{aligned}
\tag{3.30}
$$

Combining this with Lemma 2.14 demonstrates that for all $t \in [0,T]$, $s \in [t,T]$, $x \in \mathbb{R}^d$, $\theta, \vartheta \in \Theta$ it holds that $X^{\theta}_{t,s}(x) \colon \Omega \to \mathbb{R}^d$ and $X^{\vartheta}_{t,s}(x) \colon \Omega \to \mathbb{R}^d$ are identically distributed random variables. This establishes item (iv). Next we prove item (v) by induction on $n \in \mathbb{N}_0$. For the base case $n = 0$ observe that the hypothesis that for all $\theta \in \Theta$ it holds that $V^{\theta}_{M,0} = 0$ demonstrates that for all $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that $V^{\theta}_{M,0}(t,x) \colon \Omega \to \mathbb{R}^d$, $\theta \in \Theta$, are identically distributed random variables. This establishes item (v) in the base case $n = 0$. For the induction step $\mathbb{N}_0 \ni (n-1) \to n \in \mathbb{N}$ let $n \in \mathbb{N}$ and assume that for all $k \in \mathbb{N}_0 \cap [0,n)$, $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that $V^{\theta}_{M,k}(t,x) \colon \Omega \to \mathbb{R}^d$, $\theta \in \Theta$, are identically distributed random variables. This, the hypothesis that $(X^{\theta})_{\theta \in \Theta}$ are independent, the hypothesis that $(R^{\theta})_{\theta \in \Theta}$ are independent, the hypothesis that $(X^{\theta})_{\theta \in \Theta}$ and $(\mathcal{R}^{\theta})_{\theta \in \Theta}$ are independent, item (ii), Lemma 3.4, and Lemma 3.5 (with $S = [0,T] \times \mathbb{R}^d$, $E = \mathbb{R}$, $U = \big(f(s,y,V^{(\theta,k,m)}_{M,k}(s,y)) - \mathbb{1}_{\mathbb{N}}(k) f(s,y,V^{(\theta,k,-m)}_{M,k-1}(s,y))\big)_{(s,y) \in [0,T] \times \mathbb{R}^d}$, $V = \big(f(s,y,V^{(\vartheta,k,m)}_{M,k}(s,y)) - \mathbb{1}_{\mathbb{N}}(k) f(s,y,V^{(\vartheta,k,-m)}_{M,k-1}(s,y))\big)_{(s,y) \in [0,T] \times \mathbb{R}^d}$, $X = (R^{(\theta,k,m)}_t, X^{(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x))$, $Y = (R^{(\vartheta,k,m)}_t, X^{(\vartheta,k,m)}_{t,R^{(\vartheta,k,m)}_t}(x))$ for $\theta, \vartheta \in \Theta$, $t \in [0,T]$, $x \in \mathbb{R}^d$, $k \in \mathbb{N}_0 \cap [0,n)$, $m \in \mathbb{N}$ with $\theta \neq \vartheta$ in the notation of Lemma 3.5) assure that for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $k \in \mathbb{N}_0 \cap [0,n)$, $m \in \mathbb{N}$ it holds that

$$\begin{aligned} &\Big(f\big(R^{(\theta,k,m)}_t, X^{(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x), V^{(\theta,k,m)}_{M,k}\big(R^{(\theta,k,m)}_t, X^{(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x)\big)\big) \\ &- \mathbb{1}_{\mathbb{N}}(k) f\big(R^{(\theta,k,m)}_t, X^{(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x), V^{(\theta,k,-m)}_{M,k-1}\big(R^{(\theta,k,m)}_t, X^{(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x)\big)\big)\Big)_{\theta \in \Theta} \end{aligned} \tag{3.31}$$

are identically distributed random variables. Items (iii)–(iv), (3.5), and Lemma 3.4 therefore ensure that for all $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that $V^{\theta}_{M,n}(t,x) \colon \Omega \to \mathbb{R}^d$, $\theta \in \Theta$, are identically distributed random variables. Induction thus establishes item (v). The proof of Lemma 3.6 is thus completed. ☐

3.4 Analysis of approximation errors of MLP approximations

Proposition 3.15 and Corollary 3.16 in Subsection 3.4.5 below present estimates for the L^2-approximation error of the MLP scheme (cf. (3.5) in Setting 3.1 above) introduced in Setting 3.1 with respect to the solution of the stochastic fixed point equation (cf. (3.4) in Setting 3.1 above) for every iteration (cf. $n \in \mathbb{N}$ in (3.5) in Subsection 3.1 above) and every Monte Carlo accuracy (cf. $M \in \mathbb{N}$ in (3.5) in Subsection 3.1 above) of the MLP scheme. The essential idea for the proof of those statements is to decompose the L^2-approximation error into a bias and a variance part and to analyze them separately (see Subsections 3.4.1–3.4.3). This approach leads to a recursive inequality (cf. (3.89) in the proof of Proposition 3.15 below) which can be treated using an elementary Gronwall inequality, proven in Subsection 3.4.4 (see Lemma 3.12). For the proofs of the statements in this subsection we need some elementary and well-known results (see Lemma 3.7, Lemma 3.10, and Lemma 3.14) which we state and prove where they are used.

3.4.1 Expectations of MLP approximations

Lemma 3.7. *Assume Setting 3.1, let $\theta \in \Theta$, $t \in [0,T]$, let $U_1 \colon [t,T] \times \Omega \to [0,\infty]$ and $U_2 \colon [t,T] \times \Omega \to \mathbb{R}$ be continuous random fields which satisfy for all $i \in \{1,2\}$ that U_i and $\mathcal{R}^{\theta}$ are independent and $\int_t^T \mathbb{E}[|U_2(r)|]\,dr < \infty$. Then it holds for all $i \in \{1,2\}$ that $\mathrm{Borel}_{[t,T]}(\{r \in [t,T] \colon \mathbb{E}[|U_2(r)|] = \infty\}) = 0$, $\mathbb{E}\big[|U_2(R^{\theta}_t)|\big] < \infty$, and*

$$(T-t)\,\mathbb{E}\big[U_i(R^{\theta}_t)\big] = \int_t^T \mathbb{E}[U_i(r)]\,dr. \tag{3.32}$$

Proof of Lemma 3.7. Throughout this proof assume w.l.o.g. that $t < T$. Observe that (3.2) implies that R^{θ}_t is $\mathcal{U}_{[t,T]}$-distributed. Combining this with the fact that U_1 is continuous, the fact that U_1 and R_t^θ are independent, and Lemma 2.16 assures that

$$\begin{aligned}(T-t)\mathbb{E}\big[U_1(R_t^\theta)\big] &= (T-t)\int_{[t,T]} \mathbb{E}[U_1(r)]\,(R_t^\theta(\mathbb{P})_{\mathcal{B}([t,T])})(dr)\\ &= (T-t)\int_{[t,T]} \mathbb{E}[U_1(r)]\,(\mathcal{U}_{[t,T]})(dr)\\ &= \frac{(T-t)}{(T-t)}\int_t^T \mathbb{E}[U_1(r)]\,dr = \int_t^T \mathbb{E}[U_1(r)]\,dr.\end{aligned} \tag{3.33}$$

In addition, note that the fact that R_t^θ is $\mathcal{U}_{[t,T]}$-distributed, the fact that U_2 is continuous, the fact that U_2 and R_t^θ are independent, the hypothesis that $\int_t^T \mathbb{E}[|U_2(r)|]\,dr < \infty$, and Lemma 2.17 ensure that $\mathrm{Borel}_{[t,T]}(\{r \in [t,T]\colon \mathbb{E}[|U_2(r)|] = \infty\}) = 0$, $\mathbb{E}\big[|U_2(R_t^\theta)|\big] < \infty$, and

$$\begin{aligned}(T-t)\mathbb{E}\big[U_2(R_t^\theta)\big] &= (T-t)\int_{[t,T]} \mathbb{E}[U_2(r)]\,(R_t^\theta(\mathbb{P})_{\mathcal{B}([t,T])})(dr)\\ &= (T-t)\int_{[t,T]} \mathbb{E}[U_2(r)]\,(\mathcal{U}_{[t,T]})(dr)\\ &= \frac{(T-t)}{(T-t)}\int_t^T \mathbb{E}[U_2(r)]\,dr = \int_t^T \mathbb{E}[U_2(r)]\,dr.\end{aligned} \tag{3.34}$$

Combining this with (3.33) establishes (3.32). The proof of Lemma 3.7 is thus completed. □

Lemma 3.8 (Expectations of MLP approximations)**.** *Assume Setting 3.1 and assume for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that $\int_t^T \mathbb{E}\big[|f(r, X^0_{t,r}(x), 0)|\big]\,dr < \infty$. Then*

(i) for all $M \in \mathbb{N}$, $n \in \mathbb{N}_0$, $t \in [0,T]$, $s \in [t,T]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}&\mathbb{E}\big[|V^0_{M,n}(s, X^0_{t,s}(x))|\big] + (T-t)\mathbb{E}\Big[|V^0_{M,n}(R^0_t, X^0_{t,R^0_t}(x))|\Big]\\ &\quad + (T-t)\mathbb{E}\Big[|f\big(R^0_t, X^0_{t,R^0_t}(x), V^0_{M,n}(R^0_t, X^0_{t,R^0_t}(x))\big)|\Big]\\ &= \mathbb{E}\big[|V^0_{M,n}(s, X^0_{t,s}(x))|\big] + \int_t^T \mathbb{E}\big[|V^0_{M,n}(r, X^0_{t,r}(x))|\big]\,dr\\ &\quad + \int_t^T \mathbb{E}\big[|f\big(r, X^0_{t,r}(x), V^0_{M,n}(r, X^0_{t,r}(x))\big)|\big]\,dr < \infty\end{aligned} \tag{3.35}$$

and

(ii) for all $M, n \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that

$$\mathbb{E}\big[V^0_{M,n}(t,x)\big] = \mathbb{E}\left[g(X^0_{t,T}(x)) + \int_t^T f\big(r, X^0_{t,r}(x), V^0_{M,n-1}(r, X^0_{t,r}(x))\big)\,dr\right]. \tag{3.36}$$

Proof of Lemma 3.8. Throughout this proof let $M \in \mathbb{N}$, $x \in \mathbb{R}^d$. Observe that Lemma 3.7, items (i)–(ii) in Lemma 3.6, and the fact that for all $n \in \mathbb{N}$ it holds that $V^0_{M,n}$, X^0, and $\mathcal{R}^0$ are independent demonstrate that for all $n \in \mathbb{N}_0$, $t \in [0,T]$ it holds that

$$\begin{aligned}&(T-t)\mathbb{E}\Big[|V^0_{M,n}(R^0_t, X^0_{t,R^0_t}(x))|\Big] + (T-t)\mathbb{E}\Big[\big|f\big(R^0_t, X^0_{t,R^0_t}(x), V^0_{M,n}(R^0_t, X^0_{t,R^0_t}(x))\big)\big|\Big]\\ &= \int_t^T \mathbb{E}\big[|V^0_{M,n}(r, X^0_{t,r}(x))|\big]\,dr + \int_t^T \mathbb{E}\big[\big|f\big(r, X^0_{t,r}(x), V^0_{M,n}(r, X^0_{t,r}(x))\big)\big|\big]\,dr.\end{aligned} \tag{3.37}$$

Next we claim that for all $n \in \mathbb{N}_0$, $t \in [0,T]$, $s \in [t,T]$ it holds that

$$\begin{aligned} &\mathbb{E}\big[|V^0_{M,n}(s, X^0_{t,s}(x))|\big] + \int_t^T \mathbb{E}\big[|V^0_{M,n}(r, X^0_{t,r}(x))|\big]\, dr \\ &\quad + \int_t^T \mathbb{E}\big[|f(r, X^0_{t,r}(x), V^0_{M,n}(r, X^0_{t,r}(x)))|\big]\, dr < \infty. \end{aligned} \tag{3.38}$$

We now prove (3.38) by induction on $n \in \mathbb{N}_0$. For the base case $n = 0$ observe that the hypothesis that $V^0_{M,0} = 0$ and the hypothesis that for all $t \in [0,T]$ it holds that $\int_t^T \mathbb{E}\big[|f(r, X^0_{t,r}(x), 0)|\big]\ dr < \infty$ imply that for all $t \in [0,T]$, $s \in [t,T]$ it holds that

$$\begin{aligned} &\mathbb{E}\big[|V^0_{M,0}(s, X^0_{t,s}(x))|\big] + \int_t^T \mathbb{E}\big[|V^0_{M,0}(r, X^0_{t,r}(x))|\big]\, dr \\ &\quad + \int_t^T \mathbb{E}\big[\big|f\big(r, X^0_{t,r}(x), V^0_{M,0}(r, X^0_{t,r}(x))\big)\big|\big]\, dr = \int_t^T \mathbb{E}\big[|f(r, X^0_{t,r}(x), 0)|\big]\, dr < \infty. \end{aligned} \tag{3.39}$$

This establishes (3.38) in the case $n = 0$. For the induction step $\mathbb{N}_0 \ni (n-1) \to n \in \mathbb{N}$ let $n \in \mathbb{N}$ and assume that for all $k \in \mathbb{N}_0 \cap [0,n)$, $t \in [0,T]$, $s \in [t,T]$ it holds that

$$\begin{aligned} &\mathbb{E}\big[|V^0_{M,k}(s, X^0_{t,s}(x))|\big] + \int_t^T \mathbb{E}\big[|V^0_{M,k}(r, X^0_{t,r}(x))|\big]\, dr \\ &\quad + \int_t^T \mathbb{E}\big[\big|f\big(r, X^0_{t,r}(x), V^0_{M,k}(r, X^0_{t,r}(x))\big)\big|\big]\, dr < \infty. \end{aligned} \tag{3.40}$$

Note that (3.5) and the triangle inequality ensure that for all $t \in [0,T]$, $s \in [t,T]$ it holds that

$$\begin{aligned} &\mathbb{E}\big[|V^0_{M,n}(s, X^0_{t,s}(x))|\big] \\ &\leq \frac{1}{M^n}\left[\sum_{m=1}^{M^n} \mathbb{E}\Big[\big|g\big(X^{(0,n,-m)}_{s,T}(X^0_{t,s}(x))\big)\big|\Big]\right] \\ &+ \sum_{k=0}^{n-1} \frac{(T-s)}{M^{n-k}}\left[\sum_{m=1}^{M^{n-k}} \right. \\ &\quad \mathbb{E}\Big[\Big|f\Big(R^{(0,k,m)}_s, X^{(0,k,m)}_{s,R^{(0,k,m)}_s}(X^0_{t,s}(x)), V^{(0,k,m)}_{M,k}\big(R^{(0,k,m)}_s, X^{(0,k,m)}_{s,R^{(0,k,m)}_s}(X^0_{t,s}(x))\big)\Big)\Big|\Big] \\ &\quad \left. + \mathbb{1}_{\mathbb{N}}(k)\mathbb{E}\Big[\Big|f\Big(R^{(0,k,m)}_s, X^{(0,k,m)}_{s,R^{(0,k,m)}_s}(X^0_{t,s}(x)), V^{(0,k,-m)}_{M,k-1}\big(R^{(0,k,m)}_s, X^{(0,k,m)}_{s,R^{(0,k,m)}_s}(X^0_{t,s}(x))\big)\Big)\Big|\Big]\right]. \end{aligned} \tag{3.41}$$

Furthermore, observe that (3.3), (3.4), and item (iv) in Lemma 3.6 assure that for all $m \in \mathbb{Z}$, $t \in [0,T]$, $s \in [t,T]$ it holds that

$$\mathbb{E}\Big[\big|g\big(X^{(0,n,m)}_{s,T}(X^0_{t,s}(x))\big)\big|\Big] = \mathbb{E}\Big[\big|g\big(X^{(0,n,m)}_{t,T}(x)\big)\big|\Big] = \mathbb{E}\big[\big|g\big(X^0_{t,T}(x)\big)\big|\big] < \infty. \tag{3.42}$$

Moreover, note that Lemma 3.7, the hypothesis that $(X^\theta)_{\theta\in\Theta}$ are independent, the hypothesis that $(\mathcal{R}^\theta)_{\theta\in\Theta}$ are independent, the hypothesis that $(X^\theta)_{\theta\in\Theta}$ and $(\mathcal{R}^\theta)_{\theta\in\Theta}$ are independent, items (i)–(ii) & (iv)–(v) in Lemma 3.6, (3.3), and Lemma 2.16 demonstrate that for all $i,j,l,m \in \mathbb{Z}$, $k \in \mathbb{N}_0$, $t \in [0,T]$, $s \in [t,T]$ it holds that

$$\begin{aligned}
&(T-s)\mathbb{E}\Big[\Big|f\Big(R_s^{(0,j,m)}, X^{(0,j,m)}_{s,R_s^{(0,j,m)}}(X^0_{t,s}(x)), V^{(0,j,i)}_{M,k}\big(R_s^{(0,j,m)}, X^{(0,j,m)}_{s,R_s^{(0,j,m)}}(X^0_{t,s}(x))\big)\Big)\Big|\Big] \\
&= \int_s^T \mathbb{E}\Big[\Big|f\Big(r, X^{(0,j,m)}_{s,r}(X^0_{t,s}(x)), V^{(0,j,i)}_{M,k}\big(r, X^{(0,j,m)}_{s,r}(X^0_{t,s}(x))\big)\Big)\Big|\Big]\,dr \\
&= \int_s^T \int_{\mathbb{R}^d} \mathbb{E}\Big[\big|f\big(r,y,V^{(0,j,i)}_{M,k}(r,y)\big)\big|\Big]\,((X^{(0,j,m)}_{s,r}(X^0_{t,s}(x)))(\mathbb{P})_{\mathcal{B}(\mathbb{R}^d)})(dy)\,dr \\
&= \int_s^T \int_{\mathbb{R}^d} \mathbb{E}\big[\big|f\big(r,y,V^0_{M,k}(r,y)\big)\big|\big]\,(X^0_{t,r}(x)(\mathbb{P})_{\mathcal{B}(\mathbb{R}^d)})(dy)\,dr \\
&= \int_s^T \mathbb{E}\big[\big|f\big(r, X^0_{t,r}(x), V^0_{M,k}(r, X^0_{t,r}(x))\big)\big|\big]\,dr.
\end{aligned} \tag{3.43}$$

Combining this with (3.40), (3.41), and (3.42) establishes that for all $t \in [0,T]$, $s \in [t,T]$ it holds that

$$\begin{aligned}
&\mathbb{E}\big[|V^0_{M,n}(s, X^0_{t,s}(x))|\big] \\
&\le \Bigg(\sum_{k=0}^{n-1} \frac{1}{M^{n-k}} \sum_{m=1}^{M^{n-k}} \Bigg[\int_s^T \mathbb{E}\big[\big|f\big(r, X^0_{t,r}(x), V^0_{M,k}\big(r, X^0_{t,r}(x)\big)\big)\big|\big]\,dr \\
&+ \mathbb{1}_{\mathbb{N}}(k) \int_s^T \mathbb{E}\big[\big|f\big(r, X^0_{t,r}(x), V^0_{M,k-1}\big(r, X^0_{t,r}(x)\big)\big)\big|\big]\,dr\Bigg]\Bigg) + \frac{1}{M^n}\Bigg[\sum_{m=1}^{M^n} \mathbb{E}\big[\big|g\big(X^0_{t,T}(x)\big)\big|\big]\Bigg] \\
&= \Bigg[\sum_{k=0}^{n-1} \int_s^T \mathbb{E}\big[\big|f\big(r, X^0_{t,r}(x), V^0_{M,k}\big(r, X^0_{t,r}(x)\big)\big)\big|\big]\,dr \\
&\qquad + \mathbb{1}_{\mathbb{N}}(k) \int_s^T \mathbb{E}\big[\big|f\big(r, X^0_{t,r}(x), V^0_{M,k-1}\big(r, X^0_{t,r}(x)\big)\big)\big|\big]\,dr\Bigg] + \mathbb{E}\big[\big|g\big(X^0_{t,T}(x)\big)\big|\big] \\
&\le 2\Bigg[\sum_{k=0}^{n-1} \int_t^T \mathbb{E}\big[\big|f\big(r, X^0_{t,r}(x), V^0_{M,k}\big(r, X^0_{t,r}(x)\big)\big)\big|\big]\,dr\Bigg] + \mathbb{E}\big[\big|g\big(X^0_{t,T}(x)\big)\big|\big] < \infty.
\end{aligned} \tag{3.44}$$

Hence, we obtain that for all $t \in [0,T]$ it holds that

$$\begin{aligned}
&\int_t^T \mathbb{E}\big[|V^0_{M,n}(r, X^0_{t,r}(x))|\big]\,dr \le (T-t) \sup_{s\in[t,T]} \mathbb{E}\big[|V^0_{M,n}(s, X^0_{t,s}(x))|\big] \\
&\le (T-t)\Bigg(2\Bigg[\sum_{k=0}^{n-1} \int_t^T \mathbb{E}\big[\big|f\big(r, X^0_{t,r}(x), V^0_{M,k}\big(r, X^0_{t,r}(x)\big)\big)\big|\big]\,dr\Bigg] + \mathbb{E}\big[\big|g\big(X^0_{t,T}(x)\big)\big|\big]\Bigg) < \infty.
\end{aligned} \tag{3.45}$$

The hypothesis that for all $t \in [0,T]$ it holds that $\int_t^T \mathbb{E}\big[|f(r, X^0_{t,r}(x), 0)|\big]\,dr < \infty$ and the fact that for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $v \in \mathbb{R}$ it holds that $|f(t,x,v)| \le |f(t,x,0)| + L|v|$ therefore assure that for all $t \in [0,T]$ it holds that

$$\begin{aligned}
&\int_t^T \mathbb{E}\big[\big|f\big(r, X^0_{t,r}(x), V^0_{M,n}(r, X^0_{t,r}(x))\big)\big|\big]\,dr \\
&\le \int_t^T \mathbb{E}\big[|f(r, X^0_{t,r}(x), 0)|\big]\,dr + L\int_t^T \mathbb{E}\big[|V^0_{M,n}(r, X^0_{t,r}(x))|\big]\,dr < \infty.
\end{aligned} \tag{3.46}$$

This, (3.44), and (3.45) establish that for all $t \in [0,T]$, $s \in [t,T]$ it holds that

$$\begin{aligned}\mathbb{E}\big[|V^0_{M,n}(s, X^0_{t,s}(x))|\big] + \int_t^T \mathbb{E}\big[|V^0_{M,n}(r, X^0_{t,r}(x))|\big]\,dr \\ + \int_t^T \mathbb{E}\big[\big|f\big(r, X^0_{t,r}(x), V^0_{M,n}(r, X^0_{t,r}(x))\big)\big|\big]\,dr < \infty.\end{aligned} \tag{3.47}$$

Induction thus proves (3.38). Combining (3.37) and (3.38) establishes item (i). Next observe that (3.5), (3.38), items (i)–(ii) & (iv)–(v) in Lemma 3.6, the hypothesis that $(X^\theta)_{\theta\in\Theta}$ are independent, the hypothesis that $(\mathcal{R}^\theta)_{\theta\in\Theta}$ are independent, the hypothesis that $(X^\theta)_{\theta\in\Theta}$ and $(\mathcal{R}^\theta)_{\theta\in\Theta}$ are independent, and Lemma 3.5 ensure that for all $n \in \mathbb{N}$, $t \in [0,T]$ it holds that

$$\begin{aligned}&\mathbb{E}\big[V^0_{M,n}(t,x)\big] \\ &= \frac{1}{M^n}\left(\sum_{m=1}^{M^n} \mathbb{E}\Big[g\big(X^{(0,n,-m)}_{t,T}(x)\big)\Big]\right) \\ &\quad + \sum_{k=0}^{n-1} \frac{(T-t)}{M^{n-k}}\left[\sum_{m=1}^{M^{n-k}} \mathbb{E}\Big[f\Big(R^{(0,k,m)}_t, X^{(0,k,m)}_{t,R^{(0,k,m)}_t}(x), V^{(0,k,m)}_{M,k}\big(R^{(0,k,m)}_t, X^{(0,k,m)}_{t,R^{(0,k,m)}_t}(x)\big)\Big)\Big]\right. \\ &\quad \left. - \mathbb{1}_{\mathbb{N}}(k)\mathbb{E}\Big[f\Big(R^{(0,k,m)}_t, X^{(0,k,m)}_{t,R^{(0,k,m)}_t}(x), V^{(0,k,-m)}_{M,k-1}\big(R^{(0,k,m)}_t, X^{(0,k,m)}_{t,R^{(0,k,m)}_t}(x)\big)\Big)\Big]\right] \\ &= (T-t)\left(\sum_{k=0}^{n-1} \mathbb{E}\Big[f\big(R^0_t, X^0_{t,R^0_t}(x), V^0_{M,k}(R^0_t, X^0_{t,R^0_t}(x))\big)\Big]\right. \\ &\quad \left. - \mathbb{1}_{\mathbb{N}}(k)\mathbb{E}\Big[f\big(R^0_t, X^0_{t,R^0_t}(x), V^0_{M,k-1}(R^0_t, X^0_{t,R^0_t}(x))\big)\Big]\right) + \mathbb{E}\big[g\big(X^0_{t,T}(x)\big)\big] \\ &= (T-t)\,\mathbb{E}\Big[f\big(R^0_t, X^0_{t,R^0_t}(x), V^0_{M,n-1}(R^0_t, X^0_{t,R^0_t}(x))\big)\Big] + \mathbb{E}\big[g\big(X^0_{t,T}(x)\big)\big].\end{aligned} \tag{3.48}$$

Lemma 3.7, items (i)–(ii) in Lemma 3.6, the fact that for all $n \in \mathbb{N}_0$ it holds that $V^0_{M,n}$, X^0, and $\mathcal{R}^0$ are independent, (3.38), and Fubini's theorem therefore imply that for all $n \in \mathbb{N}$, $t \in [0,T]$ it holds that

$$\begin{aligned}\mathbb{E}\big[V^0_{M,n}(t,x)\big] &= \int_t^T \mathbb{E}\big[f\big(r, X^0_{t,r}(x), V^0_{M,n-1}(r, X^0_{t,r}(x))\big)\big]\,dr + \mathbb{E}\big[g(X^0_{t,T}(x))\big] \\ &= \mathbb{E}\left[g(X^0_{t,T}(x)) + \int_t^T f\big(r, X^0_{t,r}(x), V^0_{M,n-1}(r, X^0_{t,r}(x))\big)\,dr\right].\end{aligned} \tag{3.49}$$

