Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Seminar for Applied Mathematics

Overcoming the curse of dimensionality in the numerical approximation of semilinear parabolic partial differential equations

M. Hutzenthaler and A. Jentzen and Th. Kruse and T. Nguyen and

Ph. von Wurstemberger

Research Report No. 2019-46
September 2019

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

Overcoming the curse of dimensionality in the numerical approximation of semilinear parabolic partial differential equations

Martin Hutzenthaler, Arnulf Jentzen, Thomas Kruse, Tuan Anh Nguyen, and Philippe von Wurstemberger

September 3, 2019

Abstract

For a long time it is well-known that high-dimensional linear parabolic partial differential equations (PDEs) can be approximated by Monte Carlo methods with a computational effort which grows polynomially both in the dimension and in the reciprocal of the prescribed accuracy. In other words, linear PDEs do not suffer from the curse of dimensionality. For general semilinear PDEs with Lipschitz coefficients, however, it remained an open question whether these suffer from the curse of dimensionality. In this paper we partially solve this open problem. More precisely, we prove in the case of semilinear heat equations with gradient-independent and globally Lipschitz continuous nonlinearities that the computational effort of a variant of the recently introduced multilevel Picard approximations grows polynomially both in the dimension and in the reciprocal of the required accuracy.

Contents

AMS 2010 subject classification: 60H35; 65C30; 65M75

Key words and phrases: curse of dimensionality, high-dimensional PDEs, information based complexity, tractability of multivariate problems, high-dimensional semilinear BSDEs, multilevel Picard approximations, multilevel Monte Carlo method

1 Introduction and main results

Parabolic partial differential equations (PDEs) are a fundamental tool in applied mathematics for modelling phenomena in engineering, natural sciences, and man-made complex systems. For instance, semilinear PDEs appear in derivative pricing models which incorporate nonlinear risks such as default risks, interest rate risks, or liquidity risks, and PDEs are employed to model reaction diffusion systems in chemical engineering. The PDEs appearing in the above examples are often high-dimensional where the dimension corresponds to the number of financial assets such as stocks, commodities, interest rates, or exchange rates in the involved hedging portfolio.

In the literature, there exists no result which shows that essentially any of the high-dimensional semilinear PDEs appearing in the above mentioned applications can efficiently be solved approximately. More precisely, to the best of our knowledge, there exists no result in the literature which shows in the case of general semilinear PDEs with globally Lipschitz continuous coefficients that the computational effort of an approximation algorithm grows at most polynomially in both the PDE dimension and the reciprocal of the prescribed approximation accuracy. In this sense no numerical algorithm is known to not suffer from the so-called curse of dimensionality, see also the discussion after Theorem 1.1 below for details.

In this work we overcome the curse of dimensionality in the numerical approximation of semilinear heat equations with gradient-independent and globally Lipschitz continuous nonlinearities. As approximation algorithm we analyze a variant of the recently introduced multilevel Picard approximations in E et al. [11], see (1) below for the method and the paragraph after Theorem 1.1 below for a motivation hereof. The main result of this article (Theorem 1.1 below) shows in the case of general semilinear heat equations with gradient-independent and globally Lipschitz continuous nonlinearities that the computational effort of the proposed approximation algorithm grows at most polynomially in both the PDE dimension $d \in \mathbb{N}$ and the reciprocal of the required approximation accuracy $\varepsilon > 0$. More specifically, Theorem 3.8 below proves for every arbitrarily small $\delta \in (0, \infty)$ that there exists $C \in (0, \infty)$ such that for every PDE dimension $d \in \mathbb{N}$ we have that the computational cost of the proposed approximation algorithm (see (1) below) to achieve an approximation accuracy of size $\varepsilon > 0$ is bounded by $Cd^{1+p(1+\delta)}\varepsilon^{-2(1+\delta)}$, where the parameter $p \in [0, \infty)$ corresponds to the polynomial growth of the terminal condition and the nonlinearity of the PDE under consideration (see Theorem 1.1 below for details). This is essentially (up to an arbitrarily small real number $\delta \in (0, \infty)$) the same computational complexity as the plain vanilla Monte Carlo algorithm (see, e.g., [14, 19, 20, 8, 16]) achieves in the case of *linear* heat equations. In particular, in the language of information-based complexity (see, e.g., Novak & Wozniakowski [27]) this work proves, for the first time, that general semilinear heat equations with gradient-independent and globally Lipschitz continuous nonlinearities and polynomially growing terminal conditions are polynomially tractable in the setting of stochastic approximation algorithms (cf., for instance, Novak & Wozniakowski [27]). To illustrate the contribution of this article, we now present in the following result, Theorem 1.1 below, a special case of Theorem 3.8 below, which is the main result of article.