This establishes item (ii). The proof of Lemma 3.8 is thus completed. □

3.4.2 Biases of MLP approximations

Lemma 3.9 (Biases of MLP approximations)**.** *Assume Setting 3.1 and assume for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that $\int_t^T \mathbb{E}\big[|f(r, X^0_{t,r}(x), 0)|\big]\,dr < \infty$. Then it holds for all $M, n \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that*

$$\big|u(t,x) - \mathbb{E}\big[V^0_{M,n}(t,x)\big]\big|^2 \le L^2 T \int_t^T \mathbb{E}\big[\big|u(r, X^0_{t,r}(x)) - V^0_{M,n-1}(r, X^0_{t,r}(x))\big|^2\big]\,dr. \tag{3.50}$$

Proof of Lemma 3.9. Note that Lemma 3.8, the hypothesis that for all $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that $\int_t^T \mathbb{E}\big[|f(r, X^0_{t,r}(x), 0)|\big]\, dr < \infty$, (3.1), (3.4), and Tonelli's theorem demonstrate that for all $M, n \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
&\big|u(t,x) - \mathbb{E}\big[V^0_{M,n}(t,x)\big]\big| \\
&= \Bigg|\mathbb{E}\bigg[g\big(X^0_{t,T}(x)\big) + \int_t^T f\big(r, X^0_{t,r}(x), u(r, X^0_{t,r}(x))\big)\, dr\bigg] \\
&\quad - \mathbb{E}\bigg[g(X^0_{t,T}(x)) + \int_t^T f\big(r, X^0_{t,r}(x), V^0_{M,n-1}(r, X^0_{t,r}(x))\big)\, dr\bigg]\Bigg| \\
&= \Bigg|\mathbb{E}\bigg[\int_t^T f\big(r, X^0_{t,r}(x), u(r, X^0_{t,r}(x))\big) - f\big(r, X^0_{t,r}(x), V^0_{M,n-1}(r, X^0_{t,r}(x))\big)\, dr\bigg]\Bigg| \\
&\leq \mathbb{E}\bigg[\int_t^T \big|f\big(r, X^0_{t,r}(x), u(r, X^0_{t,r}(x))\big) - f\big(r, X^0_{t,r}(x), V^0_{M,n-1}(r, X^0_{t,r}(x))\big)\big|\, dr\bigg] \\
&\leq \mathbb{E}\bigg[\int_t^T L\big|u(r, X^0_{t,r}(x)) - V^0_{M,n-1}(r, X^0_{t,r}(x))\big|\, dr\bigg] \\
&= L\int_t^T \mathbb{E}\big[\big|u(r, X^0_{t,r}(x)) - V^0_{M,n-1}(r, X^0_{t,r}(x))\big|\big]\, dr.
\end{aligned} \tag{3.51}$$

Lemma 2.10 and Jensen's inequality hence show that for all $M, n \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
&\big|u(t,x) - \mathbb{E}\big[V^0_{M,n}(t,x)\big]\big|^2 \\
&\leq L^2 \left(\int_t^T \mathbb{E}\big[\big|u(r, X^0_{t,r}(x)) - V^0_{M,n-1}(r, X^0_{t,r}(x))\big|\big]\, dr\right)^2 \\
&\leq L^2 (T-t) \int_t^T \big(\mathbb{E}\big[\big|u(r, X^0_{t,r}(x)) - V^0_{M,n-1}(r, X^0_{t,r}(x))\big|\big]\big)^2\, dr \\
&\leq L^2 T \int_t^T \mathbb{E}\big[\big|u(r, X^0_{t,r}(x)) - V^0_{M,n-1}(r, X^0_{t,r}(x))\big|^2\big]\, dr.
\end{aligned} \tag{3.52}$$

The proof of Lemma 3.9 is thus completed. ☐

3.4.3 Estimates for the variances of MLP approximations

Lemma 3.10. *Let $n \in \mathbb{N}$, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $X_1, X_2, \ldots, X_n \colon \Omega \to \mathbb{R}$ be independent random variables which satisfy for all $i \in \{1, 2, \ldots, n\}$ that $\mathbb{E}[|X_i|] < \infty$. Then it holds that*

$$\begin{aligned}
\mathrm{Var}\left(\sum_{i=1}^n X_i\right) &= \mathbb{E}\Big[\big|\mathbb{E}\big[\textstyle\sum_{i=1}^n X_i\big] - \sum_{i=1}^n X_i\big|^2\Big] \\
&= \sum_{i=1}^n \mathbb{E}\Big[\big|\mathbb{E}[X_i] - X_i\big|^2\Big] = \sum_{i=1}^n \mathrm{Var}\left(X_i\right).
\end{aligned} \tag{3.53}$$

Proof of Lemma 3.10. Note that the fact that for all independent random variables $Y, Z \colon \Omega \to \mathbb{R}$ with $\mathbb{E}[|Y| + |Z|] < \infty$ it holds that $\mathbb{E}[|YZ|] < \infty$ and $\mathbb{E}[YZ] = \mathbb{E}[Y]\,\mathbb{E}[Z]$ (cf., e.g., Klenke [70, Theorem 5.4]) and the hypothesis that $X_i \colon \Omega \to \mathbb{R}$, $i \in \{1, 2, \ldots, n\}$, are independent random variables assure that

$$\begin{aligned}
&\mathbb{E}\Big[\big|\mathbb{E}\big[\textstyle\sum_{i=1}^{n} X_i\big] - \sum_{i=1}^{n} X_i\big|^2\Big] \\
&= \mathbb{E}\Big[\big|\textstyle\sum_{i=1}^{n} \big(\mathbb{E}[X_i] - X_i\big)\big|^2\Big] \\
&= \mathbb{E}\left[\sum_{i,j=1}^{n} (\mathbb{E}[X_i] - X_i)(\mathbb{E}[X_j] - X_j)\right] \\
&= \left[\sum_{i=1}^{n} \mathbb{E}\big[|\mathbb{E}[X_i] - X_i|^2\big]\right] + \left[\sum_{i,j=1, i\neq j}^{n} \mathbb{E}\big[\mathbb{E}[X_i] - X_i\big]\,\mathbb{E}\big[\mathbb{E}[X_j] - X_j\big]\right] \\
&= \sum_{i=1}^{n} \mathbb{E}\big[|\mathbb{E}[X_i] - X_i|^2\big].
\end{aligned} \tag{3.54}$$

The proof of Lemma 3.10 is thus completed. □

Lemma 3.11 (Estimates for the variances of MLP approximations)**.** *Assume Setting 3.1 and assume for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that $\int_t^T \mathbb{E}\big[|f(r, X_{t,r}^0(x), 0)|\big]\, dr < \infty$. Then it holds for all $M, n \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that*

$$\begin{aligned}
&\mathbb{E}\Big[\big|V_{M,n}^0(t,x) - \mathbb{E}\big[V_{M,n}^0(t,x)\big]\big|^2\Big] \\
&\leq \tfrac{1}{M^n}\bigg(\mathbb{E}\big[|g(X_{t,T}^0(x))|^2\big] + T\int_t^T \mathbb{E}\big[|f(r, X_{t,r}^0(x), 0)|^2\big]\, dr\bigg) \\
&\quad + \sum_{k=1}^{n-1} \tfrac{2L^2T}{M^{n-k}}\bigg(\int_t^T \mathbb{E}\Big[\big|u(r, X_{t,r}^0(x)) - V_{M,k}^0(r, X_{t,r}^0(x))\big|^2\Big]\, dr \\
&\quad + \int_t^T \mathbb{E}\Big[\big|u(r, X_{t,r}^0(x)) - V_{M,k-1}^0(r, X_{t,r}^0(x))\big|^2\Big]\, dr\bigg).
\end{aligned} \tag{3.55}$$

Proof of Lemma 3.11. Throughout this proof let $M, n \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$. Observe that Lemma 3.10, item (i) in Lemma 3.8, the fact that for all $\theta \in \Theta$ it holds that $\mathbb{E}\big[|g(X_{t,T}^0(x))|\big] < \infty$, item (iii) in Lemma 3.6, and (3.5) imply that

$$\begin{aligned}
&\mathbb{E}\Big[\big|V_{M,n}^0(t,x) - \mathbb{E}\big[V_{M,n}^0(t,x)\big]\big|^2\Big] = \mathrm{Var}(V_{M,n}^0(t,x)) \\
&= \left[\sum_{m=1}^{M^n} \mathrm{Var}\Big(\tfrac{1}{M^n} g\big(X_{t,T}^{(0,n,-m)}(x)\big)\Big)\right] \\
&\quad + \sum_{k=0}^{n-1}\sum_{m=1}^{M^{n-k}} \mathrm{Var}\Bigg(\tfrac{(T-t)}{M^{n-k}}\bigg[f\Big(R_t^{(0,k,m)}, X_{t,R_t^{(0,k,m)}}^{(0,k,m)}(x), V_{M,k}^{(0,k,m)}\big(R_t^{(0,k,m)}, X_{t,R_t^{(0,k,m)}}^{(0,k,m)}(x)\big)\Big) \\
&\quad - \mathbb{1}_{\mathbb{N}}(k) f\Big(R_t^{(0,k,m)}, X_{t,R_t^{(0,k,m)}}^{(0,k,m)}(x), V_{M,k-1}^{(0,k,-m)}\big(R_t^{(0,k,m)}, X_{t,R_t^{(0,k,m)}}^{(0,k,m)}(x)\big)\Big)\bigg]\Bigg).
\end{aligned} \tag{3.56}$$

Moreover, note that item (iv) in Lemma 3.6 and the fact that for all $Z \in \mathcal{L}^1(\mathbb{P}; \mathbb{R})$ it holds that $\mathrm{Var}(Z) \leq \mathbb{E}[|Z|^2]$ ensure that

$$\begin{aligned}
\sum_{m=1}^{M^n} \mathrm{Var}\Big(\tfrac{1}{M^n} g\big(X_{t,T}^{(0,n,-m)}(x)\big)\Big) &= M^n\, \mathrm{Var}\big(\tfrac{1}{M^n} g(X_{t,T}^0(x))\big) \\
&= \tfrac{M^n}{M^{2n}}\, \mathrm{Var}\big(g(X_{t,T}^0(x))\big) \leq \tfrac{1}{M^n}\big(\mathbb{E}\big[|g(X_{t,T}^0(x))|^2\big]\big).
\end{aligned} \tag{3.57}$$

In addition, note that items (i)–(ii) & (iv)–(v) in Lemma 3.6, the hypothesis that $(X^\theta)_{\theta\in\Theta}$ are independent, the hypothesis that $(\mathcal{R}^\theta)_{\theta\in\Theta}$ are independent, the hypothesis that $(X^\theta)_{\theta\in\Theta}$ and $(\mathcal{R}^\theta)_{\theta\in\Theta}$ are independent, the fact that for all $Z \in \mathcal{L}^1(\mathbb{P};\mathbb{R})$ it holds that $\mathrm{Var}(Z) \le \mathbb{E}[|Z|^2]$, and Lemma 3.5 show that for all $k \in \mathbb{N}_0 \cap [0,n)$ it holds that

$$\begin{aligned}
&\sum_{m=1}^{M^{n-k}} \mathrm{Var}\Bigg(\tfrac{(T-t)}{M^{n-k}}\bigg[f\Big(R_t^{(0,k,m)}, X^{(0,k,m)}_{t,R_t^{(0,k,m)}}(x), V^{(0,k,m)}_{M,k}\big(R_t^{(0,k,m)}, X^{(0,k,m)}_{t,R_t^{(0,k,m)}}(x)\big)\Big)\\
&\quad - \mathbb{1}_{\mathbb{N}}(k) f\Big(R_t^{(0,k,m)}, X^{(0,k,m)}_{t,R_t^{(0,k,m)}}(x), V^{(0,k,-m)}_{M,k-1}\big(R_t^{(0,k,m)}, X^{(0,k,m)}_{t,R_t^{(0,k,m)}}(x)\big)\Big)\bigg]\Bigg)\\
&= M^{n-k}\,\mathrm{Var}\Bigg(\tfrac{(T-t)}{M^{n-k}}\bigg[f\Big(R_t^0, X^0_{t,R_t^0}(x), V^0_{M,k}\big(R_t^0, X^0_{t,R_t^0}(x)\big)\Big)\\
&\quad - \mathbb{1}_{\mathbb{N}}(k) f\Big(R_t^0, X^0_{t,R_t^0}(x), V^1_{M,k-1}\big(R_t^0, X^0_{t,R_t^0}(x)\big)\Big)\bigg]\Bigg)\\
&= \tfrac{M^{n-k}(T-t)^2}{M^{2(n-k)}}\,\mathrm{Var}\Bigg(f\Big(R_t^0, X^0_{t,R_t^0}(x), V^0_{M,k}\big(R_t^0, X^0_{t,R_t^0}(x)\big)\Big)\\
&\quad - \mathbb{1}_{\mathbb{N}}(k) f\Big(R_t^0, X^0_{t,R_t^0}(x), V^1_{M,k-1}\big(R_t^0, X^0_{t,R_t^0}(x)\big)\Big)\Bigg)\\
&\le \tfrac{(T-t)^2}{M^{n-k}}\mathbb{E}\bigg[\Big|f\Big(R_t^0, X^0_{t,R_t^0}(x), V^0_{M,k}\big(R_t^0, X^0_{t,R_t^0}(x)\big)\Big)\\
&\quad - \mathbb{1}_{\mathbb{N}}(k) f\Big(R_t^0, X^0_{t,R_t^0}(x), V^1_{M,k-1}\big(R_t^0, X^0_{t,R_t^0}(x)\big)\Big)\Big|^2\bigg].
\end{aligned}\tag{3.58}$$

Lemma 3.7, the fact that X^0 and R^0 are independent, and the hypothesis that for all $\theta \in \Theta$ it holds that $V^\theta_{M,0} = 0$ therefore demonstrate that

$$\begin{aligned}
&\sum_{m=1}^{M^n} \mathrm{Var}\Bigg(\tfrac{(T-t)}{M^n} f\Big(R_t^{(0,0,m)}, X^{(0,0,m)}_{t,R_t^{(0,0,m)}}(x), V^{(0,0,m)}_{M,0}\big(R_t^{(0,0,m)}, X^{(0,0,m)}_{t,R_t^{(0,0,m)}}(x)\big)\Big)\Bigg)\\
&\le \frac{(T-t)^2}{M^n}\mathbb{E}\Big[|f(R_t^0, X^0_{t,R_t^0}(x), 0)|^2\Big] = \frac{(T-t)}{M^n}\int_t^T \mathbb{E}\big[|f(r, X^0_{t,r}(x), 0)|^2\big]\,dr.
\end{aligned}\tag{3.59}$$

In addition, observe that (3.1), (3.58), the fact that for all $x, y \in [0,\infty)$ it holds that $|x+y|^2 \le 2(|x|^2+|y|^2)$, items (i)–(ii) & (v) in Lemma 3.6, the hypothesis that $(X^\theta)_{\theta\in\Theta}$ are independent, the hypothesis that $(\mathcal{R}^\theta)_{\theta\in\Theta}$ are independent, the hypothesis that $(X^\theta)_{\theta\in\Theta}$ and $(\mathcal{R}^\theta)_{\theta\in\Theta}$ are independent, and Lemma 3.5 assure that for all $k \in \mathbb{N} \cap [1,n)$ it holds that

$$\begin{aligned}
&\sum_{m=1}^{M^{n-k}} \mathrm{Var}\Bigg(\tfrac{(T-t)}{M^{n-k}}\bigg[f\Big(R_t^{(0,k,m)}, X^{(0,k,m)}_{t,R_t^{(0,k,m)}}(x), V^{(0,k,m)}_{M,k}\big(R_t^{(0,k,m)}, X^{(0,k,m)}_{t,R_t^{(0,k,m)}}(x)\big)\Big)\\
&\quad - f\Big(R_t^{(0,k,m)}, X^{(0,k,m)}_{t,R_t^{(0,k,m)}}(x), V^{(0,k,-m)}_{M,k-1}\big(R_t^{(0,k,m)}, X^{(0,k,m)}_{t,R_t^{(0,k,m)}}(x)\big)\Big)\bigg]\Bigg)\\
&\le \tfrac{(T-t)^2}{M^{n-k}}\mathbb{E}\Big[L^2\big|V^0_{M,k}\big(R_t^0, X^0_{t,R_t^0}(x)\big) - V^1_{M,k-1}\big(R_t^0, X^0_{t,R_t^0}(x)\big)\big|^2\Big]\\
&\le \tfrac{2L^2(T-t)^2}{M^{n-k}}\bigg(\mathbb{E}\Big[\big|V^0_{M,k}\big(R_t^0, X^0_{t,R_t^0}(x)\big) - u\big(R_t^0, X^0_{t,R_t^0}(x)\big)\big|^2\Big]\\
&\quad + \mathbb{E}\Big[\big|V^0_{M,k-1}\big(R_t^0, X^0_{t,R_t^0}(x)\big) - u\big(R_t^0, X^0_{t,R_t^0}(x)\big)\big|^2\Big]\bigg).
\end{aligned}\tag{3.60}$$

Lemma 3.7, items (i)–(ii) in Lemma 3.6, the hypothesis that $(X^\theta)_{\theta\in\Theta}$ are independent, the hypothesis that $(\mathcal{R}^\theta)_{\theta\in\Theta}$ are independent, and the hypothesis that $(X^\theta)_{\theta\in\Theta}$ and $(\mathcal{R}^\theta)_{\theta\in\Theta}$ are independent hence ensure that for all $k \in \mathbb{N} \cap [1, n)$ it holds that

$$\begin{aligned}
&\sum_{m=1}^{M^{n-k}} \mathrm{Var}\bigg(\tfrac{(T-t)}{M^{n-k}}\Big[f\Big(R_t^{(0,k,m)}, X^{(0,k,m)}_{t,R_t^{(0,k,m)}}(x), V^{(0,k,m)}_{M,k}\big(R_t^{(0,k,m)}, X^{(0,k,m)}_{t,R_t^{(0,k,m)}}(x)\big)\Big)\\
&\quad - f\Big(R_t^{(0,k,m)}, X^{(0,k,m)}_{t,R_t^{(0,k,m)}}(x), V^{(0,k,-m)}_{M,k-1}\big(R_t^{(0,k,m)}, X^{(0,k,m)}_{t,R_t^{(0,k,m)}}(x)\big)\Big)\Big]\bigg)\\
&\le \tfrac{2L^2(T-t)}{M^{n-k}}\bigg(\int_t^T \mathbb{E}\Big[\big|V^0_{M,k}\big(r, X^0_{t,r}(x)\big) - u\big(r, X^0_{t,r}(x)\big)\big|^2\Big]\, dr\\
&\quad + \int_t^T \mathbb{E}\Big[\big|V^0_{M,k-1}\big(r, X^0_{t,r}(x)\big) - u\big(r, X^0_{t,r}(x)\big)\big|^2\Big]\, dr\bigg).
\end{aligned} \tag{3.61}$$

Combining this with (3.56), (3.57), and (3.59) establishes that

$$\begin{aligned}
&\mathbb{E}\Big[\big|V^0_{M,n}(t,x) - \mathbb{E}\big[V^0_{M,n}(t,x)\big]\big|^2\Big]\\
&\le \tfrac{1}{M^n}\Big(\mathbb{E}\big[|g(X^0_{t,T}(x))|^2\big]\Big) + \tfrac{(T-t)}{M^n}\int_t^T \mathbb{E}\big[|f(r, X^0_{t,r}(x), 0)|^2\big]\, dr\\
&\quad + \sum_{k=1}^{n-1} \tfrac{2L^2(T-t)}{M^{n-k}}\bigg(\int_t^T \mathbb{E}\Big[\big|V^0_{M,k}\big(r, X^0_{t,r}(x)\big) - u\big(r, X^0_{t,r}(x)\big)\big|^2\Big]\, dr\\
&\quad + \int_t^T \mathbb{E}\Big[\big|V^0_{M,k-1}\big(r, X^0_{t,r}(x)\big) - u\big(r, X^0_{t,r}(x)\big)\big|^2\Big]\, dr\bigg)\\
&\le \tfrac{1}{M^n}\Big(\mathbb{E}\big[|g(X^0_{t,T}(x))|^2\big] + T\int_t^T \mathbb{E}\big[|f(r, X^0_{t,r}(x), 0)|^2\big]\, dr\Big)\\
&\quad + \sum_{k=1}^{n-1} \tfrac{2L^2T}{M^{n-k}}\bigg(\int_t^T \mathbb{E}\Big[\big|u(r, X^0_{t,r}(x)) - V^0_{M,k}(r, X^0_{t,r}(x))\big|^2\Big]\, dr\\
&\quad + \int_t^T \mathbb{E}\Big[\big|u(r, X^0_{t,r}(x)) - V^0_{M,k-1}(r, X^0_{t,r}(x))\big|^2\Big]\, dr\bigg).
\end{aligned} \tag{3.62}$$

The proof of Lemma 3.11 is thus completed. □

3.4.4 On a geometric time-discrete Gronwall inequality

Lemma 3.12. *Let* $\alpha, \beta \in [0, \infty)$, $M \in (0, \infty)$, $(\epsilon_{n,q})_{n,q\in\mathbb{N}_0} \subseteq [0, \infty]$ *satisfy for all* $n, q \in \mathbb{N}_0$ *that*

$$\epsilon_{n,q} \le \frac{\alpha}{M^{n+q}} + \beta\left[\sum_{k=0}^{n-1} \frac{\epsilon_{k,q+1}}{M^{n-(k+1)}}\right]. \tag{3.63}$$

Then it holds for all $n, q \in \mathbb{N}_0$ *that*

$$\epsilon_{n,q} \le \frac{\alpha(1+\beta)^n}{M^{n+q}} < \infty. \tag{3.64}$$

Proof of Lemma 3.12. Throughout this proof assume w.l.o.g. that $\beta > 0$. We prove (3.64) by induction on $n \in \mathbb{N}_0$. For the base case $n = 0$ observe that (3.63) assures that for all $q \in \mathbb{N}_0$ it holds that

$$\epsilon_{0,q} \le \frac{\alpha}{M^{0+q}} = \frac{\alpha}{M^{0+q}}(1+\beta)^0 < \infty. \tag{3.65}$$

This proves (3.64) in the base case $n = 0$. For the induction step $\mathbb{N}_0 \ni (n-1) \to n \in \mathbb{N}$ observe that (3.63) ensures that for all $n \in \mathbb{N}$, $q \in \mathbb{N}_0$ with $\forall\, k \in \mathbb{N}_0 \cap [0,n), p \in \mathbb{N}_0\colon \epsilon_{k,p} \leq \alpha \frac{(1+\beta)^k}{M^{k+p}}$ it holds that

$$\begin{aligned}
\epsilon_{n,q} &\leq \frac{\alpha}{M^{n+q}} + \beta \left[\sum_{k=0}^{n-1} \frac{\epsilon_{k,q+1}}{M^{n-(k+1)}} \right] \leq \frac{\alpha}{M^{n+q}} + \beta \left[\sum_{k=0}^{n-1} \frac{\alpha \frac{(1+\beta)^k}{M^{k+q+1}}}{M^{n-(k+1)}} \right] \\
&= \frac{\alpha}{M^{n+q}} + \beta \left[\sum_{k=0}^{n-1} \frac{\alpha(1+\beta)^k}{M^{n-(k+1)+(k+q+1)}} \right] = \frac{\alpha}{M^{n+q}} + \beta \left[\sum_{k=0}^{n-1} \frac{\alpha(1+\beta)^k}{M^{n+q}} \right] \\
&= \frac{\alpha}{M^{n+q}} \left(1 + \beta \left[\sum_{k=0}^{n-1} (1+\beta)^k \right] \right) = \frac{\alpha}{M^{n+q}} \left(1 + \beta \frac{(1+\beta)^n - 1}{(1+\beta) - 1} \right) \\
&= \frac{\alpha}{M^{n+q}} (1+\beta)^n.
\end{aligned} \tag{3.66}$$

Induction hence establishes (3.64). The proof of Lemma 3.12 is thus completed. ☐

3.4.5 Error estimates for MLP approximations

Corollary 3.13. *Assume Setting 3.1 and assume for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that $\int_t^T \mathbb{E}\big[|f(r, X^0_{t,r}(x), 0)|\big]\, dr < \infty$. Then it holds for all $M, n \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that*

$$\begin{aligned}
&\mathbb{E}\big[|u(t,x) - V^0_{M,n}(t,x)|^2\big] \\
&\leq \tfrac{1}{M^n} \Big(\mathbb{E}\big[|g(X^0_{t,T}(x))|^2\big] + T \int_t^T \mathbb{E}\big[|f(r, X^0_{t,r}(x), 0)|^2\big]\, dr \Big) \\
&\quad + \sum_{k=0}^{n-1} \tfrac{4L^2 T}{M^{n-(k+1)}} \int_t^T \mathbb{E}\Big[\big| u(r, X^0_{t,r}(x)) - V^0_{M,k}(r, X^0_{t,r}(x)) \big|^2 \Big]\, dr.
\end{aligned} \tag{3.67}$$

Proof of Corollary 3.13. Throughout this proof let $M, n \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$, $C \in [0,\infty]$, $(e_k)_{k \in \mathbb{N}_0 \cap [0,n)} \subseteq [0,\infty]$ satisfy for all $k \in \mathbb{N}_0 \cap [0,n)$ that

$$C = \mathbb{E}\big[|g(X^0_{t,T}(x))|^2\big] + T \int_t^T \mathbb{E}\big[|f(r, X^0_{t,r}(x), 0)|^2\big]\, dr \tag{3.68}$$

and

$$e_k = \int_t^T \mathbb{E}\Big[\big| u(r, X^0_{t,r}(x)) - V^0_{M,k}(r, X^0_{t,r}(x)) \big|^2 \Big]\, dr. \tag{3.69}$$

Note that item (i) in Lemma 3.8, the bias variance decomposition of the mean square error (cf., e.g., Jentzen & von Wurstemberger [66, Lemma 2.2]), the hypothesis that for all $s \in [0,T]$, $z \in \mathbb{R}^d$ it holds that $\int_s^T \mathbb{E}\big[|f(r, X^0_{s,r}(z), 0)|\big]\, dr < \infty$, Lemma 3.9, and Lemma 3.11 demonstrate that

$$\begin{aligned}
&\mathbb{E}\big[|u(t,x) - V^0_{M,n}(t,x)|^2\big] \\
&= \mathbb{E}\Big[\big| V^0_{M,n}(t,x) - \mathbb{E}[V^0_{M,n}(t,x)] \big|^2 \Big] + \big| u(t,x) - \mathbb{E}[V^0_{M,n}(t,x)] \big|^2 \\
&\leq \tfrac{C}{M^n} + \sum_{k=1}^{n-1} \tfrac{2L^2 T}{M^{n-k}} (e_k + e_{k-1}) + L^2 T e_{n-1} \\
&\leq \tfrac{C}{M^n} + \left[\sum_{k=1}^{n-1} \tfrac{2L^2 T}{M^{n-k}} e_k \right] + \sum_{k=0}^{n-1} \tfrac{2L^2 T}{M^{n-(k+1)}} e_k \leq \tfrac{C}{M^n} + \sum_{k=0}^{n-1} \tfrac{4L^2 T}{M^{n-(k+1)}} e_k.
\end{aligned} \tag{3.70}$$

The proof of Corollary 3.13 is thus completed. ☐

Lemma 3.14. *Let $T \in [0,\infty)$, $q \in \mathbb{N}$ and let $U\colon [0,T] \to [0,\infty]$ be a $\mathcal{B}([0,T])/\mathcal{B}([0,\infty])$-measurable function. Then*

$$\int_0^T \frac{t^{q-1}}{(q-1)!} \int_t^T U(r)\,dr\,dt = \int_0^T \frac{t^q}{q!} U(t)\,dt. \tag{3.71}$$

Proof of Lemma 3.14. Observe that Tonelli's theorem assures that

$$\begin{aligned}
&\int_0^T \frac{t^{q-1}}{(q-1)!} \int_t^T U(r)\,dr\,dt \\
&= \int_0^T \int_0^T \frac{t^{q-1}}{(q-1)!} U(r)\,\mathbb{1}_{\{(\mathfrak{t},\mathfrak{r})\in[0,T]^2\colon\, \mathfrak{t}\le\mathfrak{r}\}}(t,r)\,dr\,dt \\
&= \int_0^T \int_0^T \frac{t^{q-1}}{(q-1)!} U(r)\,\mathbb{1}_{\{(\mathfrak{t},\mathfrak{r})\in[0,T]^2\colon\, \mathfrak{t}\le\mathfrak{r}\}}(t,r)\,dt\,dr \\
&= \int_0^T \int_0^r \frac{t^{q-1}}{(q-1)!}\,dt\,U(r)\,dr = \int_0^T \frac{r^q}{q!} U(r)\,dr.
\end{aligned} \tag{3.72}$$

The proof of Lemma 3.14 is thus completed. □

Proposition 3.15. *Assume Setting 3.1, let $\xi \in \mathbb{R}^d$, $C \in [0,\infty]$ satisfy that*

$$C = \left[\left(\mathbb{E}\big[|g(X^0_{0,T}(\xi))|^2\big]\right)^{1/2} + \sqrt{T} \left(\int_0^T \mathbb{E}\big[|f(t, X^0_{0,t}(\xi), 0)|^2\big]\,dt \right)^{1/2} \right] \exp(LT), \tag{3.73}$$

and assume for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that $\int_0^T \left(\mathbb{E}\big[|u(r, X^0_{0,r}(\xi))|^2\big]\right)^{1/2} dr + \int_t^T \mathbb{E}\big[|f(r, X^0_{t,r}(x), 0)|\big]\,dr < \infty$. Then it holds for all $M \in \mathbb{N}$, $n \in \mathbb{N}_0$ that

$$\left(\mathbb{E}\big[|u(0,\xi) - V^0_{M,n}(0,\xi)|^2\big]\right)^{1/2} \le \frac{C(1+2LT)^n \exp(\frac{M}{2})}{M^{n/2}}. \tag{3.74}$$

Proof of Proposition 3.15. Throughout this proof assume w.l.o.g. that $C < \infty$, let $M \in \mathbb{N}$, let $\epsilon_{n,q} \in [0,\infty]$, $n, q \in \mathbb{N}_0$, be the extended real numbers which satisfy for all $n, q \in \mathbb{N}_0$ that

$$\epsilon_{n,0} = \mathbb{E}\big[|u(0,\xi) - V^0_{M,n}(0,\xi)|^2\big] \qquad \text{and} \tag{3.75}$$

$$\epsilon_{n,q+1} = \frac{1}{T^{q+1}} \int_0^T \frac{t^q}{q!} \mathbb{E}\big[|u(t, X^0_{0,t}(\xi)) - V^0_{M,n}(t, X^0_{0,t}(\xi))|^2\big]\,dt, \tag{3.76}$$

and let $\mu_t\colon \mathcal{B}(\mathbb{R}^d) \to [0,1]$, $t \in [0,T]$, be the probability measures which satisfy for all $t \in [0,T]$, $B \in \mathcal{B}(\mathbb{R}^d)$ that

$$\mu_t(B) = \mathbb{P}(X^0_{0,t}(\xi) \in B) = \mathbb{P}(X^1_{0,t}(\xi) \in B) = \big((X^1_{0,t}(\xi))(\mathbb{P})_{\mathcal{B}(\mathbb{R}^d)}\big)(B) \tag{3.77}$$

(cf. item (iv) in Lemma 3.6). Note that the fact that for all $x, y \in [0,\infty)$ it holds that $(x+y)^2 \ge x^2 + y^2$ assures that

$$\begin{aligned}
C^2 &\ge \left[\mathbb{E}\big[|g(X^0_{0,T}(\xi))|^2\big] + T \int_0^T \mathbb{E}\big[|f(t, X^0_{0,t}(\xi), 0)|^2\big]\,dt \right] \exp(2LT) \\
&\ge \mathbb{E}\big[|g(X^0_{0,T}(\xi))|^2\big] + T \int_0^T \mathbb{E}\big[|f(t, X^0_{0,t}(\xi), 0)|^2\big]\,dt.
\end{aligned} \tag{3.78}$$

Next observe that items (i)–(ii) in Lemma 3.6, the hypothesis that $(X^\theta)_{\theta\in\Theta}$ are independent, the hypothesis that $(\mathcal{R}^\theta)_{\theta\in\Theta}$ are independent, the hypothesis that $(X^\theta)_{\theta\in\Theta}$ and

$(\mathcal{R}^\theta)_{\theta\in\Theta}$ are independent, Tonelli's theorem, Corollary 3.13, and Lemma 2.16 ensure that for all $n \in \mathbb{N}$, $t \in [0,T]$ it holds that

$$\begin{aligned}
&\mathbb{E}\big[|u(t, X^0_{0,t}(\xi)) - V^0_{M,n}(t, X^0_{0,t}(\xi))|^2\big] = \int_{\mathbb{R}^d} \mathbb{E}\big[|u(t,z) - V^0_{M,n}(t,z)|^2\big]\ \mu_t(dz)\\
&\leq \int_{\mathbb{R}^d} \Bigg[\tfrac{1}{M^n}\Big(\mathbb{E}\big[|g(X^0_{t,T}(z))|^2\big] + T\int_t^T \mathbb{E}\big[|f(r, X^0_{t,r}(z), 0)|^2\big]\ dr\Big)\\
&\quad + \sum_{k=0}^{n-1} \tfrac{4L^2T}{M^{n-(k+1)}} \int_t^T \mathbb{E}\Big[\big|u(r, X^0_{t,r}(z)) - V^0_{M,k}(r, X^0_{t,r}(z))\big|^2\Big]\ dr\Bigg]\mu_t(dz)\\
&= \tfrac{1}{M^n}\Bigg(\int_{\mathbb{R}^d} \mathbb{E}\big[|g(X^0_{t,T}(z))|^2\big]\ \mu_t(dz) + T\int_t^T\int_{\mathbb{R}^d} \mathbb{E}\big[|f(r, X^0_{t,r}(z), 0)|^2\big]\ \mu_t(dz)\, dr\Bigg)\\
&\quad + \sum_{k=0}^{n-1} \tfrac{4L^2T}{M^{n-(k+1)}} \int_t^T \int_{\mathbb{R}^d}\mathbb{E}\Big[\big|u(r, X^0_{t,r}(z)) - V^0_{M,k}(r, X^0_{t,r}(z))\big|^2\Big]\ \mu_t(dz)\, dr.
\end{aligned}\tag{3.79}$$

Moreover, observe that (3.77), (3.78), the fact that X^0 and X^1 are independent and continuous random fields, (3.3), and Lemma 2.16 imply that for all $t \in [0,T]$ it holds that

$$\begin{aligned}
&\int_{\mathbb{R}^d} \mathbb{E}\big[|g(X^0_{t,T}(z))|^2\big]\ \mu_t(dz) + T\int_t^T\int_{\mathbb{R}^d} \mathbb{E}\big[|f(r, X^0_{t,r}(z), 0)|^2\big]\ \mu_t(dz)\, dr\\
&= \mathbb{E}\big[|g(X^0_{t,T}(X^1_{0,t}(\xi)))|^2\big] + T\int_t^T \mathbb{E}\big[|f(r, X^0_{t,r}(X^1_{0,t}(\xi)), 0)|^2\big]\ dr\\
&= \mathbb{E}\big[|g(X^0_{0,T}(\xi))|^2\big] + T\int_t^T \mathbb{E}\big[|f(r, X^0_{0,r}(\xi), 0)|^2\big]\ dr \leq C^2.
\end{aligned}\tag{3.80}$$

In addition, note that (3.77), items (i)–(ii) in Lemma 3.6, the hypothesis that $(X^\theta)_{\theta\in\Theta}$ are independent, the hypothesis that $(\mathcal{R}^\theta)_{\theta\in\Theta}$ are independent, and the hypothesis that $(X^\theta)_{\theta\in\Theta}$ and $(\mathcal{R}^\theta)_{\theta\in\Theta}$ are independent, (3.3), Lemma 2.16, and Lemma 3.5 assure that for all $n \in \mathbb{N}_0$, $t \in [0,T]$, $r \in [t,T]$ it holds that

$$\begin{aligned}
&\int_{\mathbb{R}^d} \mathbb{E}\Big[\big|u(r, X^0_{t,r}(z)) - V^0_{M,n}(r, X^0_{t,r}(z))\big|^2\Big]\ \mu_t(dz)\\
&= \mathbb{E}\Big[\big|u(r, X^0_{t,r}(X^1_{0,t}(\xi))) - V^0_{M,n}(r, X^0_{t,r}(X^1_{0,t}(\xi)))\big|^2\Big]\\
&= \mathbb{E}\Big[\big|u(r, X^0_{0,r}(\xi)) - V^0_{M,n}(r, X^0_{0,r}(\xi))\big|^2\Big].
\end{aligned}\tag{3.81}$$

Combining this with (3.79) and (3.80) ensures that for all $n \in \mathbb{N}$, $t \in [0,T]$ it holds that

$$\begin{aligned}
&\mathbb{E}\big[|u(t, X^0_{0,t}(\xi)) - V^0_{M,n}(t, X^0_{0,t}(\xi))|^2\big]\\
&\leq \frac{C^2}{M^n} + \sum_{k=0}^{n-1} \tfrac{4L^2T}{M^{n-(k+1)}} \int_t^T \mathbb{E}\Big[\big|u(r, X^0_{0,r}(\xi)) - V^0_{M,k}(r, X^0_{0,r}(\xi))\big|^2\Big]\ dr.
\end{aligned}\tag{3.82}$$

The fact that $\mathbb{P}(X^0_{0,0}(\xi) = \xi) = 1$, the fact that for all $n \in \mathbb{N}$ it holds that $V^0_{M,n}$, X^0, and $\mathcal{R}^0$ are independent, Lemma 3.5, and (3.75) hence imply that for all $n \in \mathbb{N}$ it holds that

$$\begin{aligned}
\epsilon_{n,0} &= \mathbb{E}\big[|u(0, X^0_{0,0}(\xi)) - V^0_{M,n}(0, X^0_{0,0}(\xi))|^2\big]\\
&\leq \frac{C^2}{M^n} + \sum_{k=0}^{n-1} \tfrac{4L^2T^2}{M^{n-(k+1)}T} \int_0^T \mathbb{E}\Big[\big|u(r, X^0_{0,r}(\xi)) - V^0_{M,k}(r, X^0_{0,r}(\xi))\big|^2\Big]\ dr\\
&= \frac{C^2}{M^n(0!)} + 4L^2T^2\left[\sum_{k=0}^{n-1} \frac{\epsilon_{k,1}}{M^{n-(k+1)}}\right].
\end{aligned}\tag{3.83}$$

Moreover, observe that Lemma 3.14 (with $T = T$, $q = q$, $(U(r))_{r\in[0,T]} = (\mathbb{E}[|u(r, X^0_{0,r}(\xi)) - V^0_{M,n}(r, X^0_{0,r}(\xi))|^2])_{r\in[0,T]}$ for $n \in \mathbb{N}_0$, $q \in \mathbb{N}$ in the notation of Lemma 3.14) and (3.76) demonstrate that for all $n \in \mathbb{N}_0$, $q \in \mathbb{N}$ it holds that

$$\begin{aligned}
&\frac{1}{T^q}\left(\int_0^T \frac{t^{q-1}}{(q-1)!}\int_t^T \mathbb{E}\big[\big|u(r, X^0_{0,r}(\xi)) - V^0_{M,n}(r, X^0_{0,r}(\xi))\big|^2\big]\,dr\,dt\right)\\
&= \frac{T}{T^{q+1}}\left(\int_0^T \frac{t^q}{q!}\mathbb{E}\big[\big|u(t, X^0_{0,t}(\xi)) - V^0_{M,n}(t, X^0_{0,t}(\xi))\big|^2\big]\,dt\right)\\
&= T\epsilon_{n,q+1}.
\end{aligned}\tag{3.84}$$

This, (3.76), and (3.82) imply that for all $n, q \in \mathbb{N}$ it holds that

$$\begin{aligned}
\epsilon_{n,q} &= \frac{1}{T^q}\left(\int_0^T \frac{t^{q-1}}{(q-1)!}\,\mathbb{E}\big[|u(t, X^0_{0,t}(\xi)) - V^0_{M,n}(t, X^0_{0,t}(\xi))|^2\big]\;dt\right)\\
&\le \frac{C^2}{T^q M^n}\left(\int_0^T \frac{t^{q-1}}{(q-1)!}\,dt\right)\\
&\quad + \sum_{k=0}^{n-1}\tfrac{4L^2T}{M^{n-(k+1)}T^q}\left(\int_0^T \frac{t^{q-1}}{(q-1)!}\int_t^T \mathbb{E}\Big[\big|u(r, X^0_{0,r}(\xi)) - V^0_{M,k}(r, X^0_{0,r}(\xi))\big|^2\Big]\;dr\,dt\right)\\
&= \frac{C^2}{T^q M^n}\frac{T^q}{q!} + 4L^2T\left[\sum_{k=0}^{n-1}\frac{T\epsilon_{k,q+1}}{M^{n-(k+1)}}\right] = \frac{C^2}{M^n(q!)} + 4L^2T^2\left[\sum_{k=0}^{n-1}\frac{\epsilon_{k,q+1}}{M^{n-(k+1)}}\right].
\end{aligned}\tag{3.85}$$

Furthermore, note the fact that $\int_0^T \big(\mathbb{E}\big[|u(r, X^0_{0,r}(\xi))|^2\big]\big)^{1/2}\,dr < \infty$ and Lemma 3.3 prove that

$$\sup_{t\in[0,T]} \mathbb{E}\big[|u(t, X^0_{0,t}(\xi))|^2\big] \le C^2.\tag{3.86}$$

The fact that $\mathbb{P}(X^0_{0,0}(\xi) = \xi) = 1$ and the fact that $V^0_{M,0} = 0$ hence assure that

$$\epsilon_{0,0} = |u(0,\xi)|^2 = \mathbb{E}\big[|u(0, X^0_{0,0}(\xi))|^2\big] \le C^2 = \frac{C^2}{M^0 0!}.\tag{3.87}$$

Moreover, observe that (3.86) and the fact that $V^0_{M,0} = 0$ ensure that for all $q \in \mathbb{N}$ it holds that

$$\epsilon_{0,q} = \frac{1}{T^q}\int_0^T \frac{t^{q-1}}{(q-1)!}\,\mathbb{E}\big[|u(t, X^0_{0,t}(\xi))|^2\big]\;dt \le \frac{C^2}{T^q}\int_0^T \frac{t^{q-1}}{(q-1)!}\,dt = \frac{C^2}{T^q}\frac{T^q}{q!} = \frac{C^2}{M^0(q!)}.\tag{3.88}$$

Combining this, (3.83), (3.85), and (3.87) demonstrates that for all $n, q \in \mathbb{N}_0$ it holds that

$$\begin{aligned}
\epsilon_{n,q} &\le \frac{C^2}{M^n(q)!} + 4L^2T^2\left[\sum_{k=0}^{n-1}\frac{\epsilon_{k,q+1}}{M^{n-(k+1)}}\right]\\
&= \frac{C^2M^q}{M^{n+q}(q)!} + 4L^2T^2\left[\sum_{k=0}^{n-1}\frac{\epsilon_{k,q+1}}{M^{n-(k+1)}}\right] \le \frac{C^2\exp(M)}{M^{n+q}} + 4L^2T^2\left[\sum_{k=0}^{n-1}\frac{\epsilon_{k,q+1}}{M^{n-(k+1)}}\right].
\end{aligned}\tag{3.89}$$

Lemma 3.12 (with $\alpha = C^2 \exp(M)$, $\beta = 4L^2T^2$, $M = M$, $(\epsilon_{n,q})_{n,q\in\mathbb{N}_0} = (\epsilon_{n,q})_{n,q\in\mathbb{N}_0}$ in the notation of Lemma 3.12) therefore proves that for all $n, q \in \mathbb{N}_0$ it holds that

$$\epsilon_{n,q} \leq \frac{C^2 \exp(M)(1 + 4L^2T^2)^n}{M^{n+q}}. \tag{3.90}$$

This implies that for all $n \in \mathbb{N}_0$ it hold that

$$\mathbb{E}\big[|u(0,\xi) - V^0_{M,n}(0,\xi)|^2\big] = \epsilon_{n,0} \leq \frac{C^2(1 + 4L^2T^2)^n \exp(M)}{M^n}. \tag{3.91}$$

The fact that for all $x, y \in [0,\infty)$ it holds that $\sqrt{x+y} \leq \sqrt{x} + \sqrt{y}$ hence demonstrates that for all $n \in \mathbb{N}_0$ it holds that

$$\big(\mathbb{E}\big[|u(0,\xi) - V^0_{M,n}(0,\xi)|^2\big]\big)^{1/2} \leq \frac{C(\sqrt{1+4L^2T^2})^n \exp(\frac{M}{2})}{M^{n/2}} \leq \frac{C(1+2LT)^n \exp(\frac{M}{2})}{M^{n/2}}. \tag{3.92}$$

The proof of Proposition 3.15 is thus completed. □

Corollary 3.16. *Assume Setting 3.1, let $\xi \in \mathbb{R}^d$, $C \in [0,\infty]$ satisfy that*

$$C = \left[\big(\mathbb{E}\big[|g(X^0_{0,T}(\xi))|^2\big]\big)^{1/2} + \sqrt{T}\left(\int_0^T \mathbb{E}\big[|f(t, X^0_{0,t}(\xi), 0)|^2\big]\, dt\right)^{1/2}\right] \exp(LT), \tag{3.93}$$

and assume for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that $\int_0^T \big(\mathbb{E}\big[|u(r, X^0_{0,r}(\xi))|^2\big]\big)^{1/2} dr + \int_t^T \mathbb{E}\big[|f(r, X^0_{t,r}(x), 0)|\big]\, dr < \infty$. Then it holds for all $N \in \mathbb{N}$ that

$$\big(\mathbb{E}\big[|u(0,\xi) - V^0_{N,N}(0,\xi)|^2\big]\big)^{1/2} \leq C\left[\frac{\sqrt{e}(1+2LT)}{\sqrt{N}}\right]^N. \tag{3.94}$$

Proof of Corollary 3.16. Proposition 3.15 establishes Corollary 3.16. The proof of Corollary 3.16 is thus completed. □

3.5 Complexity analysis for MLP approximation algorithms

In this subsection we consider the computational effort of the MLP scheme (cf. (3.5) in Setting 3.1 above) introduced in Setting 3.1 and combine it with the L^2-error estimate in Corollary 3.16 to obtain a complexity analysis for the MLP scheme in Proposition 3.18 below. In Lemma 3.17 we think for all $M, n \in \mathbb{N}$ of $\mathcal{C}_{M,n}$ as the number of realizations of 1-dimensional random variables needed to simulate one realization of $V^\theta_{M,n}(t,x)$ for any $\theta \in \Theta$, $t \in [0,T]$, $x \in \mathbb{R}^d$. The recursive inequality in (3.95) in Lemma 3.17 is based on (3.5) and the assumption that the number of realizations of 1-dimensional random variables needed to simulate $X^\theta_{t,r}(x)$ for any $\theta \in \Theta$, $t \in [0,T]$, $r \in [t,T]$, $x \in \mathbb{R}^d$ is bounded by αd.

Lemma 3.17. *Let $d \in \mathbb{N}$, $\alpha \in [1/d, \infty)$, $(\mathcal{C}_{M,n})_{M,n\in\mathbb{Z}} \subseteq [0,\infty)$ satisfy for all $n, M \in \mathbb{N}$ that $\mathcal{C}_{M,0} = 0$ and*

$$\mathcal{C}_{M,n} \leq \alpha d M^n + \sum_{k=0}^{n-1} \Big[M^{(n-k)}(\alpha d + 1 + \mathcal{C}_{M,k} + \mathbb{1}_{\mathbb{N}}(k)\mathcal{C}_{M,k-1})\Big]. \tag{3.95}$$

Then it holds for all $n, M \in \mathbb{N}$ that $\mathcal{C}_{M,n} \leq \alpha d\,(5M)^n$.

Proof of Lemma 3.17. First, observe that (3.95) and the hypothesis that for all $M \in \mathbb{N}$ it holds that $\mathcal{C}_{M,0} = 0$ imply that for all $n \in \mathbb{N}$, $M \in \mathbb{N} \cap [2, \infty)$ it holds that

$$\begin{aligned}
(M^{-n}\mathcal{C}_{M,n}) &\leq \alpha d + \sum_{k=0}^{n-1} \left[M^{-k}(\alpha d + 1 + \mathcal{C}_{M,k} + \mathbb{1}_{\mathbb{N}}(k)\mathcal{C}_{M,k-1})\right] \\
&= \alpha d + (\alpha d + 1) \left[\sum_{k=0}^{n-1} M^{-k}\right] + \left[\sum_{k=0}^{n-1} M^{-k}\mathcal{C}_{M,k}\right] + \left[\sum_{k=0}^{n-2} M^{-(k+1)}\mathcal{C}_{M,k}\right] \\
&= \alpha d + (\alpha d + 1)\tfrac{(1-M^{-n})}{(1-M^{-1})} + \left[\sum_{k=0}^{n-1} M^{-k}\mathcal{C}_{M,k}\right] + \frac{1}{M}\left[\sum_{k=0}^{n-2} M^{-k}\mathcal{C}_{M,k}\right] \\
&\leq \alpha d + (\alpha d + 1)\tfrac{1}{(1-\frac{1}{2})} + \left(1 + \tfrac{1}{M}\right)\left[\sum_{k=0}^{n-1} M^{-k}\mathcal{C}_{M,k}\right] \\
&= 3\alpha d + 2 + \left(1 + \tfrac{1}{M}\right)\left[\sum_{k=1}^{n-1} M^{-k}\mathcal{C}_{M,k}\right].
\end{aligned} \tag{3.96}$$

The discrete Gronwall inequality in Corollary 2.3 (with $N = \infty$, $\alpha = 3\alpha d + 2$, $\beta = \left(1 + \frac{1}{M}\right)$, $(\epsilon_n)_{n\in\mathbb{N}_0} = (M^{-(n+1)}\mathcal{C}_{M,(n+1)})_{n\in\mathbb{N}_0}$ for $M \in \mathbb{N} \cap [2, \infty)$ in the notation of Corollary 2.3) hence ensures that for all $n \in \mathbb{N}_0$, $M \in \mathbb{N} \cap [2, \infty)$ it holds that

$$(M^{-(n+1)}\mathcal{C}_{M,n+1}) \leq (3\alpha d + 2)\left(2 + \tfrac{1}{M}\right)^n. \tag{3.97}$$

This establishes that for all $n \in \mathbb{N}$, $M \in \mathbb{N} \cap [2, \infty)$ it holds that

$$\mathcal{C}_{M,n} \leq (3\alpha d + 2)\left(2 + \tfrac{1}{M}\right)^{n-1} M^n \leq (5\alpha d)3^{n-1}M^n \leq \alpha d(5M)^n. \tag{3.98}$$

Moreover, observe that the fact that $\mathcal{C}_{1,0} = 0$ and (3.95) demonstrate that for all $n \in \mathbb{N}$ it holds that

$$\mathcal{C}_{1,n} \leq \alpha d + \sum_{k=0}^{n-1}(\alpha d + 1 + \mathcal{C}_{1,k} + \mathbb{1}_{\mathbb{N}}(k)\mathcal{C}_{1,k-1}) \leq \alpha d + n(\alpha d + 1) + 2\sum_{k=1}^{n-1}\mathcal{C}_{1,k}. \tag{3.99}$$

Hence, we obtain for all $n \in \mathbb{N}$, $k \in \mathbb{N} \cap (0, n]$ that

$$\mathcal{C}_{1,k} \leq \alpha d + k(\alpha d + 1) + 2\sum_{l=1}^{k-1}\mathcal{C}_{1,l} \leq \alpha d + n(\alpha d + 1) + 2\sum_{l=1}^{k-1}\mathcal{C}_{1,l}. \tag{3.100}$$

Combining this with the discrete Gronwall inequality in Corollary 2.3 (with $N = n - 1$, $\alpha = \alpha d + n(\alpha d + 1)$, $\beta = 2$, $(\epsilon_k)_{k\in\mathbb{N}_0\cap[0,N]} = (\mathcal{C}_{1,k+1})_{k\in\mathbb{N}_0\cap[0,n)}$ for $n \in \mathbb{N}$ in the notation of Corollary 2.3) proves that for all $n \in \mathbb{N}$, $k \in \mathbb{N}_0 \cap [0, n)$ it holds that

$$\mathcal{C}_{1,k+1} \leq (\alpha d + n(\alpha d + 1))3^k. \tag{3.101}$$

The fact that for all $n \in \mathbb{N}$ it holds that $(1 + 2n)3^{n-1} \leq 5^n$ hence shows that for all $n \in \mathbb{N}$ it holds that

$$\mathcal{C}_{1,n} \leq (\alpha d + n(\alpha d + 1))3^{n-1} = \alpha d\left(1 + n\left(1 + \tfrac{1}{\alpha d}\right)\right)3^{n-1} \leq \alpha d\,(1 + 2n)3^{n-1} \leq \alpha d\,5^n. \tag{3.102}$$