Theorem 1.1. *Let $T \in (0, \infty)$, $L, p \in [0, \infty)$, $\Theta = \cup_{n=1}^{\infty} \mathbb{Z}^n$, let $\xi_d \in \mathbb{R}^d$, $d \in \mathbb{N}$, satisfy $\sup_{d\in\mathbb{N}} \|\xi_d\|_{\mathbb{R}^d} < \infty$, let $g_d \colon \mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$, and $f_d \colon [0, T] \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$, $d \in \mathbb{N}$, be continuous functions which satisfy for all $t \in [0, T]$, $d \in \mathbb{N}$, $x \in \mathbb{R}^d$, $v, w \in \mathbb{R}$ that $|f_d(t, x, 0)| + |g_d(x)| \leq L(1 + \|x\|_{\mathbb{R}^d}^p)$ and $|f_d(t, x, v) - f_d(t, x, w)| \leq L|v - w|$, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $W^{d,\theta} \colon [0, T] \times \Omega \to \mathbb{R}^d$, $\theta \in \Theta$, $d \in \mathbb{N}$, be independent standard Brownian motions, let $\mathcal{R}^\theta \colon [0, T] \times \Omega \to \mathbb{R}$, $\theta \in \Theta$, be i.i.d. continuous stochastic processes which satisfy for all $t \in [0, T]$, $\theta \in \Theta$ that $\mathcal{R}_t^\theta \in [t, T]$ is uniformly distributed on $[t, T]$, assume that $(\mathcal{R}^\theta)_{\theta\in\Theta}$ and $(W^{d,\theta})_{\theta\in\Theta,d\in\mathbb{N}}$ are independent, let $U_{n,M}^{d,\theta} \colon [0, T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}$, $n, M \in \mathbb{Z}$, $\theta \in \Theta$, $d \in \mathbb{N}$, satisfy* *for all* $d, n, M \in \mathbb{N}$, $\theta \in \Theta$, $t \in [0,T]$, $x \in \mathbb{R}^d$ *that* $U^{d,\theta}_{-1,M}(t,x) = U^{d,\theta}_{0,M}(t,x) = 0$ *and*

$$U^{d,\theta}_{n,M}(t,x) = \Bigg[\sum_{l=0}^{n-1} \tfrac{(T-t)}{M^{n-l}} \sum_{i=1}^{M^{n-l}} f_d\big(\mathcal{R}^{(\theta,l,i)}_t, x + W^{d,(\theta,l,i)}_{\mathcal{R}^{(\theta,l,i)}_t - t}, U^{d,(\theta,l,i)}_{l,M}(\mathcal{R}^{(\theta,l,i)}_t, x + W^{d,(\theta,l,i)}_{\mathcal{R}^{(\theta,l,i)}_t - t})\big) \\ - \mathbb{1}_{\mathbb{N}}(l) f_d\big(\mathcal{R}^{(\theta,l,i)}_t, x + W^{d,(\theta,l,i)}_{\mathcal{R}^{(\theta,l,i)}_t - t}, U^{d,(\theta,l,i)}_{l-1,M}(\mathcal{R}^{(\theta,l,i)}_t, x + W^{d,(\theta,l,i)}_{\mathcal{R}^{(\theta,l,i)}_t - t})\big) \Bigg] + \sum_{i=1}^{M^n} \frac{g_d(x + W^{d,(\theta,0,-i)}_{T-t})}{M^n}, \tag{1}$$

and for every $d, n \in \mathbb{N}$, $\theta \in \Theta$, $t \in [0,T]$, $x \in \mathbb{R}^d$ *let* $\mathrm{Cost}_{d,n} \in \mathbb{N}$ *be the number of realizations of scalar standard normal random variables which are used to compute one realization of* $U^{d,\theta}_{n,n}(t,x)$ *(see* (127) *below for a precise definition). Then*

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

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

with $u_d(T,x) = g_d(x)$ *for* $t \in (0,T)$, $x \in \mathbb{R}^d$ *and*

(ii) for every $\delta \in (0,\infty)$ *there exist* $n \colon \mathbb{N} \times (0,\infty) \to \mathbb{N}$ *and* $C \in (0,\infty)$ *such that for all* $d \in \mathbb{N}$, $\varepsilon \in (0,\infty)$ *it holds that* $\mathrm{Cost}_{d,n_{d,\varepsilon}} \le C d^{1+p(1+\delta)} \varepsilon^{-2(1+\delta)}$ *and*

$$\big(\mathbb{E}\big[|u_d(0,\xi_d) - U^{d,0}_{n_{d,\varepsilon},n_{d,\varepsilon}}(0,\xi_d)|^2\big]\big)^{1/2} \le \varepsilon. \tag{3}$$