Combining this with (3.98) completes the proof of Lemma 3.17. □

Proposition 3.18. *Assume Setting 3.1, let $\xi \in \mathbb{R}^d$, $C \in [0,\infty)$, $\alpha \in [1,\infty)$, $(\mathcal{C}_{M,n})_{M,n\in\mathbb{Z}} \subseteq \mathbb{N}_0$ satisfy for all $n, M \in \mathbb{N}$ that*

$$\mathcal{C}_{M,0} = 0, \qquad \mathcal{C}_{M,n} \leq \alpha d M^n + \sum_{k=0}^{n-1} \left[M^{(n-k)} (\alpha d + 1 + \mathcal{C}_{M,k} + \mathbb{1}_{\mathbb{N}}(k) \mathcal{C}_{M,k-1}) \right], \tag{3.103}$$

$$\text{and} \qquad C = \left[\left(\mathbb{E}\big[|g(X^0_{0,T}(\xi))|^2 \big] \right)^{1/2} + \sqrt{T} \left(\int_0^T \mathbb{E}\big[|f(t, X^0_{0,t}(\xi), 0)|^2 \big] \, dt \right)^{1/2} \right] \exp(LT), \tag{3.104}$$

and assume for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that $\int_0^T \left(\mathbb{E}\big[|u(r, X^0_{0,r}(\xi))|^2 \big] \right)^{1/2} dr + \int_t^T \mathbb{E}\big[|f(r, X^0_{t,r}(x), 0)| \big] \, dr < \infty$. Then there exists a function $N\colon (0,\infty) \to \mathbb{N}$ such that for all $\varepsilon, \delta \in (0,\infty)$ it holds that

$$\left(\mathbb{E}\big[|u(0,\xi) - V^0_{N_\varepsilon, N_\varepsilon}(0,\xi)|^2 \big] \right)^{1/2} \leq \varepsilon \qquad \textit{and} \tag{3.105}$$

$$\mathcal{C}_{N_\varepsilon, N_\varepsilon} \leq \alpha \, d \, \max\{1, C^{2+\delta}\} \left[\sup_{n\in\mathbb{N}} \tfrac{(4+8LT)^{(3+\delta)(n+1)}}{n^{(n\delta/2)}} \right] (\min\{1,\varepsilon\})^{-(2+\delta)} < \infty. \tag{3.106}$$

Proof of Proposition 3.18. Throughout this proof let $\kappa \in (0,\infty)$ be given by

$$\kappa = \sqrt{e}(1 + 2LT), \tag{3.107}$$

let $N\colon (0,\infty) \to \mathbb{N}$ be the function which satisfies for all $\varepsilon \in (0,\infty)$ that

$$N_\varepsilon = \min \left\{ n \in \mathbb{N} \colon C \left[\frac{\kappa}{\sqrt{n}} \right]^n \leq \varepsilon \right\}, \tag{3.108}$$

and let $\delta \in (0,\infty)$. Note that (3.108) and Corollary 3.16 assure that for all $\varepsilon \in (0,\infty)$ it holds that

$$\left(\mathbb{E}\big[|u(0,\xi) - V^0_{N_\varepsilon, N_\varepsilon}(0,\xi)|^2 \big] \right)^{1/2} \leq C \left[\frac{\kappa}{\sqrt{N_\varepsilon}} \right]^{N_\varepsilon} \leq \varepsilon. \tag{3.109}$$

Moreover, observe that (3.108) ensures that for all $\varepsilon \in (0,\infty)$ with $N_\varepsilon \geq 2$ it holds that

$$C \left[\frac{\kappa}{\sqrt{N_\varepsilon - 1}} \right]^{N_\varepsilon - 1} > \varepsilon. \tag{3.110}$$

Lemma 3.17 and (3.103) hence show that for all $\varepsilon \in (0,\infty)$ with $N_\varepsilon \geq 2$ it holds that

$$\begin{aligned}
\mathcal{C}_{N_\varepsilon, N_\varepsilon} &\leq \alpha \, d \, (5N_\varepsilon)^{N_\varepsilon} \leq \alpha \, d \, (5N_\varepsilon)^{N_\varepsilon} \left[C \left[\frac{\kappa}{\sqrt{N_\varepsilon - 1}} \right]^{N_\varepsilon - 1} \varepsilon^{-1} \right]^{2+\delta} \\
&= \alpha \, d \, C^{2+\delta} \varepsilon^{-(2+\delta)} \left[\frac{(5N_\varepsilon)^{N_\varepsilon} \kappa^{(N_\varepsilon - 1)(2+\delta)}}{(N_\varepsilon - 1)^{(N_\varepsilon - 1)(1+\delta/2)}} \right] \\
&\leq \alpha \, d \, C^{2+\delta} \varepsilon^{-(2+\delta)} \sup_{n \in \mathbb{N}\cap[2,\infty)} \left[\frac{(5n)^n \kappa^{(n-1)(2+\delta)}}{(n-1)^{(n-1)(1+\delta/2)}} \right].
\end{aligned} \tag{3.111}$$

Next note that for all $n \in \mathbb{N} \cap [2,\infty)$ it holds that

$$\frac{n^n}{(n-1)^{(n-1)}} = \left(\frac{n}{n-1} \right)^{n-1} n = \left(1 + \frac{1}{n-1} \right)^{n-1} n \leq e \, n. \tag{3.112}$$

Furthermore, observe that the fact that $\kappa \geq \sqrt{e}$ and the fact that $\sqrt{5e} \leq 4$ imply that for all $n \in \mathbb{N} \cap [2, \infty)$ it holds that

$$\begin{aligned}
5^n \, e \, \kappa^{(n-1)(2+\delta)} &\leq (\sqrt{5})^{n(2+\delta)} \, (\sqrt{e})^{2+\delta} \, \kappa^{(n-1)(2+\delta)} \\
&\leq (\sqrt{5})^{n(2+\delta)} \, \kappa^{2+\delta} \, \kappa^{(n-1)(2+\delta)} \\
&= (\sqrt{5}\kappa)^{n(2+\delta)} \\
&= (\sqrt{5e}(1 + 2LT))^{n(2+\delta)} \\
&\leq (4(1 + 2LT))^{n(2+\delta)} = (4 + 8LT)^{n(2+\delta)}.
\end{aligned} \tag{3.113}$$

Combining this, (3.112), and the fact that for all $n \in \mathbb{N}$ it holds that $n \leq (4 + 8LT)^n$ demonstrates that

$$\begin{aligned}
\sup_{n \in \mathbb{N} \cap [2,\infty)} \left[\frac{(5n)^n \kappa^{(n-1)(2+\delta)}}{(n-1)^{(n-1)(1+\delta/2)}} \right] &= \sup_{n \in \mathbb{N} \cap [2,\infty)} \left[\frac{n^n}{(n-1)^{(n-1)}} \frac{5^n \kappa^{(n-1)(2+\delta)}}{(n-1)^{((n-1)\delta)/2}} \right] \\
&\leq \sup_{n \in \mathbb{N} \cap [2,\infty)} \left[\frac{e \, n \, 5^n \kappa^{(n-1)(2+\delta)}}{(n-1)^{((n-1)\delta)/2}} \right] \\
&\leq \sup_{n \in \mathbb{N} \cap [2,\infty)} \left[\frac{n(4 + 8LT)^{n(2+\delta)}}{(n-1)^{((n-1)\delta)/2}} \right] \\
&\leq \sup_{n \in \mathbb{N} \cap [2,\infty)} \left[\frac{(4 + 8LT)^{n(3+\delta)}}{(n-1)^{((n-1)\delta)/2}} \right].
\end{aligned} \tag{3.114}$$

In addition, observe that

$$\begin{aligned}
\sup_{n \in \mathbb{N} \cap [2,\infty)} \left[\frac{(4 + 8LT)^{n(3+\delta)}}{(n-1)^{((n-1)\delta)/2}} \right] &= (4 + 8LT)^{3+\delta} \sup_{n \in \mathbb{N}} \left[\frac{(4 + 8LT)^{n(3+\delta)}}{n^{(n\delta)/2}} \right] \\
&= (4 + 8LT)^{3+\delta} \sup_{n \in \mathbb{N}} \left[\frac{(4 + 8LT)^{(3+\delta)}}{n^{\delta/2}} \right]^n < \infty.
\end{aligned} \tag{3.115}$$

This, (3.111), and (3.114) prove that for all $\varepsilon \in (0, \infty)$ with $N_\varepsilon \geq 2$ it holds that

$$\mathcal{C}_{N_\varepsilon, N_\varepsilon} \leq \alpha \, d \, C^{2+\delta} \varepsilon^{-(2+\delta)} \sup_{n \in \mathbb{N}} \left[\frac{(4 + 8LT)^{(n+1)(3+\delta)}}{n^{(n\delta)/2}} \right] < \infty. \tag{3.116}$$

Next note that the hypothesis that $\mathcal{C}_{1,0} = 0$, (3.103), and the fact that $3 \leq \sup_{n \in \mathbb{N}} \left[\frac{(4+8LT)^{(n+1)(3+\delta)}}{n^{(n\delta)/2}} \right] < \infty$ assure that for all $\varepsilon \in (0, \infty)$ with $N_\varepsilon = 1$ it holds that

$$\begin{aligned}
\mathcal{C}_{N_\varepsilon, N_\varepsilon} &= \mathcal{C}_{1,1} \leq 2\alpha d + 1 \leq 3\alpha d \\
&\leq \alpha \, d \, \max\{1, C^{2+\delta}\} \left[\sup_{n \in \mathbb{N}} \frac{(4 + 8LT)^{(n+1)(3+\delta)}}{n^{(n\delta)/2}} \right] (\min\{1, \varepsilon\})^{-(2+\delta)} < \infty.
\end{aligned} \tag{3.117}$$

This and (3.116) demonstrate that for all $\varepsilon \in (0, \infty)$ it holds that

$$\mathcal{C}_{N_\varepsilon, N_\varepsilon} \leq \alpha \, d \, \max\{1, C^{2+\delta}\} \left[\sup_{n \in \mathbb{N}} \frac{(4 + 8LT)^{(n+1)(3+\delta)}}{n^{(n\delta)/2}} \right] (\min\{1, \varepsilon\})^{-(2+\delta)} < \infty. \tag{3.118}$$

Combining this with (3.109) completes the proof of Proposition 3.18. ☐

3.6 MLP approximations for semilinear partial differential equations (PDEs)

In order to carry over the complexity analysis for the approximation of solutions of stochastic fixed point equations in Subsection 3.5 to the approximation of solutions of semilinear Kolmogorov PDEs we present in Subsection 3.6.1 a result establishing a one to one correspondence between viscosity solutions of semilinear Kolmogorov PDEs and solutions of stochastic fixed point equations (see Proposition 3.22 in Subsection 3.6.1 below). This correspondence combined with the complexity analysis in Proposition 3.18 yields Proposition 3.23 in Subsection 3.6.2 below. Considering this complexity analysis over variable dimensions shows that our proposed MLP algorithm overcomes the curse of dimensionality in the approximation of solutions of certain semilinear Kolmogorov PDEs (see Theorem 3.24 in Subsection 3.6.3 below, the main result of this paper, for details).

3.6.1 Viscosity solutions of semilinear PDEs and stochastic fixed point equations

In this subsection we briefly recall the notion of viscosity solutions in the case of a suitable class of parabolic PDEs (see, e.g., Crandall, Ishii, & Lions [27] and cf. Definition 3.21 below) and we display in Proposition 3.22 a well-known correspondence between viscosity solutions of semilinear Kolmogorov PDEs and solutions of stochastic fixed point equations. The following presentations are modified extracts from Beck et al. [5].

Definition 3.19 (Symmetric matrices)**.** *Let $d \in \mathbb{N}$. Then we denote by $\mathbb{S}_d \subseteq \mathbb{R}^{d\times d}$ the set given by*

$$\mathbb{S}_d = \{A \in \mathbb{R}^{d\times d} \colon A = A^*\} \tag{3.119}$$

and we call $\mathbb{S}_d$ the set of all real symmetric $(d \times d)$-matrices (we call $\mathbb{S}_d$ the set of all symmetric $(d \times d)$-matrices).

Definition 3.20 (Degenerate elliptic functions)**.** *Let $d \in \mathbb{N}$, $T \in (0,\infty)$, let $O \subseteq \mathbb{R}^d$ be an open set, and let $F \in C((0,T) \times O \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}_d, \mathbb{R})$ (cf. Definition 3.19). Then we say that F is a degenerate elliptic function on $(0,T) \times O \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}_d$ (we say that F is a degenerate elliptic function) if and only if for all $t \in (0,T)$, $x \in O$, $r \in \mathbb{R}$, $p \in \mathbb{R}^d$, $A, B \in \mathbb{S}_d$ with $\forall\, v \in \mathbb{R}^d \colon v^*(Av) \le v^*(Bv)$ it holds that*

$$F(t,x,r,p,A) \le F(t,x,r,p,B). \tag{3.120}$$

Definition 3.21 (Viscosity solution)**.** *Let $d \in \mathbb{N}$, $T \in (0,\infty)$, let $O \subseteq \mathbb{R}^d$ be a non-empty open set, and let $F \in C((0,T) \times O \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}_d, \mathbb{R})$ be a degenerate elliptic function (cf. Definitions 3.19 and 3.20). Then we say that u is a viscosity solution of*

$$\big(\tfrac{\partial}{\partial t}u\big)(t,x) + F(t,x,u(t,x),(\nabla_x u)(t,x),(\operatorname{Hess}_x u)(t,x)) = 0 \tag{3.121}$$

for $(t,x) \in (0,T) \times O$ if and only if $u \in C((0,T) \times O, \mathbb{R})$ is a continuous function from $(0,T) \times O$ to $\mathbb{R}$ which satisfies for all $t \in (0,T)$, $x \in O$, $\varphi \in C^{1,2}((0,T) \times O, \mathbb{R})$, $k \in \{0,1\}$ with $\varphi(t,x) = u(t,x)$ and $(-1)^k(\varphi - u) \le 0$ that

$$(-1)^k \left[\big(\tfrac{\partial}{\partial t}\varphi\big)(t,x) + F(t,x,\varphi(t,x),(\nabla_x \varphi)(t,x),(\operatorname{Hess}_x \varphi)(t,x))\right] \le 0. \tag{3.122}$$

Proposition 3.22. *Let $d, m \in \mathbb{N}$, $L, T \in (0,\infty)$, let $\langle \cdot,\cdot \rangle \colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be the Euclidean scalar product on $\mathbb{R}^d$, let $\|\cdot\| \colon \mathbb{R}^d \to [0,\infty)$ be the Euclidean norm on $\mathbb{R}^d$, let $|||\cdot||| \colon \mathbb{R}^{d\times m} \to [0,\infty)$ be the Frobenius norm on $\mathbb{R}^{d\times m}$, let $\mu \colon [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma \colon [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ be globally Lipschitz continuous functions, let $f \in C([0,T] \times \mathbb{R}^d \times \mathbb{R}, \mathbb{R})$, $g \in C(\mathbb{R}^d, \mathbb{R})$ be at most polynomially growing functions, assume for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $v, w \in \mathbb{R}$ that*

$$|f(t,x,v) - f(t,x,w)| \le L|v - w|, \tag{3.123}$$

let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_t)_{t\in[0,T]})$ be a filtered probability space which satisfies the usual conditions, and let $W \colon [0,T] \times \Omega \to \mathbb{R}^d$ be a standard $(\mathbb{F}_t)_{t\in[0,T]}$-Brownian motion. Then *(i) for every $t \in [0,T]$, $x \in \mathbb{R}^d$ there exists an up to indistinguishability unique $(\mathbb{F}_s)_{s\in[t,T]}$-adapted stochastic process $(X_{t,s}(x))_{s\in[t,T]}\colon [t,T]\times\Omega \to \mathbb{R}^d$ with continuous sample paths which satisfies that for all $s \in [t,T]$ it holds $\mathbb{P}$-a.s. that*

$$X_{t,s}(x) = x + \int_t^s \mu(r, X_{t,r}(x))\,dr + \int_t^s \sigma(r, X_{t,r}(x))\,dW_r, \tag{3.124}$$

(ii) there exists a unique at most polynomially growing function $u \in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ which satisfies that $u|_{(0,T)\times\mathbb{R}^d}\colon (0,T)\times\mathbb{R}^d \to \mathbb{R}$ is a viscosity solution of

$$\begin{aligned}(\tfrac{\partial u}{\partial t})(t,x) + \tfrac{1}{2}\operatorname{Trace}\big(\sigma(t,x)[\sigma(t,x)]^*(\operatorname{Hess}_x u)(t,x)\big)&\\ + \langle \mu(t,x), (\nabla_x u)(t,x)\rangle + f(t,x,u(t,x)) &= 0\end{aligned} \tag{3.125}$$

for $(t,x) \in (0,T)\times\mathbb{R}^d$ (cf. Definition 3.21) and which satisfies for all $x \in \mathbb{R}^d$ that $u(T,x) = g(x)$,

(iii) there exists a unique at most polynomially growing function $v \in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ which satisfies for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that $\mathbb{E}\big[|g(X_{t,T}(x))| + \int_t^T |f(s, X_{t,s}(x), v(s, X_{t,s}(x)))|\,ds\big] < \infty$ and

$$v(t,x) = \mathbb{E}\left[g\big(X_{t,T}(x)\big) + \int_t^T f\big(s, X_{t,s}(x), v(s, X_{t,s}(x))\big)\,ds\right], \tag{3.126}$$

and

(iv) it holds for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that $u(t,x) = v(t,x)$.

Proof of Proposition 3.22. First, observe that the hypothesis that μ and σ are globally Lipschitz continuous functions implies that

$$\sup_{t\in[0,T],x\in\mathbb{R}^d} \frac{|\langle x, \mu(t,x)\rangle| + |||\sigma(t,x)|||^2}{1 + \|x\|^2} < \infty. \tag{3.127}$$

Combining this, the hypothesis that f and g are at most polynomially growing functions, the hypothesis that μ and σ are globally Lipschitz continuous functions, and (3.123) with Beck et al. [5, Corollary 3.9] establishes items (i)–(iv). The proof of Proposition 3.22 is thus completed. ☐

3.6.2 MLP approximations in fixed space dimensions

Proposition 3.23. *Let $d, m \in \mathbb{N}$, $T \in (0,\infty)$, $L, K, p, C_1, C_2, \mathfrak{C} \in [0,\infty)$, $\alpha \in [1,\infty)$, $\xi \in \mathbb{R}^d$, $\Theta = \cup_{n=1}^\infty \mathbb{Z}^n$, let $\langle\cdot,\cdot\rangle : \mathbb{R}^d\times\mathbb{R}^d \to \mathbb{R}$ be the Euclidean scalar product on $\mathbb{R}^d$, let $\|\cdot\| : \mathbb{R}^d \to [0,\infty)$ be the Euclidean norm on $\mathbb{R}^d$, let $|||\cdot|||\colon \mathbb{R}^{d\times m} \to [0,\infty)$ be the Frobenius norm on $\mathbb{R}^{d\times m}$, assume that*

$$\mathfrak{C} = 4Ke^{T(L+2+p(p+2)(C_2+1))}\left((1+\|\xi\|^2)^{p/2} + (2p+1)|C_1|^{p/2}\right), \tag{3.128}$$

let $g \in C(\mathbb{R}^d,\mathbb{R})$, $f \in C([0,T]\times\mathbb{R}^d\times\mathbb{R},\mathbb{R})$ satisfy for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $v, w \in \mathbb{R}$ that

$$\max\{|g(x)|, |f(t,x,0)|\} \le K(1+\|x\|^p) \quad \text{and} \quad |f(t,x,v) - f(t,x,w)| \le L|v-w|, \tag{3.129}$$

let $\mu\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ and $\sigma\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}^{d\times m}$ be globally Lipschitz continuous functions which satisfy for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$\max\{\langle x, \mu(t,x)\rangle, |||\sigma(t,x)|||^2\} \le C_1 + C_2\|x\|^2, \tag{3.130}$$

let $(\Omega, \mathcal{F}, \mathbb{P})$ *be a complete probability space, let* $\mathcal{R}^\theta\colon \Omega \to [0,1]$*,* $\theta \in \Theta$*, be independent* $\mathcal{U}_{[0,1]}$*-distributed random variables, let* $R^\theta = (R^\theta_t)_{t\in[0,T]}\colon [0,T]\times\Omega \to [0,T]$*,* $\theta\in\Theta$*, be the stochastic processes which satisfy for all* $t\in[0,T]$*,* $\theta\in\Theta$ *that*

$$R^\theta_t = t + (T-t)\mathcal{R}^\theta, \tag{3.131}$$

let $(\mathbb{F}^\theta_t)_{t\in[0,T]}$*,* $\theta\in\Theta$*, be filtrations on* $(\Omega,\mathcal{F},\mathbb{P})$ *which satisfy the usual conditions, assume that* $(\mathbb{F}^\theta_T)_{\theta\in\Theta}$ *is an independent family of sigma-algebras, assume that* $(\mathbb{F}^\theta_T)_{\theta\in\Theta}$ *and* $\big(\mathcal{R}^\theta\big)_{\theta\in\Theta}$ *are independent, for every* $\theta\in\Theta$ *let* $W^\theta\colon[0,T]\times\Omega\to\mathbb{R}^m$ *be a standard* $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}^\theta_t)_{t\in[0,T]})$*-Brownian motion, for every* $\theta\in\Theta$ *let* $X^\theta = (X^\theta_{t,s}(x))_{s\in[t,T],t\in[0,T],x\in\mathbb{R}^d}\colon \{(t,s)\in[0,T]^2\colon t\le s\}\times\mathbb{R}^d\times\Omega\to\mathbb{R}^d$ *be a continuous random field which satisfies for every* $t\in[0,T]$*,* $x\in\mathbb{R}^d$ *that* $(X^\theta_{t,s}(x))_{s\in[t,T]}\colon[t,T]\times\Omega\to\mathbb{R}^d$ *is an* $(\mathbb{F}^\theta_s)_{s\in[t,T]}$*/*$\mathcal{B}(\mathbb{R}^d)$*-adapted stochastic process and which satisfies that for all* $t\in[0,T]$*,* $s\in[t,T]$*,* $x\in\mathbb{R}^d$ *it holds* $\mathbb{P}$*-a.s. that*

$$X^\theta_{t,s}(x) = x + \int_t^s \mu\big(r, X^\theta_{t,r}(x)\big)\,dr + \int_t^s \sigma\big(r, X^\theta_{t,r}(x)\big)\,dW^\theta_r, \tag{3.132}$$

let $V^\theta_{M,n}\colon[0,T]\times\mathbb{R}^d\times\Omega\to\mathbb{R}$*,* $M,n\in\mathbb{Z}$*,* $\theta\in\Theta$*, be functions which satisfy for all* $M,n\in\mathbb{N}$*,* $\theta\in\Theta$*,* $t\in[0,T]$*,* $x\in\mathbb{R}^d$ *that* $V^\theta_{M,0}(t,x)=0$ *and*

$$\begin{aligned}
&V^\theta_{M,n}(t,x)\\
&= \frac{1}{M^n}\left[\sum_{m=1}^{M^n} g\big(X^{(\theta,n,-m)}_{t,T}(x)\big)\right]\\
&\quad+\sum_{k=0}^{n-1}\frac{(T-t)}{M^{n-k}}\left[\sum_{m=1}^{M^{n-k}} f\Big(R^{(\theta,k,m)}_t, X^{(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x), V^{(\theta,k,m)}_{M,k}\big(R^{(\theta,k,m)}_t, X^{(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x)\big)\Big)\right.\\
&\quad\left.-\mathbb{1}_{\mathbb{N}}(k) f\Big(R^{(\theta,k,m)}_t, X^{(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x), V^{(\theta,k,-m)}_{M,k-1}\big(R^{(\theta,k,m)}_t, X^{(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x)\big)\Big)\right],
\end{aligned} \tag{3.133}$$

and let $(\mathcal{C}_{M,n})_{M,n\in\mathbb{Z}}\subseteq\mathbb{N}_0$ *satisfy for all* $n,M\in\mathbb{N}$ *that* $\mathcal{C}_{M,0}=0$ *and*

$$\mathcal{C}_{M,n}\le\alpha d M^n + \sum_{k=0}^{n-1}\Big[M^{(n-k)}(\alpha d + 1 + \mathcal{C}_{M,k} + \mathbb{1}_{\mathbb{N}}(k)\mathcal{C}_{M,k-1})\Big]. \tag{3.134}$$

Then

(i) *there exists a unique at most polynomially growing function* $u\in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ *which satisfies that* $u|_{(0,T)\times\mathbb{R}^d}\colon(0,T)\times\mathbb{R}^d\to\mathbb{R}$ *is a viscosity solution of*

$$\begin{aligned}(\tfrac{\partial u}{\partial t})(t,x) + \tfrac12\operatorname{Trace}\big(\sigma(t,x)[\sigma(t,x)]^*(\operatorname{Hess}_x u)(t,x)\big)&\\ + \langle\mu(t,x),(\nabla_x u)(t,x)\rangle + f(t,x,u(t,x)) &= 0\end{aligned} \tag{3.135}$$

for $(t,x)\in(0,T)\times\mathbb{R}^d$ *(cf. Definition 3.21) and which satisfies for all* $x\in\mathbb{R}^d$ *that* $u(T,x)=g(x)$*,*

(ii) *it holds for all* $M\in\mathbb{N}$*,* $n\in\mathbb{N}_0$ *that*

$$\big(\mathbb{E}\big[|u(0,\xi)-V^0_{M,n}(0,\xi)|^2\big]\big)^{1/2}\le\frac{\mathfrak{C}(1+2LT)^n\exp(\frac{M}{2})}{M^{n/2}}<\infty, \tag{3.136}$$

and

(iii) there exists a function $N\colon (0,\infty)\to\mathbb{N}$ such that for all $\varepsilon,\delta\in(0,\infty)$ it holds that

$$\big(\mathbb{E}\big[|u(0,\xi)-V^0_{N_\varepsilon,N_\varepsilon}(0,\xi)|^2\big]\big)^{1/2}\le\varepsilon \qquad \textit{and} \tag{3.137}$$

$$\mathcal{C}_{N_\varepsilon,N_\varepsilon}\le\alpha\, d\,\max\{1,\mathfrak{C}^{2+\delta}\}\left[\sup_{n\in\mathbb{N}}\tfrac{(4+8LT)^{(3+\delta)(n+1)}}{n^{(n\delta/2)}}\right](\min\{1,\varepsilon\})^{-(2+\delta)}<\infty. \tag{3.138}$$

Proof of Proposition 3.23. Throughout this proof let $(\rho_1^{(q)})_{q\in[0,\infty)},(\rho_2^{(q)})_{q\in[0,\infty)}\subseteq[0,\infty)$, $C\in[0,\infty]$ satisfy for all $q\in[0,\infty)$ that

$$\rho_1^{(q)}=\tfrac{q(q+3)(C_2+1)}{2},\qquad \rho_2^{(q)}=(q+1)|C_1|^{q/2},\qquad\text{and} \tag{3.139}$$

$$C=\left[\left(\mathbb{E}\big[|g(X^0_{0,T}(\xi))|^2\big]\right)^{1/2}+\sqrt{T}\left(\int_0^T\mathbb{E}\big[|f(t,X^0_{0,t}(\xi),0)|^2\big]\,dt\right)^{1/2}\right]e^{LT}. \tag{3.140}$$

Observe that the fact that μ and σ are globally Lipschitz continuous functions, (3.129), and item (ii) in Proposition 3.22 assure that there exists a unique at most polynomially growing function $u\in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ which satisfies that $u|_{(0,T)\times\mathbb{R}^d}\colon(0,T)\times\mathbb{R}^d\to\mathbb{R}$ is a viscosity solution of

$$\begin{aligned}(\tfrac{\partial u}{\partial t})(t,x)+\tfrac12\operatorname{Trace}\big(\sigma(t,x)[\sigma(t,x)]^*(\operatorname{Hess}_x u)(t,x)\big)&\\ +\langle\mu(t,x),(\nabla_x u)(t,x)\rangle+f(t,x,u(t,x))&=0\end{aligned} \tag{3.141}$$

for $(t,x)\in(0,T)\times\mathbb{R}^d$ and which satisfies for all $x\in\mathbb{R}^d$ that $u(T,x)=g(x)$. This proves item (i). In addition, note that the fact that μ and σ are globally Lipschitz continuous functions, (3.129), (3.132), and Proposition 3.22 assure that for all $t\in[0,T]$, $x\in\mathbb{R}^d$ it holds that $\mathbb{E}\big[|g(X^0_{t,T}(x))|+\int_t^T|f(r,X^0_{t,r}(x),u(r,X^0_{t,r}(x)))|\,dr\big]<\infty$ and

$$u(t,x)=\mathbb{E}\left[g\big(X^0_{t,T}(x)\big)+\int_t^T f\big(r,X^0_{t,r}(x),u(r,X^0_{t,r}(x))\big)\,dr\right]. \tag{3.142}$$

Moreover, observe that the hypothesis that μ and σ are globally Lipschitz continuous functions, the fact that for all $\theta,\vartheta\in\Theta$ with $\theta\ne\vartheta$ it holds that $\mathbb{F}^\theta_T$ and $\mathbb{F}^\vartheta_T$ are independent, (3.132), and Lemma 2.20 assure that for all $\theta,\vartheta\in\Theta$, $r,s,t\in[0,T]$, $x\in\mathbb{R}^d$, $B\in\mathcal{B}(\mathbb{R}^d)$ with $t\le s\le r$ and $\theta\ne\vartheta$ it holds that $\mathbb{P}(X^\theta_{t,t}(x)=x)=1$ and

$$\mathbb{P}\big(X^\theta_{s,r}(X^\vartheta_{t,s}(x))\in B\big)=\mathbb{P}\big(X^\theta_{t,r}(x)\in B\big). \tag{3.143}$$

Next note that the hypothesis that μ and σ are globally Lipschitz continuous functions, (3.130), (3.132), and Lemma 2.7 (with $d=d$, $m=m$, $T=T-t$, $C_1=C_1$, $C_2=C_2$, $\xi=x$, $(\mu(r,y))_{r\in[0,T],y\in\mathbb{R}^d}=(\mu(t+r,y))_{r\in[0,T-t],y\in\mathbb{R}^d}$, $(\sigma(r,y))_{r\in[0,T],y\in\mathbb{R}^d}=(\sigma(t+r,y))_{r\in[0,T-t],y\in\mathbb{R}^d}$, $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_r)_{r\in[0,T]})=(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}^0_{t+r})_{r\in[0,T-t]})$, $(W_r)_{r\in[0,T]}=(W^0_{t+r}-W^0_t)_{r\in[0,T-t]}$, $(X_r)_{r\in[0,T]}=(X^0_{t,t+r}(x))_{r\in[0,T-t]}$ for $t\in[0,T]$, $x\in\mathbb{R}^d$ in the notation of Lemma 2.7) assure that for all $x\in\mathbb{R}^d$, $t\in[0,T]$, $s\in[t,T]$, $q\in[0,\infty)$ it holds that

$$\mathbb{E}\Big[\big\|X^0_{t,s}(x)\big\|^q\Big]\le\max\{T,1\}\left((1+\|x\|^2)^{q/2}+\rho_2^{(q)}\right)e^{\rho_1^{(q)}T}. \tag{3.144}$$

For the next step let $\mathfrak{K}, \mathfrak{p} \in [0, \infty)$ satisfy for all $t \in [0, T]$, $x \in \mathbb{R}^d$ that

$$|u(t,x)| \le \mathfrak{K}(1 + \|x\|^{\mathfrak{p}}). \tag{3.145}$$

This, Tonelli's theorem, and (3.129) assure that for all $t \in [0, T]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
&\mathbb{E}\left[\big|g(X^0_{t,T}(x))\big| + \int_t^T \big|f(r, X^0_{t,r}(x), u(r, X^0_{t,r}(x)))\big|\, dr\right] \\
&\le \mathbb{E}\big[K\big(1 + \|X^0_{t,T}(x)\|^p\big)\big] + \int_t^T \mathbb{E}\big[K\big(1 + \|X^0_{t,r}(x)\|^p\big) + L\,\big|u(r, X^0_{t,r}(x))\big|\big]\; dr \\
&\le K\big(1 + \mathbb{E}\big[\|X^0_{t,T}(x)\|^p\big]\big) + \int_t^T K\big(1 + \mathbb{E}\big[\|X^0_{t,r}(x)\|^p\big]\big) + L\mathfrak{K}\left(1 + \mathbb{E}\big[\|X^0_{t,r}(x)\|^{\mathfrak{p}}\big]\right)\; dr \\
&\le (1 + (T-t))K\Big(1 + \max\{T,1\}\left((1 + \|x\|^2)^{p/2} + \rho_2^{(p)}\right) e^{\rho_1^{(p)}T}\Big) \\
&\quad + (T-t)L\mathfrak{K}\left(1 + \max\{T,1\}\left((1 + \|x\|^2)^{\mathfrak{p}/2} + \rho_2^{(\mathfrak{p})}\right) e^{\rho_1^{(\mathfrak{p})}T}\right) < \infty.
\end{aligned} \tag{3.146}$$

Moreover, observe that (3.129), (3.144), (3.145), and the triangle inequality demonstrate that for all $t \in [0, T]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
&\int_0^T \left(\mathbb{E}\big[|u(r, X^0_{0,r}(\xi))|^2\big]\right)^{1/2} dr + \int_t^T \mathbb{E}\big[|f(r, X^0_{t,r}(x), 0)|\big]\, dr \\
&\le \int_0^T \left(\mathbb{E}\Big[\big|\mathfrak{K}(1 + \|X^0_{0,r}(\xi)\|^{\mathfrak{p}})\big|^2\Big]\right)^{1/2} dr + \int_t^T \mathbb{E}\big[K(1 + \|X^0_{t,r}(x)\|^p)\big]\, dr \\
&\le \int_0^T \mathfrak{K}\big(1 + \left(\mathbb{E}\big[\|X^0_{0,r}(\xi)\|^{2\mathfrak{p}}\big]\right)^{1/2}\big)\, dr + \int_t^T K\big(1 + \mathbb{E}\big[\|X^0_{t,r}(x)\|^p\big]\big)\, dr \\
&\le T\mathfrak{K}\left(1 + \left[\max\{T,1\}\left((1 + \|\xi\|^2)^{\mathfrak{p}} + \rho_2^{(2\mathfrak{p})}\right) e^{\rho_1^{(2\mathfrak{p})}T}\right]^{1/2}\right) \\
&\quad + (T-t)K\left(1 + \max\{T,1\}\left((1 + \|x\|^2)^{p/2} + \rho_2^{(p)}\right) e^{\rho_1^{(p)}T}\right) < \infty.
\end{aligned} \tag{3.147}$$

Combining this, (3.129), (3.133), (3.142), (3.143), (3.146), the fact that $(X^\theta)_{\theta\in\Theta}$ are independent, and the fact that $(X^\theta)_{\theta\in\Theta}$ and $(\mathcal{R}^\theta)_{\theta\in\Theta}$ are independent with Proposition 3.15 (with $d = d$, $T = T$, $L = L$, $u = u$, $g = g$, $f = f$, $\mathcal{R}^\theta = \mathcal{R}^\theta$, $X^\theta = X^\theta$, $V^\theta_{M,n} = V^\theta_{M,n}$, $\xi = \xi$, $C = C$ for $M, n \in \mathbb{Z}$, $\theta \in \Theta$ in the notation of Proposition 3.15) proves that for all $M \in \mathbb{N}$, $n \in \mathbb{N}_0$ it holds that

$$\left(\mathbb{E}\big[|u(0,\xi) - V^0_{M,n}(0,\xi)|^2\big]\right)^{1/2} \le \frac{C(1 + 2LT)^n \exp(\frac{M}{2})}{M^{n/2}}. \tag{3.148}$$

Next observe that (3.129), (3.144), and the triangle inequality imply that

$$\begin{aligned}
\left(\mathbb{E}\big[|g(X^0_{0,T}(\xi))|^2\big]\right)^{1/2} &\le \left(\mathbb{E}\Big[\big|K(1 + \|X^0_{0,T}(\xi)\|^p)\big|^2\Big]\right)^{1/2} \\
&\le K\left(1 + \left(\mathbb{E}\big[\|X^0_{0,T}(\xi)\|^{2p}\big]\right)^{1/2}\right) \\
&\le K\left(1 + \left[\max\{T,1\}\left((1 + \|\xi\|^2)^p + \rho_2^{(2p)}\right) e^{\rho_1^{(2p)}T}\right]^{1/2}\right).
\end{aligned} \tag{3.149}$$

In addition note that (3.129), (3.144), and the triangle inequality imply that

$$\begin{aligned} &\left(\int_0^T \mathbb{E}\big[|f(t,X_{0,t}^0(\xi),0)|^2\big]\,dt\right)^{1/2} \\ &\leq \left(\int_0^T \mathbb{E}\Big[\big|K(1+\|X_{0,t}^0(\xi)\|^p)\big|^2\Big]\,dt\right)^{1/2} \\ &\leq K\left(\sqrt{T}+\left(\int_0^T \mathbb{E}\big[\|X_{0,t}^0(\xi)\|^{2p}\big]\,dt\right)^{1/2}\right) \\ &\leq K\sqrt{T}\left(1+\Big[\max\{T,1\}\Big((1+\|\xi\|^2)^p+\rho_2^{(2p)}\Big)e^{\rho_1^{(2p)}T}\Big]^{1/2}\right). \end{aligned} \tag{3.150}$$

Combining this and (3.149) with (3.139) and (3.140) demonstrates that

$$\begin{aligned} C &\leq 2K\max\{T,1\}\left(1+\Big[\max\{T,1\}\Big((1+\|\xi\|^2)^p+\rho_2^{(2p)}\Big)e^{\rho_1^{(2p)}T}\Big]^{1/2}\right)e^{LT} \\ &\leq 2K(\max\{T,1\})^{3/2}\left(1+\left[\Big((1+\|\xi\|^2)^p+(2p+1)|C_1|^p\Big)e^{\frac{2p(2p+3)(C_2+1)}{2}T}\right]^{1/2}\right)e^{LT} \\ &\leq 2Ke^{3T/2}\left(1+\Big((1+\|\xi\|^2)^{p/2}+\sqrt{(2p+1)}|C_1|^{p/2}\Big)e^{p(p+3/2)(C_2+1)T}\right)e^{LT} \\ &\leq 4Ke^{T(L+2+p(p+2)(C_2+1))}\Big((1+\|\xi\|^2)^{p/2}+(2p+1)|C_1|^{p/2}\Big)\leq \mathfrak{C}<\infty. \end{aligned} \tag{3.151}$$

This and (3.148) establish item (ii). In addition, observe that (3.129), (3.133), (3.134), (3.142), (3.143), (3.146) (3.147), the fact that $(X^\theta)_{\theta\in\Theta}$ are independent, the fact that $(X^\theta)_{\theta\in\Theta}$ and $(\mathcal{R}^\theta)_{\theta\in\Theta}$ are independent, (3.151), and Proposition 3.18 (with $d=d$, $T=T$, $L=L$, $u=u$, $g=g$, $f=f$, $X^\theta=X^\theta$, $V_{M,n}^\theta=V_{M,n}^\theta$, $\xi=\xi$, $C=C$, $\alpha=\alpha$, $\mathcal{C}_{M,n}=\mathcal{C}_{M,n}$ for $M,n\in\mathbb{Z}$, $\theta\in\Theta$ in the notation of Proposition 3.15) prove that there exists a function $N\colon(0,\infty)\to\mathbb{N}$ such that for all $\varepsilon,\delta\in(0,\infty)$ it holds that

$$\big(\mathbb{E}\big[|u(0,\xi)-V_{N_\varepsilon,N_\varepsilon}^0(0,\xi)|^2\big]\big)^{1/2}\leq\varepsilon \qquad \text{and} \tag{3.152}$$

$$\begin{aligned} \mathcal{C}_{N_\varepsilon,N_\varepsilon} &\leq \alpha\,d\,\max\{1,C^{2+\delta}\}\left[\sup_{n\in\mathbb{N}}\tfrac{(4+8LT)^{(3+\delta)(n+1)}}{n^{(n\delta/2)}}\right](\min\{1,\varepsilon\})^{-(2+\delta)} \\ &\leq \alpha\,d\,\max\{1,\mathfrak{C}^{2+\delta}\}\left[\sup_{n\in\mathbb{N}}\tfrac{(4+8LT)^{(3+\delta)(n+1)}}{n^{(n\delta/2)}}\right](\min\{1,\varepsilon\})^{-(2+\delta)}<\infty. \end{aligned} \tag{3.153}$$

This establishes item (iii). The proof of Proposition 3.23 is thus completed. □

3.6.3 MLP approximations in variable space dimensions

Theorem 3.24. *Let $T\in(0,\infty)$, $\alpha,c,K\in[1,\infty)$, $L,p,P,\mathfrak{P},q,C_1,C_2\in[0,\infty)$, $\Theta=\cup_{n=1}^\infty\mathbb{Z}^n$, for every $d\in\mathbb{N}$ let $\|\cdot\|_{\mathbb{R}^d}:\mathbb{R}^d\to[0,\infty)$ be the Euclidean norm on $\mathbb{R}^d$, let $\langle\cdot,\cdot\rangle_{\mathbb{R}^d}:\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}$ be the Euclidean scalar product on $\mathbb{R}^d$, and let $|\!|\!|\cdot|\!|\!|_d\colon\mathbb{R}^{d\times d}\to[0,\infty)$ be the Frobenius norm on $\mathbb{R}^{d\times d}$, for every $d\in\mathbb{N}$ let $\xi_d\in\mathbb{R}^d$ satisfy that $\|\xi_d\|_{\mathbb{R}^d}\leq cd^q$, for every $d\in\mathbb{N}$ let $g_d\in C(\mathbb{R}^d,\mathbb{R})$, $f_d\in C([0,T]\times\mathbb{R}^d\times\mathbb{R},\mathbb{R})$ satisfy for all $t\in[0,T]$, $x\in\mathbb{R}^d$, $v,w\in\mathbb{R}$ that*

$$\max\{|g_d(x)|,|f_d(t,x,0)|\}\leq Kd^{\mathfrak{P}}(1+\|x\|_{\mathbb{R}^d}^p) \qquad \textit{and} \qquad |f_d(t,x,v)-f_d(t,x,w)|\leq L|v-w|, \tag{3.154}$$

for every $d\in\mathbb{N}$ let $\mu_d\colon[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$ and $\sigma_d\colon[0,T]\times\mathbb{R}^d\to\mathbb{R}^{d\times d}$ be globally Lipschitz continuous functions which satisfy for all $t\in[0,T]$, $x\in\mathbb{R}^d$ that

$$\max\{\langle x,\mu_d(t,x)\rangle_{\mathbb{R}^d},|\!|\!|\sigma_d(t,x)|\!|\!|_d^2\}\leq C_1d^P+C_2\,\|x\|_{\mathbb{R}^d}^2\,, \tag{3.155}$$

let $(\Omega, \mathcal{F}, \mathbb{P})$ *be a complete probability space, let* $\mathcal{R}^\theta\colon \Omega \to [0,1]$, $\theta \in \Theta$, *be independent* $\mathcal{U}_{[0,1]}$*-distributed random variables, let* $R^\theta = (R^\theta_t)_{t\in[0,T]}\colon [0,T]\times\Omega \to [0,T]$, $\theta\in\Theta$, *be the stochastic processes which satisfy for all* $t \in [0,T]$, $\theta \in \Theta$ *that*

$$R^\theta_t = t + (T-t)\mathcal{R}^\theta, \tag{3.156}$$

let $(\mathbb{F}^{d,\theta}_t)_{t\in[0,T]}$, $d \in \mathbb{N}$, $\theta\in\Theta$, *be filtrations on* $(\Omega, \mathcal{F}, \mathbb{P})$ *which satisfy the usual conditions, assume for every* $d \in \mathbb{N}$ *that* $(\mathbb{F}^{d,\theta}_T)_{\theta\in\Theta}$ *is an independent family of sigma-algebras, assume that* $(\mathbb{F}^{d,\theta}_T)_{d\in\mathbb{N},\theta\in\Theta}$ *and* $(\mathcal{R}^\theta)_{\theta\in\Theta}$ *are independent, for every* $d\in\mathbb{N}$, $\theta\in\Theta$ *let* $W^{d,\theta}\colon [0,T]\times\Omega\to\mathbb{R}^d$ *be a standard* $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}^{d,\theta}_t)_{t\in[0,T]})$*-Brownian motion, for every* $d\in\mathbb{N}$, $\theta\in\Theta$ *let* $X^{d,\theta} = (X^{d,\theta}_{t,s}(x))_{s\in[t,T],t\in[0,T],x\in\mathbb{R}^d}\colon \{(t,s)\in[0,T]^2\colon t\le s\}\times\mathbb{R}^d\times\Omega\to\mathbb{R}^d$ *be a continuous random field which satisfies for every* $t\in[0,T]$, $x\in\mathbb{R}^d$ *that* $(X^{d,\theta}_{t,s}(x))_{s\in[t,T]}\colon [t,T]\times\Omega\to\mathbb{R}^d$ *is an* $(\mathbb{F}^{d,\theta}_s)_{s\in[t,T]}$/$\mathcal{B}(\mathbb{R}^d)$*-adapted stochastic process and which satisfies that for all* $t\in[0,T]$, $s\in[t,T]$, $x\in\mathbb{R}^d$ *it holds* $\mathbb{P}$*-a.s. that*

$$X^{d,\theta}_{t,s}(x) = x + \int_t^s \mu_d\big(r, X^{d,\theta}_{t,r}(x)\big)\,dr + \int_t^s \sigma_d\big(r, X^{d,\theta}_{t,r}(x)\big)\,dW^{d,\theta}_r, \tag{3.157}$$

let $V^{d,\theta}_{M,n}\colon [0,T]\times\mathbb{R}^d\times\Omega\to\mathbb{R}$, $M,n\in\mathbb{Z}$, $\theta\in\Theta$, $d\in\mathbb{N}$, *be functions which satisfy for all* $d,M,n\in\mathbb{N}$, $\theta\in\Theta$, $t\in[0,T]$, $x\in\mathbb{R}^d$ *that* $V^{d,\theta}_{M,0}(t,x) = 0$ *and*

$$\begin{aligned}
&V^{d,\theta}_{M,n}(t,x)\\
&= \frac{1}{M^n}\left[\sum_{m=1}^{M^n} g_d\big(X^{d,(\theta,n,-m)}_{t,T}(x)\big)\right]\\
&+ \sum_{k=0}^{n-1}\frac{(T-t)}{M^{n-k}}\left[\sum_{m=1}^{M^{n-k}} f_d\Big(R^{(\theta,k,m)}_t, X^{d,(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x), V^{d,(\theta,k,m)}_{M,k}\big(R^{(\theta,k,m)}_t, X^{d,(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x)\big)\Big)\right.\\
&\left.- \mathbb{1}_{\mathbb{N}}(k) f_d\Big(R^{(\theta,k,m)}_t, X^{d,(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x), V^{d,(\theta,k,-m)}_{M,k-1}\big(R^{(\theta,k,m)}_t, X^{d,(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x)\big)\Big)\right],
\end{aligned} \tag{3.158}$$

and let $(\mathcal{C}_{d,M,n})_{M,n\in\mathbb{Z},d\in\mathbb{N}}\subseteq\mathbb{N}_0$ *satisfy for all* $d,n,M\in\mathbb{N}$ *that* $\mathcal{C}_{d,M,0}=0$ *and*

$$\mathcal{C}_{d,M,n} \le \alpha d M^n + \sum_{k=0}^{n-1}\Big[M^{(n-k)}(\alpha d + 1 + \mathcal{C}_{d,M,k} + \mathbb{1}_{\mathbb{N}}(k)\mathcal{C}_{d,M,k-1})\Big]. \tag{3.159}$$

Then

(i) *for every* $d\in\mathbb{N}$ *there exists a unique at most polynomially growing function* $u_d\in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ *which satisfies that* $u_d|_{(0,T)\times\mathbb{R}^d}\colon (0,T)\times\mathbb{R}^d\to\mathbb{R}$ *is a viscosity solution of*

$$\begin{aligned}
(\tfrac{\partial u_d}{\partial t})(t,x) + \tfrac{1}{2}\operatorname{Trace}\big(\sigma_d(t,x)[\sigma_d(t,x)]^*(\operatorname{Hess}_x u_d)(t,x)\big)&\\
+ \langle \mu_d(t,x), (\nabla_x u_d)(t,x)\rangle_{\mathbb{R}^d} + f_d(t,x,u_d(t,x)) &= 0
\end{aligned} \tag{3.160}$$

for $(t,x)\in(0,T)\times\mathbb{R}^d$ *(cf. Definition 3.21) and which satisfies for all* $x\in\mathbb{R}^d$ *that* $u_d(T,x) = g_d(x)$ *and*

(ii) *there exists a function* $N = (N_{d,\varepsilon})_{d\in\mathbb{N},\varepsilon\in(0,\infty)}\colon \mathbb{N}\times(0,\infty)\to\mathbb{N}$ *such that for all*

$d \in \mathbb{N}$, $\varepsilon, \delta \in (0,\infty)$ *it holds that*

$$\big(\mathbb{E}\big[|u_d(0,\xi_d) - V^{d,0}_{N_{d,\varepsilon},N_{d,\varepsilon}}(0,\xi_d)|^2\big]\big)^{1/2} \leq \varepsilon \qquad \textit{and} \tag{3.161}$$

$$\begin{aligned}\mathcal{C}_{d,N_{d,\varepsilon},N_{d,\varepsilon}} \leq \alpha\Big[4^{p+2} K e^{T(L+2+p(p+2)(C_2+1))}\left(c^p + |C_1|^{p/2}\right)\Big]^{(2+\delta)}\left[\sup_{n\in\mathbb{N}} \tfrac{(4+8LT)^{(3+\delta)(n+1)}}{n^{(n\delta/2)}}\right] \\ \cdot\, d^{1+(\mathfrak{P}+\max\{pq,(Pp)/2\})(2+\delta)}(\min\{1,\varepsilon\})^{-(2+\delta)} < \infty.\end{aligned} \tag{3.162}$$

Proof of Theorem 3.24. Throughout this proof let $(\beta_\delta)_{\delta\in(0,\infty)} \subseteq (0,\infty)$, $(\mathfrak{C}_d)_{d\in\mathbb{N}} \subseteq [0,\infty)$ satisfy for all $\delta \in (0,\infty)$, $d \in \mathbb{N}$ that $\beta_\delta = \left[\sup_{n\in\mathbb{N}} \frac{(4+8LT)^{(3+\delta)(n+1)}}{n^{(n\delta/2)}}\right]$ and

$$\mathfrak{C}_d = 4Kd^{\mathfrak{P}} e^{T(L+2+p(p+2)(C_2+1))}\left((1+\|\xi_d\|^2_{\mathbb{R}^d})^{p/2} + (2p+1)|C_1 d^P|^{p/2}\right). \tag{3.163}$$

Observe that Proposition 3.23 (with $d = d$, $m = d$, $T = T$, $L = L$, $K = Kd^{\mathfrak{P}}$, $p = p$, $C_1 = C_1 d^P$, $C_2 = C_2$, $\alpha = \alpha$, $\xi = \xi_d$, $g = g_d$, $f = f_d$, $\mu = \mu_d$, $\sigma = \sigma_d$, $\mathcal{R}^\theta = \mathcal{R}^\theta$, $\mathbb{F}^\theta = \mathbb{F}^{d,\theta}$, $W^\theta = W^{d,\theta}$, $X^\theta = X^{d,\theta}$, $V^\theta_{M,n} = V^{d,\theta}_{M,n}$, $\mathcal{C}_{M,n} = \mathcal{C}_{d,M,n}$ for $d \in \mathbb{N}$, $M, n \in \mathbb{Z}$, $\theta \in \Theta$ in the notation of Proposition 3.23) proves that for every $d \in \mathbb{N}$

(I) there exists a unique at most polynomially growing function $u_d \in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ which satisfies that $u_d|_{(0,T)\times\mathbb{R}^d} \colon (0,T)\times\mathbb{R}^d \to \mathbb{R}$ is a viscosity solution of

$$\begin{aligned}(\tfrac{\partial u_d}{\partial t})(t,x) + \tfrac{1}{2}\operatorname{Trace}\big(\sigma_d(t,x)[\sigma_d(t,x)]^*(\operatorname{Hess}_x u_d)(t,x)\big) \\ + \langle \mu_d(t,x), (\nabla_x u_d)(t,x)\rangle_{\mathbb{R}^d} + f_d(t,x,u_d(t,x)) = 0\end{aligned} \tag{3.164}$$

for $(t,x) \in (0,T)\times\mathbb{R}^d$ and which satisfies for all $x \in \mathbb{R}^d$ that $u_d(T,x) = g_d(x)$ and

(II) there exists a function $N_d = (N_{d,\varepsilon})_{\varepsilon\in(0,\infty)} \colon (0,\infty) \to \mathbb{N}$ such that for all $\varepsilon, \delta \in (0,\infty)$ it holds that

$$\Big(\mathbb{E}\Big[|u_d(0,\xi_d) - V^{d,0}_{N_{d,\varepsilon},N_{d,\varepsilon}}(0,\xi_d)|^2\Big]\Big)^{1/2} \leq \varepsilon \qquad \text{and} \tag{3.165}$$

$$\mathcal{C}_{d,N_{d,\varepsilon},N_{d,\varepsilon}} \leq \alpha\, d\, \max\{1, \mathfrak{C}_d^{2+\delta}\}\beta_\delta(\min\{1,\varepsilon\})^{-(2+\delta)} < \infty. \tag{3.166}$$

Observe that item (I) proves item (i). Moreover, note that the hypothesis that for all $d \in \mathbb{N}$ it holds that $\|\xi_d\|_{\mathbb{R}^d} \leq cd^q$ and the fact that $(2p+1) \leq 4^{p+1}$ imply that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned}\mathfrak{C}_d &\leq 4Kd^{\mathfrak{P}} e^{T(L+2+p(p+2)(C_2+1))}\left((1+|cd^q|^2)^{p/2} + (2p+1)|C_1 d^P|^{p/2}\right) \\ &\leq 4Kd^{\mathfrak{P}} e^{T(L+2+p(p+2)(C_2+1))}\left(d^{pq}(1+c^2)^{p/2} + d^{(Pp)/2}(2p+1)|C_1|^{p/2}\right) \\ &\leq d^{\mathfrak{P}+\max\{pq,(Pp)/2\}} 4Ke^{T(L+2+p(p+2)(C_2+1))}\left(2^{p/2}c^p + 4^{p+1}|C_1|^{p/2}\right) \\ &\leq d^{\mathfrak{P}+\max\{pq,(Pp)/2\}} 4^{p+2}Ke^{T(L+2+p(p+2)(C_2+1))}\left(c^p + |C_1|^{p/2}\right) \geq 1.\end{aligned} \tag{3.167}$$

This and (3.166) demonstrate that for all $d \in \mathbb{N}$, $\delta, \varepsilon \in (0,\infty)$ it holds that

$$\begin{aligned}\mathcal{C}_{d,N_{d,\varepsilon},N_{d,\varepsilon}} \leq \alpha\left[4^{p+2}Ke^{T(L+2+p(p+2)(C_2+1))}\left(c^p + |C_1|^{p/2}\right)\right]^{(2+\delta)} \\ \cdot\, d^{1+(\mathfrak{P}+\max\{pq,(Pp)/2\})(2+\delta)}\beta_\delta(\min\{1,\varepsilon\})^{-(2+\delta)} < \infty.\end{aligned} \tag{3.168}$$

Combining this and (3.165) establishes item (ii). The proof of Theorem 3.24 is thus completed. □

4 MLP approximations for PDE models

The MLP scheme for semilinear Kolmogorov PDEs (cf. (3.158) in Theorem 3.24 above) proposed in Subsection 3.6 can only be implemented for semilinear Kolmogorov PDEs for which an explicit solution of the corresponding SDE is known. In this section, we consider the MLP algorithm for two examples of such semilinear Kolmogorov PDEs, semilinear heat equations (see Subsection 4.1 below) and semilinear Black-Scholes equations (see Subsections 4.2–4.3 below). Apart from specifying the linear part of the PDE we also choose a particular nonlinearity (cf. (4.46) in Corollary 4.5 below) in Subsection 4.3 to obtain a PDE, which is used in the pricing of financial derivatives with default risk (cf., e.g., Han et al. [54, (10)] and Duffie et al. [37]).