Theorem 1.1 is an immediate consequence of Theorem 3.8, Corollary 3.11, and the Feynman-Kac formula. We now motivate the multilevel Picard approximations in (1). For this assume the setting of Theorem 1.1 and let $d \in \mathbb{N}$. The Feynman-Kac formula then implies that the exact solution u_d of the PDE (2) satisfies for all $t \in (0,T)$, $x \in \mathbb{R}^d$ that

$$u_d(t,x) = \mathbb{E}\Big[g_d(x + W^{d,0}_{T-t})\Big] + \int_t^T \mathbb{E}\Big[f_d(s, x + W^{d,0}_{s-t}, u_d(s, x + W^{d,0}_{s-t}))\Big]\, ds. \tag{4}$$

This is a fixed-point equation for u_d. To this fixed-point equation we apply the well-known Picard approximation method and a telescope sum and let $u_{d,n} \colon [0,T] \times \mathbb{R}^d \to \mathbb{R}$, $n \in \mathbb{Z}$, be functions which satisfy for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $n \in \mathbb{N}$ that $u_{d,-1}(t,x) = u_{d,0}(t,x) = 0$ and that

$$\begin{aligned} &u_{d,n}(t,x) - \mathbb{E}\Big[g_d(x + W^{d,0}_{T-t})\Big] \\ &= \int_t^T \mathbb{E}\Big[f_d(s, x + W^{d,0}_{s-t}, u_{d,n-1}(s, x + W^{d,0}_{s-t}))\Big]\, ds. \\ &= \sum_{l=0}^{n-1} \int_t^T \mathbb{E}\Big[f_d(s, x + W^{d,0}_{s-t}, u_{d,l}(s, x + W^{d,0}_{s-t})) - \mathbb{1}_{\mathbb{N}}(l) f_d(s, x + W^{d,0}_{s-t}, u_{d,l-1}(s, x + W^{d,0}_{s-t}))\Big]\, ds. \end{aligned} \tag{5}$$

Next we apply a multilevel Monte Carlo approach to the non-discrete expectations and time integrals. The crucial idea for this is that the summands on the right-hand side of (5) are cheap to calculate for small $l \in \mathbb{N}_0$ and are small for large $l \in \mathbb{N}_0$ since then $u_{d,l} - u_{d,l-1}$ is small. For this reason, for every $n \in \mathbb{N}$ we approximate the expectation and the time integral on level $l \in \mathbb{N}_0$ with an average over M^{n-l} independent copies for the n-th approximation. This motivates the multilevel Picard approximations (1). For more details on the derivation of the multilevel Picard approximations see E et al. [11]. The main difference between the method (1) and the method introduced in [11] is that here we approximate time integrals by the Monte

Carlo method (this is inspired by [23, 32]) instead of quadrature rules with fixed time grids, and this modification simplifies the analysis considerably.

Next we relate Theorem 1.1 to results in the literature. Classical deterministic methods such as finite elements or sparse grid methods suffer from the curse of dimensionality. Also methods based on backward stochastic differential equations (introduced in Pardoux & Peng [28]) such as the Malliavin calculus based regression method (introduced in Bouchard & Touzi [6]), the projection on function spaces method (introduced in Gobet et al. [15]), cubature on Wiener space (introduced in Crisan & Manolarakis [9]), or the Wiener chaos decomposition method (introduced in Briand & Labart [7]) have not been shown to not suffer from the curse of dimensionality, see Subsections 4.3–4.6 in E et al. [12] for a more detailed discussion. Moreover, recently a nested Monte Carlo method has been proposed in Warin [32, 31]. Simulations show that the nested Monte Carlo method is efficient for non-large T but the method has not been shown to not suffer from the curse of dimensionality. Branching diffusion methods (cf., e.g., [21, 24, 23, 5]) exploit that solutions of semilinear PDEs with polynomial nonlinearities are equal to expectations of certain functionals of branching diffusion processes and these expectations are then approximated by the Monte Carlo method. Branching diffusion methods have been shown to not suffer from the curse of dimensionality under restrictive conditions on the initial value, on the time horizon and on the nonlinearity; see, e.g., Henry-Labordere et al. [23, Theorem 3.12]. If these conditions are not satisfied, then the approximations have not been shown to not suffer from the curse of dimensionality and simulations, e.g., for Allen-Cahn equations, indicate that the method fails to converge in this case. Moreover, the multilevel Picard approximations introduced in E et al. [11] have been shown to not suffer from the curse of dimensionality under very restrictive assumptions on the regularity of the exact solution; see [11, 25]. In addition, numerical simulations for deep learning based numerical approximation methods for PDEs (cf., for example, [10, 17, 3, 33, 13, 18, 22, 29, 30, 4, 2]) indicate that such approximation methods seem to overcome the curse of dimensionality in the numerical approximation of nonlinear PDEs but there exist no rigorous mathematical results which demonstrate this conjecture. To the best of our knowledge, the scheme (1) in Theorem 1.1 above is the first numerical approximation scheme in the scientific literature for which it has been proven that it overcomes the curse of dimensionality in the numerical approximation of general gradient-independent semilinear heat PDEs.