4.1 MLP approximations for semilinear heat equations

Theorem 4.1. *Let $T \in (0,\infty)$, $\kappa, p, \mathfrak{P}, q \in [0,\infty)$, $\Theta = \cup_{n=1}^{\infty} \mathbb{Z}^n$, for every $d \in \mathbb{N}$ let $\|\cdot\|_{\mathbb{R}^d} : \mathbb{R}^d \to [0,\infty)$ be the Euclidean norm on $\mathbb{R}^d$, for every $d \in \mathbb{N}$ let $\xi_d \in \mathbb{R}^d$ satisfy that $\|\xi_d\|_{\mathbb{R}^d} \leq \kappa d^q$, for every $d \in \mathbb{N}$ let $g_d \in C(\mathbb{R}^d, \mathbb{R})$, $f_d \in C([0,T] \times \mathbb{R}^d \times \mathbb{R}, \mathbb{R})$ satisfy for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $v, w \in \mathbb{R}$ that*

$$\max\{|g_d(x)|, |f_d(t,x,0)|\} \leq \kappa d^{\mathfrak{P}}(1 + \|x\|_{\mathbb{R}^d}^p) \qquad \text{and} \qquad |f_d(t,x,v) - f_d(t,x,w)| \leq \kappa |v-w|, \tag{4.1}$$

let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\mathcal{R}^\theta \colon \Omega \to [0,1]$, $\theta \in \Theta$, be independent $\mathcal{U}_{[0,1]}$-distributed random variables, let $R^\theta = (R^\theta_t)_{t\in[0,T]} \colon [0,T] \times \Omega \to [0,T]$, $\theta \in \Theta$, be the stochastic processes which satisfy for all $t \in [0,T]$, $\theta \in \Theta$ that

$$R^\theta_t = t + (T-t)\mathcal{R}^\theta, \tag{4.2}$$

for every $d \in \mathbb{N}$ let $W^{d,\theta} \colon [0,T] \times \Omega \to \mathbb{R}^d$, $\theta \in \Theta$, be independent standard Brownian motions, assume that $\big(W^{d,\theta}\big)_{d\in\mathbb{N},\theta\in\Theta}$ and $\big(\mathcal{R}^\theta\big)_{\theta\in\Theta}$ are independent, for every $d \in \mathbb{N}$, $\theta \in \Theta$ let $X^{d,\theta} = (X^{d,\theta}_{t,s}(x))_{s\in[t,T],t\in[0,T],x\in\mathbb{R}^d} \colon \{(t,s) \in [0,T]^2 \colon t \leq s\} \times \mathbb{R}^d \times \Omega \to \mathbb{R}^d$ be the function which satisfies for all $t \in [0,T]$, $s \in [t,T]$, $x \in \mathbb{R}^d$ that

$$X^{d,\theta}_{t,s}(x) = x + W^{d,\theta}_s - W^{d,\theta}_t, \tag{4.3}$$

let $V^{d,\theta}_{M,n} \colon [0,T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}$, $M, n \in \mathbb{Z}$, $\theta \in \Theta$, $d \in \mathbb{N}$, be functions which satisfy for all $d, M, n \in \mathbb{N}$, $\theta \in \Theta$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that $V^{d,\theta}_{M,0}(t,x) = 0$ and

$$\begin{aligned} &V^{d,\theta}_{M,n}(t,x) \\ &= \frac{1}{M^n}\left[\sum_{m=1}^{M^n} g_d\big(X^{d,(\theta,n,-m)}_{t,T}(x)\big)\right] \\ &+ \sum_{k=0}^{n-1} \frac{(T-t)}{M^{n-k}}\left[\sum_{m=1}^{M^{n-k}} f_d\Big(R^{(\theta,k,m)}_t, X^{d,(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x), V^{d,(\theta,k,m)}_{M,k}\big(R^{(\theta,k,m)}_t, X^{d,(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x)\big)\Big)\right. \\ &\left. - \mathbb{1}_{\mathbb{N}}(k) f_d\Big(R^{(\theta,k,m)}_t, X^{d,(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x), V^{d,(\theta,k,-m)}_{M,k-1}\big(R^{(\theta,k,m)}_t, X^{d,(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x)\big)\Big)\right], \end{aligned} \tag{4.4}$$

and let $(\mathcal{C}_{d,M,n})_{M,n\in\mathbb{Z},d\in\mathbb{N}} \subseteq \mathbb{N}_0$ satisfy for all $d, n, M \in \mathbb{N}$ that $\mathcal{C}_{d,M,0} = 0$ and

$$\mathcal{C}_{d,M,n} \leq dM^n + \sum_{k=0}^{n-1}\Big[M^{(n-k)}(d + 1 + \mathcal{C}_{d,M,k} + \mathbb{1}_{\mathbb{N}}(k)\mathcal{C}_{d,M,k-1})\Big]. \tag{4.5}$$

Then

(i) *for every* $d \in \mathbb{N}$ *there exists a unique at most polynomially growing function* $u_d \in C([0,T] \times \mathbb{R}^d, \mathbb{R})$ *which satisfies that* $u_d|_{(0,T)\times\mathbb{R}^d} \colon (0,T) \times \mathbb{R}^d \to \mathbb{R}$ *is a viscosity solution of*

$$(\tfrac{\partial u_d}{\partial t})(t,x) + \tfrac{1}{2}(\Delta_x u_d)(t,x) + f_d(t,x,u_d(t,x)) = 0 \tag{4.6}$$

for $(t,x) \in (0,T) \times \mathbb{R}^d$ *(cf. Definition 3.21) and which satisfies for all* $x \in \mathbb{R}^d$ *that* $u_d(T,x) = g_d(x)$ *and*

(ii) *there exist functions* $N = (N_{d,\varepsilon})_{d\in\mathbb{N},\varepsilon\in(0,\infty)} \colon \mathbb{N} \times (0,\infty) \to \mathbb{N}$ *and* $c = (c_\delta)_{\delta\in(0,\infty)} \colon (0,\infty) \to (0,\infty)$ *such that for all* $d \in \mathbb{N}$*,* $\varepsilon, \delta \in (0,\infty)$ *it holds that*

$$\big(\mathbb{E}\big[|u_d(0,\xi_d) - V^{d,0}_{N_{d,\varepsilon},N_{d,\varepsilon}}(0,\xi_d)|^2\big]\big)^{1/2} \leq \varepsilon \qquad \textit{and} \tag{4.7}$$

$$\mathcal{C}_{d,N_{d,\varepsilon},N_{d,\varepsilon}} \leq c_\delta\, d^{1+(\mathfrak{P}+\max\{pq,p/2\})(2+\delta)}(\min\{1,\varepsilon\})^{-(2+\delta)}. \tag{4.8}$$

Proof of Theorem 4.1. Throughout this proof assume w.l.o.g. that $\kappa \geq 1$, assume w.l.o.g. that $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space, for every $d \in \mathbb{N}$ let $\langle\cdot,\cdot\rangle_{\mathbb{R}^d} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be the Euclidean scalar product on $\mathbb{R}^d$ and let $|\!|\!|\cdot|\!|\!|_d \colon \mathbb{R}^{d\times d} \to [0,\infty)$ be the Frobenius norm on $\mathbb{R}^{d\times d}$, let $\mu_d \in C([0,T] \times \mathbb{R}^d, \mathbb{R}^d)$, $d \in \mathbb{N}$, and $\sigma_d \in C([0,T] \times \mathbb{R}^d, \mathbb{R}^{d\times d})$, $d \in \mathbb{N}$, satisfy for all $d \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$\mu_d(t,x) = 0 \qquad \text{and} \qquad \sigma_d(t,x) = I_{\mathbb{R}^d}, \tag{4.9}$$

and for every $d \in \mathbb{N}$, $\theta \in \Theta$, $t \in [0,T]$ let $\mathbb{F}^{d,\theta}_t \subseteq \mathcal{F}$ be the sigma-algebra which satisfies that

$$\mathbb{F}^{d,\theta}_t = \begin{cases} \bigcap_{s\in(t,T]} \mathfrak{S}\big(\mathfrak{S}(W^{d,\theta}_r \colon r \in [0,t]) \cup \{A \in \mathcal{F} \colon \mathbb{P}(A) = 0\}\big) & : t < T \\ \mathfrak{S}\big(\mathfrak{S}(W^{d,\theta}_r \colon r \in [0,T]) \cup \{A \in \mathcal{F} \colon \mathbb{P}(A) = 0\}\big) & : t = T. \end{cases} \tag{4.10}$$

Note that (4.10) implies that for every $d \in \mathbb{N}$, $\theta \in \Theta$ it holds that $(\mathbb{F}^{d,\theta}_t)_{t\in[0,T]}$ is a filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ which satisfies the usual conditions. Moreover, observe that (4.10) and Lemma 2.18 demonstrate that for every $d \in \mathbb{N}$, $\theta \in \Theta$ it holds that $W^{d,\theta} \colon [0,T] \times \Omega \to \mathbb{R}^d$ is a standard $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}^{d,\theta}_t)_{t\in[0,T]})$-Brownian motion. Next note that (4.3) and (4.9) assure that for every $d \in \mathbb{N}$, $\theta \in \Theta$ it holds that $X^{d,\theta}$ is a continuous random field which satisfies for every $t \in [0,T]$, $x \in \mathbb{R}^d$ that $(X^{d,\theta}_{t,s}(x))_{s\in[t,T]} \colon [t,T] \times \Omega \to \mathbb{R}^d$ is an $(\mathbb{F}^{d,\theta}_s)_{s\in[t,T]}$/$\mathcal{B}(\mathbb{R}^d)$-adapted stochastic process and which satisfies that for all $t \in [0,T]$, $s \in [t,T]$, $x \in \mathbb{R}^d$ it holds $\mathbb{P}$-a.s. that

$$x + \int_t^s \mu_d\big(r, X^{d,\theta}_{t,r}(x)\big)\,dr + \int_t^s \sigma_d\big(r, X^{d,\theta}_{t,r}(x)\big)\,dW^{d,\theta}_r = x + W^{d,\theta}_s - W^{d,\theta}_t = X^{d,\theta}_{t,s}(x). \tag{4.11}$$

In addition, note that for all $d \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that

$$\max\{\langle x, \mu_d(t,x)\rangle_{\mathbb{R}^d}, |\!|\!|\sigma_d(t,x)|\!|\!|^2_d\} = \max\{0,d\} = d. \tag{4.12}$$

This, (4.1), (4.2), (4.4), (4.5), (4.9), (4.11), and Theorem 3.24 (with $T = T$, $\alpha = 1$, $c = \kappa$, $K = \kappa$, $L = \kappa$, $p = p$, $P = 1$, $\mathfrak{P} = \mathfrak{P}$, $q = q$, $C_1 = 1$, $C_2 = 0$, $\xi_d = \xi_d$, $g_d = g_d$, $f_d = f_d$, $\mu_d = \mu_d$, $\sigma_d = \sigma_d$, $\mathcal{R}^\theta = \mathcal{R}^\theta$, $\mathbb{F}^{d,\theta} = \mathbb{F}^{d,\theta}$, $W^{d,\theta} = W^{d,\theta}$, $X^{d,\theta} = X^{d,\theta}$, $V^{d,\theta}_{M,n} = V^{d,\theta}_{M,n}$, $\mathcal{C}_{d,M,n} = \mathcal{C}_{d,M,n}$ for $d \in \mathbb{N}$, $M, n \in \mathbb{Z}$, $\theta \in \Theta$ in the notation of Theorem 3.24) establish that

(I) for every $d \in \mathbb{N}$ there exists a unique at most polynomially growing function $u_d \in C([0,T] \times \mathbb{R}^d, \mathbb{R})$ which satisfies that $u_d|_{(0,T)\times\mathbb{R}^d} \colon (0,T) \times \mathbb{R}^d \to \mathbb{R}$ is a viscosity solution of

$$(\tfrac{\partial u_d}{\partial t})(t,x) + \tfrac{1}{2}\operatorname{Trace}\big(I_{\mathbb{R}^d}[I_{\mathbb{R}^d}]^*(\operatorname{Hess}_x u_d)(t,x)\big) \\ + \langle 0, (\nabla_x u_d)(t,x)\rangle_{\mathbb{R}^d} + f_d(t,x,u_d(t,x)) = 0 \quad (4.13)$$

for $(t,x) \in (0,T) \times \mathbb{R}^d$ and which satisfies for all $x \in \mathbb{R}^d$ that $u_d(T,x) = g_d(x)$ and

(II) there exists a function $N = (N_{d,\varepsilon})_{d\in\mathbb{N},\varepsilon\in(0,\infty)} \colon \mathbb{N} \times (0,\infty) \to \mathbb{N}$ such that for all $d \in \mathbb{N}$, $\varepsilon, \delta \in (0,\infty)$ it holds that

$$\big(\mathbb{E}\big[|u_d(0,\xi_d) - V^{d,0}_{N_{d,\varepsilon},N_{d,\varepsilon}}(0,\xi_d)|^2\big]\big)^{1/2} \leq \varepsilon \qquad \text{and} \quad (4.14)$$

$$\mathcal{C}_{d,N_{d,\varepsilon},N_{d,\varepsilon}} \leq \Big[4^{p+2}\kappa e^{T(\kappa+2+p(p+2))}\,(\kappa^p+1)\Big]^{(2+\delta)} \left[\sup_{n\in\mathbb{N}} \tfrac{(4+8\kappa T)^{(3+\delta)(n+1)}}{n^{(n\delta/2)}}\right] \\ \cdot d^{1+(\mathfrak{P}+\max\{pq,p/2\})(2+\delta)}(\min\{1,\varepsilon\})^{-(2+\delta)} < \infty. \quad (4.15)$$

Note that item (I) establishes item (i). Moreover, observe that item (II) establishes item (ii). The proof of Theorem 4.1 is thus completed. ☐

4.2 MLP approximations for semilinear Black-Scholes equations

Lemma 4.2. *Let $d \in \mathbb{N}$, $T \in (0,\infty)$, $(\alpha_i)_{i\in\{1,2,\ldots,d\}}$, $(\beta_i)_{i\in\{1,2,\ldots,d\}} \subseteq \mathbb{R}$, let $\langle\cdot,\cdot\rangle \colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be the Euclidean scalar product on $\mathbb{R}^d$, let $\Sigma = (\zeta_1,\ldots,\zeta_d) \in \mathbb{R}^{d\times d}$ satisfy for all $i \in \{1,2,\ldots,d\}$ that $\langle\zeta_i,\zeta_i\rangle = 1$, let $(\Omega,\mathcal{F},\mathbb{P})$ be a complete probability space, let $W \colon [0,T]\times\Omega \to \mathbb{R}^d$ be a d-dimensional standard Brownian motion, and let $X = (X^{(i)}_{t,s}(x))_{s\in[t,T],t\in[0,T],x\in\mathbb{R}^d,i\in\{1,2,\ldots,d\}} \colon \{(t,s)\in[0,T]^2 \colon t\leq s\}\times\mathbb{R}^d\times\Omega \to \mathbb{R}^d$ be the function which satisfies for all $i \in \{1,2,\ldots,d\}$, $t\in[0,T]$, $s\in[t,T]$, $x=(x_1,x_2,\ldots,x_d)\in\mathbb{R}^d$ that*

$$X^{(i)}_{t,s}(x) = x_i \exp\Big(\big(\alpha_i - \tfrac{|\beta_i|^2}{2}\big)(s-t) + \beta_i\langle\zeta_i, W_s - W_t\rangle\Big). \quad (4.16)$$

Then it holds that X is a continuous random field which satisfies that for all $t\in[0,T]$, $s\in[t,T]$, $x\in\mathbb{R}^d$ it holds $\mathbb{P}$-a.s. that

$$X_{t,s}(x) = x + \int_t^s \begin{pmatrix} \alpha_1 X^{(1)}_{t,r}(x) \\ \vdots \\ \alpha_d X^{(d)}_{t,r}(x)\end{pmatrix} dr + \int_t^s \operatorname{diag}\big(\beta_1 X^{(1)}_{t,r}(x),\ldots,\beta_d X^{(d)}_{t,r}(x)\big)\Sigma^*\,dW_r. \quad (4.17)$$

Proof of Lemma 4.2. Throughout this proof let $t\in[0,T]$, $s\in(0,T]$, $x=(x_1,x_2,\ldots,x_d)\in\mathbb{R}^d$, let $f_i \colon [0,T]\times\mathbb{R}^d \to \mathbb{R}$, $i\in\{1,2,\ldots,d\}$, be the functions which satisfy for all $i\in\{1,2,\ldots,d\}$, $r\in[0,T]$, $w\in\mathbb{R}^d$ that

$$f_i(r,w) = x_i \exp\big((\alpha_i - \tfrac{|\beta_i|^2}{2})r + \beta_i\,\langle\zeta_i,w\rangle\,\big), \quad (4.18)$$

let $B = (B^{(i)})_{i\in\{1,2,\ldots,d\}} \colon [0,s-t]\times\Omega \to \mathbb{R}^d$ satisfy for all $r\in[0,s-t]$ that $B_r = W_{t+r} - W_t$, and let $\zeta^{(j)}_i \in \mathbb{R}$, $i,j\in\{1,2,\ldots,d\}$, satisfy for all $i\in\{1,2,\ldots,d\}$ that $\zeta_i = (\zeta^{(j)}_i)_{j\in\{1,2,\ldots,d\}}$. Observe that Itô's formula (cf., e.g., Karatzas & Shreve [68, Theorem 3.3.6]) assures that for all $i \in \{1,2,\dots,d\}$ it holds $\mathbb{P}$-a.s. that

$$\begin{aligned} X_{t,s}^{(i)}(x) &= f_i(s-t, W_s - W_t) = f_i(s-t, B_{s-t}) \\ &= f_i(0,B_0) + \int_0^{s-t} \left(\tfrac{\partial f_i}{\partial r}\right)(r,B_r)\,dr + \sum_{j=1}^d \int_0^{s-t} \left(\tfrac{\partial f_i}{\partial w_j}\right)(r,B_r)\,dB_r^{(j)} \\ &\quad + \frac{1}{2}\sum_{j=1}^d \int_0^{s-t} \left(\tfrac{\partial^2 f_i}{\partial w_j^2}\right)(r,B_r)\,dr \\ &= f_i(0,B_0) + \int_0^{s-t} (\alpha_i - \tfrac{|\beta_i|^2}{2}) f_i(r,B_r)\,dr + \sum_{j=1}^d \int_0^{s-t} \beta_i \zeta_i^{(j)} f_i(r,B_r)\,dB_r^{(j)} \\ &\quad + \frac{1}{2}\sum_{j=1}^d \int_0^{s-t} |\beta_i|^2 \big|\zeta_i^{(j)}\big|^2 f_i(r,B_r)\,dr. \end{aligned} \tag{4.19}$$

The fact that for all $i \in \{1,2,\dots,d\}$ it holds that $\sum_{j=1}^d \big|\zeta_i^{(j)}\big|^2 = \langle \zeta_i, \zeta_i \rangle = 1$ and the fact that for all $i \in \{1,2,\dots,d\}$, $r \in [0,s-t]$ it holds that $f_i(r,B_r) = X_{t,t+r}^{(i)}(x)$ hence assure that for all $i \in \{1,2,\dots,d\}$ it holds $\mathbb{P}$-a.s. that

$$\begin{aligned} X_{t,s}^{(i)}(x) &= x_i + \int_0^{s-t} \left(\left(\alpha_i - \tfrac{|\beta_i|^2}{2}\right) + \tfrac{1}{2}|\beta_i|^2 \left[\sum_{j=1}^d \big|\zeta_i^{(j)}\big|^2 \right] \right) X_{t,t+r}^{(i)}(x)\,dr \\ &\quad + \int_0^{s-t} \beta_i X_{t,t+r}^{(i)}(x) (\zeta_i)^* dB_r \\ &= x_i + \int_0^{s-t} \alpha_i X_{t,t+r}^{(i)}(x)\,dr + \int_0^{s-t} \beta_i X_{t,t+r}^{(i)}(x)(\zeta_i)^* dB_r \\ &= x_i + \int_t^s \alpha_i X_{t,r}^{(i)}(x)\,dr + \int_t^s \beta_i X_{t,r}^{(i)}(x)(\zeta_i)^* dW_r. \end{aligned} \tag{4.20}$$

This implies (4.17). The proof of Lemma 4.2 is thus completed. □

Theorem 4.3. *Let* $T \in (0,\infty)$, $\kappa, p, \mathfrak{P}, q \in [0,\infty)$, $(\alpha_{d,i})_{i\in\{1,2,\dots,d\},d\in\mathbb{N}}$, $(\beta_{d,i})_{i\in\{1,2,\dots,d\},d\in\mathbb{N}} \subseteq \mathbb{R}$, $\Theta = \cup_{n=1}^\infty \mathbb{Z}^n$ *satisfy that* $\sup_{d\in\mathbb{N},i\in\{1,2,\dots,d\}} \max\{|\alpha_{d,i}|, |\beta_{d,i}|^2\} \le \kappa$, *for every* $d \in \mathbb{N}$ *let* $\langle\cdot,\cdot\rangle_{\mathbb{R}^d} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ *be the Euclidean scalar product on* $\mathbb{R}^d$ *and let* $\|\cdot\|_{\mathbb{R}^d} : \mathbb{R}^d \to [0,\infty)$ *be the Euclidean norm on* $\mathbb{R}^d$, *for every* $d \in \mathbb{N}$ *let* $\xi_d \in \mathbb{R}^d$, $\Sigma_d = (\zeta_{d,1},\dots,\zeta_{d,d}) \in \mathbb{R}^{d\times d}$ *satisfy for all* $i \in \{1,2,\dots,d\}$ *that* $\|\xi_d\|_{\mathbb{R}^d} \le \kappa d^q$ *and* $\|\zeta_{d,i}\|_{\mathbb{R}^d} = 1$, *for every* $d \in \mathbb{N}$ *let* $\mu_d\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ *and* $\sigma_d\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}^{d\times d}$ *be the functions which satisfy for all* $t \in [0,T]$, $x = (x_1,x_2,\dots,x_d) \in \mathbb{R}^d$ *that*

$$\mu_d(t,x) = (\alpha_{d,1}x_1,\dots,\alpha_{d,d}x_d) \qquad \textit{and} \qquad \sigma_d(t,x) = \operatorname{diag}(\beta_{d,1}x_1,\dots,\beta_{d,d}x_d)\Sigma_d^*, \tag{4.21}$$

for every $d \in \mathbb{N}$ *let* $g_d \in C(\mathbb{R}^d,\mathbb{R})$, $f_d \in C([0,T]\times\mathbb{R}^d\times\mathbb{R},\mathbb{R})$ *satisfy for all* $t \in [0,T]$, $x \in \mathbb{R}^d$, $v,w \in \mathbb{R}$ *that*

$$\max\{|g_d(x)|, |f_d(t,x,0)|\} \le \kappa d^{\mathfrak{P}}(1+\|x\|_{\mathbb{R}^d}^p) \qquad \textit{and} \qquad |f_d(t,x,v) - f_d(t,x,w)| \le \kappa|v-w|, \tag{4.22}$$

let $(\Omega,\mathcal{F},\mathbb{P})$ *be a probability space, let* $\mathcal{R}^\theta\colon \Omega \to [0,1]$, $\theta \in \Theta$, *be independent* $\mathcal{U}_{[0,1]}$*-distributed random variables, let* $R^\theta = (R_t^\theta)_{t\in[0,T]}\colon [0,T]\times\Omega \to [0,T]$, $\theta \in \Theta$, *be the stochastic processes which satisfy for all* $t \in [0,T]$, $\theta \in \Theta$ *that*

$$R_t^\theta = t + (T-t)\mathcal{R}^\theta, \tag{4.23}$$

for every $d \in \mathbb{N}$ *let* $W^{d,\theta}\colon [0,T] \times \Omega \to \mathbb{R}^d$, $\theta \in \Theta$, *be independent standard Brownian motions, assume that* $\big(W^{d,\theta}\big)_{d\in\mathbb{N},\theta\in\Theta}$ *and* $\big(\mathcal{R}^\theta\big)_{\theta\in\Theta}$ *are independent, for every* $d \in \mathbb{N}$, $\theta \in \Theta$ *let* $X^{d,\theta} = (X^{d,\theta,i}_{t,s}(x))_{s\in[t,T],t\in[0,T],x\in\mathbb{R}^d,i\in\{1,2,\ldots,d\}}\colon \{(t,s) \in [0,T]^2\colon t \leq s\} \times \mathbb{R}^d \times \Omega \to \mathbb{R}^d$ *be the function which satisfies for all* $t \in [0,T]$, $s \in [t,T]$, $x = (x_1,x_2,\ldots,x_d) \in \mathbb{R}^d$, $i \in \{1,2,\ldots,d\}$ *that*

$$X^{d,\theta,i}_{t,s}(x) = x_i \exp\Big(\big(\alpha_{d,i} - \tfrac{|\beta_{d,i}|^2}{2}\big)(s-t) + \beta_{d,i}\langle \zeta_{d,i}, W^{d,\theta}_s - W^{d,\theta}_t\rangle_{\mathbb{R}^d}\Big), \tag{4.24}$$

let $V^{d,\theta}_{M,n}\colon [0,T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}$, $M,n \in \mathbb{Z}$, $\theta \in \Theta$, $d \in \mathbb{N}$, *be functions which satisfy for all* $d,M,n \in \mathbb{N}$, $\theta \in \Theta$, $t \in [0,T]$, $x \in \mathbb{R}^d$ *that* $V^{d,\theta}_{M,0}(t,x) = 0$ *and*

$$\begin{aligned}
&V^{d,\theta}_{M,n}(t,x)\\
&= \frac{1}{M^n}\left[\sum_{m=1}^{M^n} g_d\big(X^{d,(\theta,n,-m)}_{t,T}(x)\big)\right]\\
&+ \sum_{k=0}^{n-1} \frac{(T-t)}{M^{n-k}}\left[\sum_{m=1}^{M^{n-k}} f_d\Big(R^{(\theta,k,m)}_t, X^{d,(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x), V^{d,(\theta,k,m)}_{M,k}\big(R^{(\theta,k,m)}_t, X^{d,(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x)\big)\Big)\right.\\
&\left. - \mathbb{1}_{\mathbb{N}}(k) f_d\Big(R^{(\theta,k,m)}_t, X^{d,(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x), V^{d,(\theta,k,-m)}_{M,k-1}\big(R^{(\theta,k,m)}_t, X^{d,(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x)\big)\Big)\right],
\end{aligned} \tag{4.25}$$

and let $(\mathcal{C}_{d,M,n})_{M,n\in\mathbb{Z},d\in\mathbb{N}} \subseteq \mathbb{N}_0$ *satisfy for all* $d,n,M \in \mathbb{N}$ *that* $\mathcal{C}_{d,M,0} = 0$ *and*

$$\mathcal{C}_{d,M,n} \leq dM^n + \sum_{k=0}^{n-1}\Big[M^{(n-k)}(d + 1 + \mathcal{C}_{d,M,k} + \mathbb{1}_{\mathbb{N}}(k)\mathcal{C}_{d,M,k-1})\Big]. \tag{4.26}$$

Then

(i) for every $d \in \mathbb{N}$ *there exists a unique at most polynomially growing function* $u_d \in C([0,T] \times \mathbb{R}^d, \mathbb{R})$ *which satisfies that* $u_d|_{(0,T)\times\mathbb{R}^d}\colon (0,T) \times \mathbb{R}^d \to \mathbb{R}$ *is a viscosity solution of*

$$\begin{aligned}
(\tfrac{\partial u_d}{\partial t})(t,x) + \tfrac{1}{2}\operatorname{Trace}\big(\sigma_d(t,x)[\sigma_d(t,x)]^*(\operatorname{Hess}_x u_d)(t,x)\big)&\\
+ \langle \mu_d(t,x), (\nabla_x u_d)(t,x)\rangle_{\mathbb{R}^d} + f_d(t,x,u_d(t,x)) &= 0
\end{aligned} \tag{4.27}$$

for $(t,x) \in (0,T) \times \mathbb{R}^d$ *(cf. Definition 3.21) and which satisfies for all* $x \in \mathbb{R}^d$ *that* $u_d(T,x) = g_d(x)$ *and*

(ii) there exist functions $N = (N_{d,\varepsilon})_{d\in\mathbb{N},\varepsilon\in(0,\infty)}\colon \mathbb{N} \times (0,\infty) \to \mathbb{N}$ *and* $c = (c_\delta)_{\delta\in(0,\infty)}\colon (0,\infty) \to (0,\infty)$ *such that for all* $d \in \mathbb{N}$, $\varepsilon,\delta \in (0,\infty)$ *it holds that*

$$\big(\mathbb{E}\big[|u_d(0,\xi_d) - V^{d,0}_{N_{d,\varepsilon},N_{d,\varepsilon}}(0,\xi_d)|^2\big]\big)^{1/2} \leq \varepsilon \qquad \text{and} \tag{4.28}$$

$$\mathcal{C}_{d,N_{d,\varepsilon},N_{d,\varepsilon}} \leq c_\delta\, d^{1+(\mathfrak{P}+pq)(2+\delta)}(\min\{1,\varepsilon\})^{-(2+\delta)}. \tag{4.29}$$

Proof of Theorem 4.3. Throughout this proof assume w.l.o.g. that $\kappa \geq 1$, assume w.l.o.g. that $(\Omega,\mathcal{F},\mathbb{P})$ is a complete probability space, for every $d \in \mathbb{N}$ let $|||\cdot|||_d\colon \mathbb{R}^{d\times d} \to [0,\infty)$ be the Frobenius norm on $\mathbb{R}^{d\times d}$, for every $d \in \mathbb{N}$, $i \in \{1,2,\ldots,d\}$ let $\zeta^{(j)}_{d,i} \in \mathbb{R}$, $j \in \{1,2,\ldots,d\}$, satisfy that $\zeta_{d,i} = (\zeta^{(j)}_{d,i})_{j\in\{1,2,\ldots,d\}}$, and for every $d \in \mathbb{N}$, $\theta \in \Theta$, $t \in [0,T]$ let $\mathbb{F}^{d,\theta}_t \subseteq \mathcal{F}$ be the sigma-algebra which satisfies that

$$\mathbb{F}^{d,\theta}_t = \begin{cases} \bigcap_{s\in(t,T]} \mathfrak{S}\big(\mathfrak{S}(W^{d,\theta}_r\colon r \in [0,t]) \cup \{A \in \mathcal{F}\colon \mathbb{P}(A) = 0\}\big) & \colon t < T \\ \mathfrak{S}\big(\mathfrak{S}(W^{d,\theta}_r\colon r \in [0,T]) \cup \{A \in \mathcal{F}\colon \mathbb{P}(A) = 0\}\big) & \colon t = T. \end{cases} \tag{4.30}$$

Note that (4.30) implies that for every $d \in \mathbb{N}$, $\theta \in \Theta$ it holds that $(\mathbb{F}_t^{d,\theta})_{t\in[0,T]}$ is a filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ which satisfies the usual conditions. Moreover, observe that (4.30) and Lemma 2.18 demonstrate that for every $d \in \mathbb{N}$, $\theta \in \Theta$ it holds that $W^{d,\theta} \colon [0,T] \times \Omega \to \mathbb{R}^d$ is a standard $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_t^{d,\theta})_{t\in[0,T]})$-Brownian motion. In addition, note that (4.21) and the fact that $\sup_{d\in\mathbb{N}, i\in\{1,2,\dots,d\}} |\alpha_{d,i}| \le \kappa$ imply that for all $d \in \mathbb{N}$, $t \in [0,T]$, $x = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d$ it holds that

$$\langle x, \mu_d(t,x)\rangle_{\mathbb{R}^d} = \sum_{i=1}^d x_i \alpha_{d,i} x_i \le \sum_{i=1}^d |x_i|^2 |\alpha_{d,i}| \le \kappa \, \|x\|_{\mathbb{R}^d}^2 . \tag{4.31}$$

Furthermore, observe that (4.21), the fact that $\sup_{d\in\mathbb{N}, i\in\{1,2,\dots,d\}} |\beta_{d,i}|^2 \le \kappa$, and the hypothesis that for all $d \in \mathbb{N}$, $i \in \{1,2,\dots,d\}$ it holds that $\|\zeta_{d,i}\|_{\mathbb{R}^d} = 1$ assure that for all $d \in \mathbb{N}$, $t \in [0,T]$, $x = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d$ it holds that

$$\begin{aligned} |||\sigma_d(t,x)|||_d^2 &= \sum_{i,j=1}^d \big|\beta_{d,i} x_i \zeta_{d,i}^{(j)}\big|^2 = \sum_{i=1}^d \left[|\beta_{d,i}|^2 |x_i|^2 \sum_{j=1}^d \big|\zeta_{d,i}^{(j)}\big|^2 \right] \\ &\le \sum_{i=1}^d \kappa \, |x_i|^2 \, \|\zeta_{d,i}\|_{\mathbb{R}^d}^2 = \kappa \, \|x\|_{\mathbb{R}^d}^2 . \end{aligned} \tag{4.32}$$

This and (4.31) ensure that for all $d \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that

$$\max\{\langle x, \mu_d(t,x)\rangle_{\mathbb{R}^d} , |||\sigma_d(t,x)|||_d^2\} \le \kappa \, \|x\|_{\mathbb{R}^d}^2 . \tag{4.33}$$

Next note that (4.21), (4.24), and Lemma 4.2 (with $d = d$, $T = T$, $(\alpha_i)_{i\in\{1,\dots,d\}} = (\alpha_{d,i})_{i\in\{1,\dots,d\}}$, $(\beta_i)_{i\in\{1,\dots,d\}} = (\beta_{d,i})_{i\in\{1,\dots,d\}}$, $\Sigma = \Sigma_d$, $(\Omega, \mathcal{F}, \mathbb{P}) = (\Omega, \mathcal{F}, \mathbb{P})$, $W = W^{d,\theta}$, $X = X^{d,\theta}$ for $\theta \in \Theta$, $d \in \mathbb{N}$ in the notation of Lemma 4.2) demonstrate that for all $d \in \mathbb{N}$, $\theta \in \Theta$ it holds that $X^{d,\theta}$ is a continuous random field which satisfies for every $t \in [0,T]$, $x \in \mathbb{R}^d$ that $(X_{t,s}^{d,\theta}(x))_{s\in[t,T]} \colon [t,T] \times \Omega \to \mathbb{R}^d$ is an $(\mathbb{F}_s^{d,\theta})_{s\in[t,T]}$/$\mathcal{B}(\mathbb{R}^d)$-adapted stochastic process and which satisfies that for all $t \in [0,T]$, $s \in [t,T]$, $x \in \mathbb{R}^d$ it holds $\mathbb{P}$-a.s. that

$$X_{t,s}^{d,\theta}(x) = x + \int_t^s \mu_d\big(r, X_{t,r}^{d,\theta}(x)\big)\, dr + \int_t^s \sigma_d\big(r, X_{t,r}^{d,\theta}(x)\big)\, dW_r^{d,\theta}. \tag{4.34}$$

Combining this, (4.22), the fact that for all $d \in \mathbb{N}$ it holds that μ_d and σ_d are globally Lipschitz continuous functions, and (4.33) with Theorem 3.24 (with $T = T$, $\alpha = 1$, $c = \kappa$, $K = \kappa$, $L = \kappa$, $p = p$, $P = 0$, $\mathfrak{P} = \mathfrak{P}$, $q = q$, $C_1 = 0$, $C_2 = \kappa$, $\xi_d = \xi_d$, $g_d = g_d$, $f_d = f_d$, $\mu_d = \mu_d$, $\sigma_d = \sigma_d$, $(\Omega, \mathcal{F}, \mathbb{P}) = (\Omega, \mathcal{F}, \mathbb{P})$, $\mathcal{R}^\theta = \mathcal{R}^\theta$, $\mathbb{F}^{d,\theta} = \mathbb{F}^{d,\theta}$, $W^{d,\theta} = W^{d,\theta}$, $X^{d,\theta} = X^{d,\theta}$, $V_{M,n}^{d,\theta} = V_{M,n}^{d,\theta}$, $\mathcal{C}_{d,M,n} = \mathcal{C}_{d,M,n}$ for $d \in \mathbb{N}$, $\theta \in \Theta$, $M, n \in \mathbb{Z}$, in the notation of Theorem 3.24) establishes that

(I) for every $d \in \mathbb{N}$ there exists a unique at most polynomially growing function $u_d \in C([0,T] \times \mathbb{R}^d, \mathbb{R})$ which satisfies that $u_d|_{(0,T)\times\mathbb{R}^d} \colon (0,T) \times \mathbb{R}^d \to \mathbb{R}$ is a viscosity solution of

$$\begin{aligned} &(\tfrac{\partial u_d}{\partial t})(t,x) + \tfrac{1}{2} \operatorname{Trace}\big(\sigma_d(t,x)[\sigma_d(t,x)]^* (\operatorname{Hess}_x u_d)(t,x)\big) \\ &\qquad + \langle \mu_d(t,x), (\nabla_x u_d)(t,x)\rangle_{\mathbb{R}^d} + f_d(t,x,u_d(t,x)) = 0 \end{aligned} \tag{4.35}$$

for $(t,x) \in (0,T) \times \mathbb{R}^d$ and which satisfies for all $x \in \mathbb{R}^d$ that $u_d(T,x) = g_d(x)$ and

(II) there exists a function $N = (N_{d,\varepsilon})_{d\in\mathbb{N},\varepsilon\in(0,\infty)} \colon \mathbb{N} \times (0,\infty) \to \mathbb{N}$ such that for all

$d \in \mathbb{N}$, $\varepsilon, \delta \in (0,\infty)$ it holds that

$$\left(\mathbb{E}\left[|u_d(0,\xi_d) - V^{d,0}_{N_{d,\varepsilon},N_{d,\varepsilon}}(0,\xi_d)|^2\right]\right)^{1/2} \leq \varepsilon \qquad \text{and} \tag{4.36}$$

$$\begin{aligned}\mathcal{C}_{d,N_{d,\varepsilon},N_{d,\varepsilon}} \leq \left[4^{p+2}\kappa e^{T(\kappa+2+p(p+2)(\kappa+1))}\kappa^p\right]^{(2+\delta)}\left[\sup_{n\in\mathbb{N}} \tfrac{(4+8\kappa T)^{(3+\delta)(n+1)}}{n^{(n\delta/2)}}\right] \\ \cdot\, d^{1+(\mathfrak{P}+pq)(2+\delta)}(\min\{1,\varepsilon\})^{-(2+\delta)} < \infty.\end{aligned} \tag{4.37}$$

Observe that item (I) proves item (i). Furthermore, note that item (II) establishes item (ii). The proof of Theorem 4.3 is thus completed. ☐

Theorem 4.4 in an immediate consequence of Theorem 4.3 above. The proof of Theorem 4.4 is therefore omitted.

Theorem 4.4. *Let $T \in (0,\infty)$, $\kappa, p, \mathfrak{P}, q \in [0,\infty)$, $(\alpha_{d,i})_{i\in\{1,2,\ldots,d\},d\in\mathbb{N}}$, $(\beta_{d,i})_{i\in\{1,2,\ldots,d\},d\in\mathbb{N}} \subseteq \mathbb{R}$, $\Theta = \cup_{n=1}^\infty \mathbb{Z}^n$ satisfy that $\sup_{d\in\mathbb{N},i\in\{1,2,\ldots,d\}} \max\{|\alpha_{d,i}|, |\beta_{d,i}|^2\} \leq \kappa$, for every $d \in \mathbb{N}$ let $\langle\cdot,\cdot\rangle_{\mathbb{R}^d} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be the Euclidean scalar product on $\mathbb{R}^d$ and let $\|\cdot\|_{\mathbb{R}^d} : \mathbb{R}^d \to [0,\infty)$ be the Euclidean norm on $\mathbb{R}^d$, for every $d \in \mathbb{N}$ let $\xi_d \in \mathbb{R}^d$, $\Sigma_d = (\zeta_{d,1},\ldots,\zeta_{d,d}) \in \mathbb{R}^{d\times d}$ satisfy for all $i \in \{1,2,\ldots,d\}$ that $\|\xi_d\|_{\mathbb{R}^d} \leq \kappa d^q$ and $\|\zeta_{d,i}\|_{\mathbb{R}^d} = 1$, let $f\colon \mathbb{R} \to \mathbb{R}$ be a Lipschitz continuous function, for every $d \in \mathbb{N}$ let $g_d \in C(\mathbb{R}^d,\mathbb{R})$ satisfy for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that*

$$|g_d(x)| \leq \kappa d^{\mathfrak{P}}(1 + \|x\|^p_{\mathbb{R}^d}), \tag{4.38}$$

let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $\mathcal{R}^\theta\colon \Omega \to [0,1]$, $\theta \in \Theta$, be independent $\mathcal{U}_{[0,1]}$-distributed random variables, let $R^\theta = (R^\theta_t)_{t\in[0,T]}\colon [0,T]\times\Omega \to [0,T]$, $\theta \in \Theta$, be the stochastic processes which satisfy for all $t \in [0,T]$, $\theta \in \Theta$ that

$$R^\theta_t = t + (T-t)\mathcal{R}^\theta, \tag{4.39}$$

for every $d \in \mathbb{N}$ let $W^{d,\theta}\colon [0,T]\times\Omega \to \mathbb{R}^d$, $\theta \in \Theta$, be independent standard Brownian motions, assume that $\left(W^{d,\theta}\right)_{d\in\mathbb{N},\theta\in\Theta}$ and $\left(\mathcal{R}^\theta\right)_{\theta\in\Theta}$ are independent, for every $d \in \mathbb{N}$, $\theta \in \Theta$ let $X^{d,\theta} = (X^{d,\theta,i}_{t,s}(x))_{s\in[t,T],t\in[0,T],x\in\mathbb{R}^d,i\in\{1,2,\ldots,d\}}\colon \{(t,s) \in [0,T]^2\colon t \leq s\}\times\mathbb{R}^d\times\Omega \to \mathbb{R}^d$ be the function which satisfies for all $t \in [0,T]$, $s \in [t,T]$, $x = (x_1,x_2,\ldots,x_d) \in \mathbb{R}^d$, $i \in \{1,2,\ldots,d\}$ that

$$X^{d,\theta,i}_{t,s}(x) = x_i \exp\left(\left(\alpha_{d,i} - \tfrac{|\beta_{d,i}|^2}{2}\right)(s-t) + \beta_{d,i}\langle\zeta_{d,i}, W^{d,\theta}_s - W^{d,\theta}_t\rangle_{\mathbb{R}^d}\right), \tag{4.40}$$

let $V^{d,\theta}_{M,n}\colon [0,T]\times\mathbb{R}^d\times\Omega \to \mathbb{R}$, $M,n \in \mathbb{Z}$, $\theta \in \Theta$, $d \in \mathbb{N}$, be functions which satisfy for all $d,M,n \in \mathbb{N}$, $\theta \in \Theta$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that $V^{d,\theta}_{M,0}(t,x) = 0$ and

$$\begin{aligned}V^{d,\theta}_{M,n}(t,x) = \sum_{k=0}^{n-1} \frac{(T-t)}{M^{n-k}}\Bigg[\sum_{m=1}^{M^{n-k}} f\Big(V^{d,(\theta,k,m)}_{M,k}\big(R^{(\theta,k,m)}_t, X^{d,(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x)\big)\Big) \\ - \mathbb{1}_{\mathbb{N}}(k) f\Big(V^{d,(\theta,k,-m)}_{M,k-1}\big(R^{(\theta,k,m)}_t, X^{d,(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x)\big)\Big)\Bigg] + \left[\sum_{m=1}^{M^n} \frac{g_d(X^{d,(\theta,n,-m)}_{t,T}(x))}{M^n}\right],\end{aligned} \tag{4.41}$$

and let $(\mathcal{C}_{d,M,n})_{M,n\in\mathbb{Z},d\in\mathbb{N}} \subseteq \mathbb{N}_0$ satisfy for all $d,n,M \in \mathbb{N}$ that $\mathcal{C}_{d,M,0} = 0$ and

$$\mathcal{C}_{d,M,n} \leq dM^n + \sum_{k=0}^{n-1}\left[M^{(n-k)}(d + 1 + \mathcal{C}_{d,M,k} + \mathbb{1}_{\mathbb{N}}(k)\mathcal{C}_{d,M,k-1})\right]. \tag{4.42}$$

Then

(i) *for every* $d \in \mathbb{N}$ *there exists a unique at most polynomially growing function* $u_d \in C([0,T] \times \mathbb{R}^d, \mathbb{R})$ *which satisfies that* $u_d|_{(0,T)\times\mathbb{R}^d} : (0,T) \times \mathbb{R}^d \to \mathbb{R}$ *is a viscosity solution of*

$$(\tfrac{\partial u_d}{\partial t})(t,x) + \left[\sum_{i,j=1}^{d} \tfrac{\beta_{d,i}\beta_{d,j}x_i x_j \langle \zeta_{d,i},\zeta_{d,j}\rangle_{\mathbb{R}^d}}{2} \big(\tfrac{\partial^2 u_d}{\partial x_i \partial x_j}\big)(t,x)\right] + \left[\sum_{i=1}^{d} \alpha_{d,i} x_i \big(\tfrac{\partial u_d}{\partial x_i}\big)(t,x)\right] + f(u_d(t,x)) = 0 \quad (4.43)$$

for $(t,x) \in (0,T) \times \mathbb{R}^d$ *(cf. Definition 3.21) and which satisfies for all* $x \in \mathbb{R}^d$ *that* $u_d(T,x) = g_d(x)$ *and*

(ii) *there exist functions* $N = (N_{d,\varepsilon})_{d\in\mathbb{N},\varepsilon\in(0,\infty)} : \mathbb{N} \times (0,\infty) \to \mathbb{N}$ *and* $c = (c_\delta)_{\delta\in(0,\infty)} : (0,\infty) \to (0,\infty)$ *such that for all* $d \in \mathbb{N}$, $\varepsilon, \delta \in (0,\infty)$ *it holds that*

$$\big(\mathbb{E}\big[|u_d(0,\xi_d) - V^{d,0}_{N_{d,\varepsilon},N_{d,\varepsilon}}(0,\xi_d)|^2\big]\big)^{1/2} \le \varepsilon \qquad \text{and} \quad (4.44)$$

$$\mathcal{C}_{d,N_{d,\varepsilon},N_{d,\varepsilon}} \le c_\delta\, d^{1+(\mathfrak{P}+pq)(2+\delta)} (\min\{1,\varepsilon\})^{-(2+\delta)}. \quad (4.45)$$

4.3 MLP approximations for the pricing of financial derivatives with default risks

The next result, Corollary 4.5 below, is an immediate consequence of Theorem 4.4 above. The proof of Corollary 4.5 is therefore omitted.

Corollary 4.5. *Let* $T, R, \gamma_l, \gamma_h, v_l, v_h \in (0,\infty)$, $p, \mathfrak{P}, q \in [0,\infty)$, $\epsilon \in [0,1)$, $\alpha, \beta \in \mathbb{R}$, $f \in C(\mathbb{R},\mathbb{R})$, $\Theta = \cup_{n=1}^{\infty} \mathbb{Z}^n$ *satisfy for all* $u \in \mathbb{R}$ *that* $\gamma_l < \gamma_h$, $v_l > v_h$, *and*

$$f(u) = -Ru - (1-\epsilon)\left[\min\left\{\gamma_h, \max\left\{\gamma_l, \tfrac{(\gamma_h-\gamma_l)}{(v_h-v_l)}(u-v_h) + \gamma_h\right\}\right\}\right] u, \quad (4.46)$$

for every $d \in \mathbb{N}$ *let* $\|\cdot\|_{\mathbb{R}^d} : \mathbb{R}^d \to [0,\infty)$ *be the Euclidean norm on* $\mathbb{R}^d$, *let* $\xi_d \in \mathbb{R}^d$, $d \in \mathbb{N}$, *satisfy that* $\sup_{d\in\mathbb{N}} \frac{\|\xi_d\|_{\mathbb{R}^d}}{d^q} < \infty$, *let* $g_d \in C(\mathbb{R}^d,\mathbb{R})$, $d \in \mathbb{N}$, *satisfy that* $\sup_{d\in\mathbb{N},x\in\mathbb{R}^d} \frac{|g_d(x)|}{d^{\mathfrak{P}}(1+\|x\|^p_{\mathbb{R}^d})} < \infty$, *let* $(\Omega,\mathcal{F},\mathbb{P})$ *be a probability space, let* $\mathcal{R}^\theta : \Omega \to [0,1]$, $\theta \in \Theta$, *be independent* $\mathcal{U}_{[0,1]}$*-distributed random variables, let* $R^\theta = (R^\theta_t)_{t\in[0,T]} : [0,T] \times \Omega \to [0,T]$, $\theta \in \Theta$, *be the stochastic processes which satisfy for all* $t \in [0,T]$, $\theta \in \Theta$ *that* $R^\theta_t = t + (T-t)\mathcal{R}^\theta$, *for every* $d \in \mathbb{N}$ *let* $W^{d,\theta} = (W^{d,\theta,i})_{i\in\{1,2,\dots,d\}} : [0,T] \times \Omega \to \mathbb{R}^d$, $\theta \in \Theta$, *be independent standard Brownian motions, assume that* $\big(W^{d,\theta}\big)_{d\in\mathbb{N},\theta\in\Theta}$ *and* $\big(\mathcal{R}^\theta\big)_{\theta\in\Theta}$ *are independent, for every* $d \in \mathbb{N}$, $\theta \in \Theta$ *let* $X^{d,\theta} = (X^{d,\theta,i}_{t,s}(x))_{s\in[t,T],t\in[0,T],x\in\mathbb{R}^d,i\in\{1,2,\dots,d\}} : \{(t,s) \in [0,T]^2 : t \le s\} \times \mathbb{R}^d \times \Omega \to \mathbb{R}^d$ *be the function which satisfies for all* $i \in \{1,2,\dots,d\}$, $t \in [0,T]$, $s \in [t,T]$, $x = (x_1,x_2,\dots,x_d) \in \mathbb{R}^d$ *that*

$$X^{d,\theta,i}_{t,s}(x) = x_i \exp\Big(\big(\alpha - \tfrac{\beta^2}{2}\big)(s-t) + \beta\big(W^{d,\theta,i}_s - W^{d,\theta,i}_t\big)\Big), \quad (4.47)$$

let $V^{d,\theta}_{M,n} : [0,T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}$, $M, n \in \mathbb{Z}$, $\theta \in \Theta$, $d \in \mathbb{N}$, *be functions which satisfy for all* $d, M, n \in \mathbb{N}$, $\theta \in \Theta$, $t \in [0,T]$, $x \in \mathbb{R}^d$ *that* $V^{d,\theta}_{M,0}(t,x) = 0$ *and*

$$V^{d,\theta}_{M,n}(t,x) = \sum_{k=0}^{n-1} \frac{(T-t)}{M^{n-k}} \left[\sum_{m=1}^{M^{n-k}} f\Big(V^{d,(\theta,k,m)}_{M,k}\big(R^{(\theta,k,m)}_t, X^{d,(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x)\big)\Big) - \mathbb{1}_{\mathbb{N}}(k) f\Big(V^{d,(\theta,k,-m)}_{M,k-1}\big(R^{(\theta,k,m)}_t, X^{d,(\theta,k,m)}_{t,R^{(\theta,k,m)}_t}(x)\big)\Big)\right] + \left[\sum_{m=1}^{M^n} \frac{g_d(X^{d,(\theta,n,-m)}_{t,T}(x))}{M^n}\right], \quad (4.48)$$

and let $(\mathcal{C}_{d,M,n})_{M,n\in\mathbb{Z},d\in\mathbb{N}} \subseteq \mathbb{N}_0$ *satisfy for all* $d,n,M \in \mathbb{N}$ *that* $\mathcal{C}_{d,M,0} = 0$ *and*

$$\mathcal{C}_{d,M,n} \leq dM^n + \sum_{k=0}^{n-1} \Big[M^{(n-k)}(d+1+\mathcal{C}_{d,M,k} + \mathbb{1}_{\mathbb{N}}(k)\mathcal{C}_{d,M,k-1}) \Big]. \tag{4.49}$$

Then

(i) for every $d \in \mathbb{N}$ *there exists a unique at most polynomially growing function* $u_d \in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ *which satisfies that* $u_d|_{(0,T)\times\mathbb{R}^d} \colon (0,T)\times\mathbb{R}^d \to \mathbb{R}$ *is a viscosity solution of*

$$\begin{aligned}(\tfrac{\partial u_d}{\partial t})(t,x) + \left[\sum_{i=1}^{d} \tfrac{|\beta|^2|x_i|^2}{2}\big(\tfrac{\partial^2 u_d}{\partial (x_i)^2}\big)(t,x) + \alpha x_i \big(\tfrac{\partial u_d}{\partial x_i}\big)(t,x)\right] - Ru_d(t,x)&\\ - (1-\epsilon)\left[\min\left\{\gamma_h, \max\left\{\gamma_l, \tfrac{(\gamma_h-\gamma_l)}{(v_h-v_l)}(u_d(t,x)-v_h)+\gamma_h\right\}\right\}\right] u_d(t,x) = 0&\end{aligned} \tag{4.50}$$

for $(t,x) = (t,(x_1,x_2,\ldots,x_d)) \in (0,T)\times\mathbb{R}^d$ *(cf. Definition 3.21) and which satisfies for all* $x \in \mathbb{R}^d$ *that* $u_d(T,x) = g_d(x)$ *and*

(ii) there exist functions $N = (N_{d,\varepsilon})_{d\in\mathbb{N},\varepsilon\in(0,1]} \colon \mathbb{N}\times(0,1] \to \mathbb{N}$ *and* $c = (c_\delta)_{\delta\in(0,\infty)} \colon (0,\infty) \to (0,\infty)$ *such that for all* $d \in \mathbb{N}$*,* $\varepsilon \in (0,1]$*,* $\delta \in (0,\infty)$ *it holds that* $\mathcal{C}_{d,N_{d,\varepsilon},N_{d,\varepsilon}} \leq c_\delta\, d^{1+(\mathfrak{P}+pq)(2+\delta)}\varepsilon^{-(2+\delta)}$ *and*

$$\Big(\mathbb{E}\big[|u_d(0,\xi_d) - V^{d,0}_{N_{d,\varepsilon},N_{d,\varepsilon}}(0,\xi_d)|^2\big]\Big)^{1/2} \leq \varepsilon. \tag{4.51}$$

References

[1] Agarwal, R. P. *Difference equations and inequalities: theory, methods, and applications*. CRC Press, 2000. MR-1740241

[2] Bally, V., Pages, G., et al. A quantization algorithm for solving multidimensional discrete-time optimal stopping problems. *Bernoulli 9*, 6 (2003), 1003–1049. MR-2046816

[3] Beck, C., Becker, S., Grohs, P., Jaafari, N., and Jentzen, A. Solving stochastic differential equations and Kolmogorov equations by means of deep learning. arXiv:1806.00421 (2018), 56 pages.

[4] Beck, C., Gonon, L., Hutzenthaler, M., and Jentzen, A. On existence and uniqueness properties for solutions of stochastic fixed point equations. arXiv:1908.03382 (2019), 33 pages.

[5] Beck, C., Hutzenthaler, M., and Jentzen, A. On nonlinear Feynman-Kac formulas for viscosity solutions of semilinear parabolic partial differential equations. arXiv:2004.03389 (2020), 54 pages.

[6] Beck, C., Weinan, E., and Jentzen, A. Machine learning approximation algorithms for high-dimensional fully nonlinear partial differential equations and second-order backward stochastic differential equations. *Journal of Nonlinear Science* (2017), 1–57. MR-3993178

[7] Becker, S., Cheridito, P., and Jentzen, A. Deep optimal stopping. *J. Mach. Learn. Res. 20* (2019), Paper No. 74, 25. MR-3960928

[8] Bellman, R. Dynamic programming. *Science 153*, 3731 (1966), 34–37. MR-0191704

[9] Bender, C., and Denk, R. A forward scheme for backward SDEs. *Stochastic processes and their applications 117*, 12 (2007), 1793–1812. MR-2437729

[10] Bender, C., Schweizer, N., and Zhuo, J. A primal-dual algorithm for BSDEs. *Mathematical Finance. An International Journal of Mathematics, Statistics and Financial Economics 27*, 3 (2017), 866–901. MR-3668160

[11] Berg, J., and Nyström, K. A unified deep artificial neural network approach to partial differential equations in complex geometries. *Neurocomputing 317* (2018), 28–41. MR-3796894

[12] Black, F., and Scholes, M. The pricing of options and corporate liabilities. *Journal of political economy 81*, 3 (1973), 637–654. MR-3363443

[13] Bouchard, B., Elie, R., and Touzi, N. Discrete-time approximation of BSDEs and probabilistic schemes for fully nonlinear PDEs. *Advanced financial modelling 8* (2009), 91–124. MR-2648459

[14] Bouchard, B., Elie, R., and Touzi, N. Discrete-time approximation of BSDEs and probabilistic schemes for fully nonlinear PDEs. In *Advanced financial modelling*, vol. 8 of *Radon Ser. Comput. Appl. Math.* Walter de Gruyter, Berlin, 2009, pp. 91–124. MR-2648459

[15] Bouchard, B., and Touzi, N. Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Processes and their Applications 111*, 2 (2004), 175–206. MR-2056536

[16] Briand, P., and Labart, C. Simulation of BSDEs by Wiener chaos expansion. *The Annals of Applied Probability 24*, 3 (2014), 1129–1171. MR-3199982

[17] Burgard, C., and Kjaer, M. Partial differential equation representations of derivatives with bilateral counterparty risk and funding costs. *The Journal of Credit Risk 7*, 3 (2011), 1–19.

[18] Chan-Wai-Nam, Q., Mikael, J., and Warin, X. Machine learning for semi linear PDEs. arXiv:1809.07609 (2018). MR-3946472

[19] Chang, D., Liu, H., and Xiong, J. A branching particle system approximation for a class of FBSDEs. *Probability, Uncertainty and Quantitative Risk 1* (2016), Paper No. 9, 34. MR-3583180

[20] Chassagneux, J.-F. Linear multistep schemes for BSDEs. *SIAM Journal on Numerical Analysis 52*, 6 (2014), 2815–2836. MR-3284573

[21] Chassagneux, J.-F., and Crisan, D. Runge-Kutta schemes for backward stochastic differential equations. *The Annals of Applied Probability 24*, 2 (2014), 679–720. MR-3178495

[22] Chassagneux, J.-F., and Richou, A. Numerical stability analysis of the Euler scheme for BSDEs. *SIAM Journal on Numerical Analysis 53*, 2 (2015), 1172–1193. MR-3338675

[23] Chassagneux, J.-F., and Richou, A. Numerical simulation of quadratic BSDEs. *The Annals of Applied Probability 26*, 1 (2016), 262–304. MR-3449318

[24] Cheridito, P., Soner, H. M., Touzi, N., and Victoir, N. Second-order backward stochastic differential equations and fully nonlinear parabolic PDEs. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences 60*, 7 (2007), 1081–1110. MR-2319056

[25] Cox, S., Hutzenthaler, M., and Jentzen, A. Local Lipschitz continuity in the initial value and strong completeness for nonlinear stochastic differential equations. arXiv:1309.5595 (2013).

[26] Cox, S., Hutzenthaler, M., Jentzen, A., van Neerven, J., and Welti, T. Convergence in Hölder norms with applications to Monte Carlo methods in infinite dimensions. *To appear in IMA J. Numer. Anal.* arXiv:1605.00856 (2016), 48 pages.

[27] Crandall, M. G., Ishii, H., and Lions, P.-L. User's guide to viscosity solutions of second order partial differential equations. *Bulletin of the American mathematical society 27*, 1 (1992), 1–67. MR-1118699

[28] Crépey, S., Gerboud, R., Grbac, Z., and Ngor, N. Counterparty risk and funding: The four wings of the TVA. *International Journal of Theoretical and Applied Finance 16*, 02 (2013), 1350006. MR-3056818

[29] Crisan, D., and Manolarakis, K. Probabilistic methods for semilinear partial differential equations. Applications to finance. *M2AN. Mathematical Modelling and Numerical Analysis 44*, 5 (2010), 1107–1133. MR-2731405

[30] Crisan, D., and Manolarakis, K. Solving backward stochastic differential equations using the cubature method: application to nonlinear pricing. *SIAM Journal on Financial Mathematics 3*, 1 (2012), 534–571. MR-2968045

[31] Crisan, D., and Manolarakis, K. Second order discretization of backward SDEs and simulation with the cubature method. *The Annals of Applied Probability 24*, 2 (2014), 652–678. MR-3178494

[32] Crisan, D., Manolarakis, K., and Touzi, N. On the Monte Carlo simulation of BSDEs: an improvement on the Malliavin weights. *Stochastic Processes and their Applications 120*, 7 (2010), 1133–1158. MR-2639741

[33] Da Prato, G., Jentzen, A., and Röckner, M. A mild Itô formula for SPDEs. *Trans. Amer. Math. Soc. 372*, 6 (2019), 3755–3807. MR-4009384

[34] Dehghan, M., Nourian, M., and Menhaj, M. B. Numerical solution of Helmholtz equation by the modified Hopfield finite difference techniques. *Numerical Methods for Partial Differential Equations: An International Journal 25*, 3 (2009), 637–656. MR-2510752

[35] Delarue, F., and Menozzi, S. A forward-backward stochastic algorithm for quasi-linear PDEs. *The Annals of Applied Probability 16*, 1 (2006), 140–184. MR-2209339

[36] Douglas, Jr., J., Ma, J., and Protter, P. Numerical methods for forward-backward stochastic differential equations. *The Annals of Applied Probability 6*, 3 (1996), 940–968. MR-1410123

[37] Duffie, D., Schroder, M., Skiadas, C., et al. Recursive valuation of defaultable securities and the timing of resolution of uncertainty. *The Annals of Applied Probability 6*, 4 (1996), 1075–1090. MR-1422978

[38] E, W., Han, J., and Jentzen, A. Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. *Communications in Mathematics and Statistics 5*, 4 (2017), 349–380. MR-3736669

[39] E, W., Hutzenthaler, M., Jentzen, A., and Kruse, T. Multilevel Picard iterations for solving smooth semilinear parabolic heat equations. arXiv:1607.03295 (2016), 18 pages.

[40] E, W., Hutzenthaler, M., Jentzen, A., and Kruse, T. On Multilevel Picard Numerical Approximations for High-Dimensional Nonlinear Parabolic Partial Differential Equations and High-Dimensional Nonlinear Backward Stochastic Differential Equations. *Journal of Scientific Computing* (Mar 2019). MR-3946468

[41] Fahim, A., Touzi, N., and Warin, X. A probabilistic numerical method for fully nonlinear parabolic PDEs. *The Annals of Applied Probability 21*, 4 (2011), 1322–1364. MR-2857450

[42] Fu, Y., Zhao, W., and Zhou, T. Efficient spectral sparse grid approximations for solving multidimensional forward backward SDEs. *Discrete and Continuous Dynamical Systems. Series B. A Journal Bridging Mathematics and Sciences 22*, 9 (2017), 3439–3458. MR-3674067

[43] Geiss, C., and Labart, C. Simulation of BSDEs with jumps by Wiener chaos expansion. *Stochastic Processes and their Applications 126*, 7 (2016), 2123–2162. MR-3483749

[44] Geiss, S., and Ylinen, J. Decoupling on the Wiener space, related Besov spaces, and applications to BSDEs. arXiv:1409.5322 (2014).

[45] Gobet, E., and Labart, C. Solving BSDE with adaptive control variate. *SIAM Journal on Numerical Analysis 48*, 1 (2010), 257–277. MR-2608369

[46] Gobet, E., and Lemor, J.-P. Numerical simulation of BSDEs using empirical regression methods: theory and practice. arXiv:0806.4447 (2008).

[47] Gobet, E., Lemor, J.-P., Warin, X., et al. A regression-based Monte Carlo method to solve backward stochastic differential equations. *The Annals of Applied Probability 15*, 3 (2005), 2172–2202. MR-2152657

[48] Gobet, E., López-Salas, J. G., Turkedjiev, P., and Vàzquez, C. Stratified regression Monte-Carl scheme for semilinear PDEs and BSDEs with large scale parallelization on GPUs. *SIAM Journal on Scientific Computing 38*, 6 (2016), C652–C677. MR-3573315

[49] Gobet, E., and Turkedjiev, P. Linear regression MDP scheme for discrete backward stochastic differential equations under general conditions. *Mathematics of Computation 85*, 299 (2016), 1359–1391. MR-3454368

[50] Gobet, E., Turkedjiev, P., et al. Approximation of backward stochastic differential equations using Malliavin weights and least-squares regression. *Bernoulli 22*, 1 (2016), 530–562. MR-3449792

[51] Graham, C., and Talay, D. *Stochastic simulation and Monte Carlo methods: mathematical foundations of stochastic simulation*, vol. 68. Springer Science & Business Media, 2013. MR-3097957

[52] Grohs, P., Hornung, F., Jentzen, A., and von Wurstemberger, P. A proof that artificial neural networks overcome the curse of dimensionality in the numerical approximation of Black-Scholes partial differential equations. arXiv:1809.02362 (2018), 124 pages.

[53] Guo, W., Zhang, J., and Zhuo, J. A monotone scheme for high-dimensional fully nonlinear PDEs. *The Annals of Applied Probability 25*, 3 (2015), 1540–1580. MR-3325281

[54] Han, J., Jentzen, A., and E, W. Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences 115*, 34 (2018), 8505–8510. MR-3847747

[55] He, J., Li, L., Xu, J., and Zheng, C. Relu deep neural networks and linear finite elements. arXiv:1807.03973 (2018). MR-3796894

[56] Henry-Labordere, P. Counterparty Risk Valuation: A Marked Branching Diffusion Approach. arXiv:1203.2369 (2012).

[57] Henry-Labordere, P. Deep primal-dual algorithm for BSDEs: Applications of machine learning to CVA and IM. *Available at SSRN 3071506* (2017).

[58] Henry-Labordere, P., Oudjane, N., Tan, X., Touzi, N., Warin, X., et al. Branching diffusion representation of semilinear pdes and monte carlo approximation. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* (2019), vol. 55, Institut Henri Poincaré, pp. 184–210. MR-3901645

[59] Henry-Labordere, P., Tan, X., and Touzi, N. A numerical algorithm for a class of BSDEs via the branching process. *Stochastic Processes and their Applications 124*, 2 (2014), 1112–1140. MR-3138609

[60] Huijskens, T., Ruijter, M. J., and Oosterlee, C. W. Efficient numerical Fourier methods for coupled forward–backward SDEs. *Journal of Computational and Applied Mathematics 296* (2016), 593–612. MR-3430160

[61] Huré, C., Pham, H., and Warin, X. Some machine learning schemes for high-dimensional nonlinear PDEs. arXiv:1902.01599 (2019). MR-4081911

[62] Hutzenthaler, M., Jentzen, A., Kruse, T., and Nguyen, T. A. A proof that rectified deep neural networks overcome the curse of dimensionality in the numerical approximation of semilinear heat equations. *SN Partial Differ. Equ. Appl. 1*, 10 (2020).

[63] Hutzenthaler, M., Jentzen, A., Kruse, T., Nguyen, T. A., and von Wurstemberger, P. Overcoming the curse of dimensionality in the numerical approximation of semilinear parabolic partial differential equations. arXiv:1807.01212 (2018), 27 pages.

[64] Hutzenthaler, M., and Kruse, T. Multilevel Picard approximations of high-dimensional semilinear parabolic differential equations with gradient-dependent nonlinearities. *SIAM Journal on Numerical Analysis 58*, 2 (2020), 929–961. MR-4075337

[65] Jentzen, A., and Kloeden, P. E. *Taylor approximations for stochastic partial differential equations*, vol. 83. SIAM, 2011. MR-2856611

[66] Jentzen, A., and von Wurstemberger, P. Lower error bounds for the stochastic gradient descent optimization algorithm: Sharp convergence rates for slowly and fast decaying learning rates. arXiv:1803.08600 (2018), 42 pages. MR-4055054

[67] Jentzen, A., Welti, T., and Salimova, D. Strong convergence for explicit space-time discrete numerical approximation methods for stochastic Burgers equations. *J. Math. Anal. Appl. 469*, 2 (2019), 661–704. MR-3860443

[68] Karatzas, I., and Shreve, S. *Brownian motion and stochastic calculus*, vol. 113. Springer Science & Business Media, 2012. MR-3184878

[69] Khoo, Y., Lu, J., and Ying, L. Solving parametric PDE problems with artificial neural networks. arXiv:1707.03351 (2017).

[70] Klenke, A. *Probabilitly Theory*, 2 ed. Universitext. Springer-Verlag London Ltd., 2014. MR-3112259

[71] Kloeden, P. E., and Platen, E. *Numerical solution of stochastic differential equations*, vol. 23. Springer Science & Business Media, 2013. MR-1214374

[72] Kong, T., Zhao, W., and Zhou, T. Probabilistic high order numerical schemes for fully nonlinear parabolic PDEs. *Communications in Computational Physics 18*, 5 (2015), 1482–1503. MR-3427591

[73] Labart, C., and Lelong, J. A parallel algorithm for solving BSDEs. *Monte Carlo Methods and Applications 19*, 1 (2013), 11–39. MR-3039401

[74] Lemor, J.-P., Gobet, E., Warin, X., et al. Rate of convergence of an empirical regression method for solving generalized backward stochastic differential equations. *Bernoulli 12*, 5 (2006), 889–916. MR-2265667

[75] Lionnet, A., Dos Reis, G., Szpruch, L., et al. Time discretization of FBSDE with polynomial growth drivers and reaction–diffusion PDEs. *The Annals of Applied Probability 25*, 5 (2015), 2563–2625. MR-3375884

[76] Ma, J., Protter, P., San Martín, J., and Torres, S. Numerical method for backward stochastic differential equations. *The Annals of Applied Probability 12*, 1 (2002), 302–316. MR-1890066

[77] Ma, J., Protter, P., and Yong, J. M. Solving forward-backward stochastic differential equations explicitly—a four step scheme. *Probability Theory and Related Fields 98*, 3 (1994), 339–359. MR-1262970

[78] Ma, J., and Yong, J. *Forward-backward stochastic differential equations and their applications*, vol. 1702 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1999. MR-1704232

[79] McKean, H. P. Application of Brownian Motion to the Equation of Kolmogorov-Petrovskii-Piskunov. *Communications on pure and applied mathematics 28*, 3 (1975), 323–331. MR-0400428

[80] Merton, R. C. Theory of rational option pricing. *Theory of Valuation* (1973), 229–288. MR-0496534

[81] Milstein, G. N. *Numerical integration of stochastic differential equations*, vol. 313. Springer Science & Business Media, 1994. MR-1335454

[82] Milstein, G. N., and Tretyakov, M. V. Numerical algorithms for forward-backward stochastic differential equations. *SIAM Journal on Scientific Computing 28*, 2 (2006), 561–582. MR-2231721

[83] Milstein, G. N., and Tretyakov, M. V. Discretization of forward-backward stochastic differential equations and related quasi-linear parabolic equations. *IMA Journal of Numerical Analysis 27*, 1 (2007), 24–44. MR-2289270

[84] Nabian, M. A., and Meidani, H. A deep neural network surrogate for high-dimensional random partial differential equations. arXiv:1806.02957 (2018).

[85] Novak, E., and Woźniakowski, H. *Tractability of Multivariate Problems: Standard information for functionals*, vol. 12. European Mathematical Society, 2008. MR-2676032

[86] Pardoux, E., and Peng, S. Adapted solution of a backward stochastic differential equation. *Systems & Control Letters 14*, 1 (1990), 55–61. MR-1037747

[87] Pardoux, E., and Peng, S. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic partial differential equations and their applications*. Springer, 1992, pp. 200–217. MR-1176785

[88] Pardoux, E., and Tang, S. Forward-backward stochastic differential equations and quasilinear parabolic PDEs. *Probability Theory and Related Fields 114*, 2 (1999), 123–150. MR-1701517

[89] Pham, H. Feynman-Kac representation of fully nonlinear PDEs and applications. *Acta Mathematica Vietnamica 40*, 2 (2015), 255–269. MR-3366170

[90] Prévôt, C., and Röckner, M. *A concise course on stochastic partial differential equations*, vol. 1905. Springer, 2007. MR-2329435

[91] Raissi, M. Forward-backward stochastic neural networks: Deep learning of high-dimensional partial differential equations. arXiv:1804.07010 (2018).

[92] Rasulov, A., Raimova, G., and Mascagni, M. Monte Carlo solution of Cauchy problem for a nonlinear parabolic equation. *Mathematics and Computers in Simulation 80*, 6 (2010), 1118–1123. MR-2610073

[93] Ruijter, M. J., and Oosterlee, C. W. A Fourier cosine method for an efficient computation of solutions to BSDEs. *SIAM Journal on Scientific Computing 37*, 2 (2015), A859–A889. MR-3336306

[94] Ruijter, M. J., and Oosterlee, C. W. Numerical Fourier method and second-order Taylor scheme for backward SDEs in finance. *Applied Numerical Mathematics. An IMACS Journal 103* (2016), 1–26. MR-3458021

[95] Ruszczynski, A., and Yao, J. A dual method for backward stochastic differential equations with application to risk valuation. arXiv:1701.06234 (2017).

[96] Sirignano, J., and Spiliopoulos, K. DGM: A deep learning algorithm for solving partial differential equations. *Journal of Computational Physics 375* (2018), 1339–1364. MR-3874585

[97] Skorokhod, A. V. Branching diffusion processes. *Theory of Probability & Its Applications 9*, 3 (1964), 445–449. MR-0168030

[98] Tadmor, E. A review of numerical methods for nonlinear partial differential equations. *Bulletin of the American Mathematical Society 49*, 4 (2012), 507–554. MR-2958929

[99] Thomée, V. *Galerkin finite element methods for parabolic problems*, vol. 1054. Springer, 1984. MR-0744045

[100] Turkedjiev, P. Two algorithms for the discrete time approximation of Markovian backward stochastic differential equations under local conditions. *Electronic Journal of Probability 20* (2015), no. 50, 49. MR-3347919

[101] Von Petersdorff, T., and Schwab, C. Numerical solution of parabolic equations in high dimensions. *ESAIM: Mathematical Modelling and Numerical Analysis 38*, 1 (2004), 93–127. MR-2073932

[102] Warin, X. Variations on branching methods for non linear PDEs. arXiv:1701.07660 (2017).

[103] Warin, X. Monte Carlo for high-dimensional degenerated Semi Linear and Full Non Linear PDEs. arXiv:1805.05078 (2018). MR-3891748

[104] Warin, X. Nesting Monte Carlo for high-dimensional non-linear PDEs. *Monte Carlo Methods and Applications 24*, 4 (2018), 225–247. MR-3891748

[105] Watanabe, S. On the branching process for Brownian particles with an absorbing boundary. *Journal of Mathematics of Kyoto University 4*, 2 (1965), 385–398. MR-0178505

[106] Wu, Z., and Yu, Z. Probabilistic interpretation for a system of quasilinear parabolic partial differential equation combined with algebra equations. *Stochastic Processes and their Applications 124*, 12 (2014), 3921–3947. MR-3264433

[107] Zhang, G., Gunzburger, M., and Zhao, W. A sparse-grid method for multi-dimensional backward stochastic differential equations. *Journal of Computational Mathematics 31*, 3 (2013), 221–248. MR-3063734

[108] Zhang, J. A numerical scheme for BSDEs. *The Annals of Applied Probability 14*, 1 (2004), 459–488. MR-2023027

[109] Zhao, W., Zhou, T., and Kong, T. High order numerical schemes for second-order FBSDEs with applications to stochastic optimal control. *Communications in Computational Physics 21*, 3 (2017), 808–834. MR-3639253

Acknowledgments. This project has been partially funded by the SNSF-Research project 200020_175699 “Higher order numerical approximation methods for stochastic partial differential equations”. The second author acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy EXC 2044-390685587, Mathematics Muenster: Dynamics-Geometry-Structure. The first author acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) via research grant HU 1889/6-1.