2 Analysis of semi-norms

2.1 Setting

Throughout this section we frequently consider the following setting.

Setting 2.1. *Let $d \in \mathbb{N}$, $T \in (0,\infty)$, $L \in [0,\infty)$, $\xi \in \mathbb{R}^d$, let $F\colon C([0,T] \times \mathbb{R}^d,\mathbb{R}) \to C([0,T] \times \mathbb{R}^d,\mathbb{R})$ be a function which satisfies for all $u,v \in C([0,T] \times \mathbb{R}^d,\mathbb{R})$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that*

$$|(F(u))(t,x) - (F(v))(t,x)| \leq L\,|u(t,x) - v(t,x)|\,, \tag{6}$$

let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $\mathbf{W}\colon [0,T]\times\Omega \to \mathbb{R}^d$ be a standard Brownian motion with continuous sample paths, and for every $k \in \mathbb{N}_0$ and every $(\mathcal{B}([0,T]\times\mathbb{R}^d)\otimes\mathcal{F})/\mathcal{B}(\mathbb{R})$-measurable function $V\colon [0,T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}$ let $\|V\|_k \in [0,\infty]$ be the extended real number given by

$$\|V\|_k^2 = \begin{cases} \mathbb{E}\big[|V(0,\xi)|^2\big] & : k = 0 \\ \frac{1}{T^k}\int_0^T \tfrac{t^{k-1}}{(k-1)!}\,\mathbb{E}\big[|V(t,\xi+\mathbf{W}_t)|^2\big]\,dt & : k \geq 1. \end{cases} \tag{7}$$

2.2 Expectations of random fields

In this subsection we derive in Lemma 2.2, Lemma 2.3, Lemma 2.4, and Corollary 2.5 below some elementary consequences of Fubini's theorem.

Lemma 2.2. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\mathcal{G} \subseteq \mathcal{F}$ be a sigma-algebra on Ω, let $(S, \mathcal{S})$ be a measurable space, let $U = (U(s))_{s \in S} = (U(s, \omega))_{s \in S, \omega \in \Omega} \colon S \times \Omega \to [0, \infty)$ be an $(\mathcal{S} \otimes \mathcal{G})/\mathcal{B}([0, \infty))$-measurable function, let $Y \colon \Omega \to S$ be a $\mathcal{F}/\mathcal{S}$-measurable function, assume that Y and $\mathcal{G}$ are independent, and let $\Phi \colon S \to [0, \infty]$ be the function which satisfies for all $s \in S$ that $\Phi(s) = \mathbb{E}[U(s)]$. Then*

(i) it holds that $U(Y) = (\Omega \ni \omega \mapsto U(Y(\omega), \omega) \in [0, \infty))$ is an $\mathcal{F}/\mathcal{B}([0, \infty))$-measurable function and

(ii) it holds that

$$\mathbb{E}[U(Y)] = \mathbb{E}[\Phi(Y)] = \int_S \mathbb{E}[U(s)] \, (Y(\mathbb{P})_{\mathcal{S}})(ds). \tag{8}$$

Proof of Lemma 2.2. Throughout this proof let $Z \colon \Omega \to \Omega$ be the function which satisfies for all $\omega \in \Omega$ that $Z(\omega) = \omega$. Observe that the hypothesis that $\mathcal{G} \subseteq \mathcal{F}$ ensures that Z is an $\mathcal{F}/\mathcal{G}$-measurable function. The hypothesis that Y is an $\mathcal{F}/\mathcal{S}$-measurable function hence proves that $\Omega \ni \omega \mapsto (Y(\omega), Z(\omega)) = (Y(\omega), \omega) \in S \times \Omega$ is an $\mathcal{F}/(\mathcal{S} \otimes \mathcal{G})$-measurable function. Combining this with the hypothesis that U is an $(\mathcal{S} \otimes \mathcal{G})/\mathcal{B}([0, \infty))$-measurable function establishes Item (i). It thus remains to prove Item (ii). For this observe that the hypothesis that Y and $\mathcal{G}$ are independent proves that

$$\big((Y, Z)\big)(\mathbb{P})_{(\mathcal{S} \otimes \mathcal{G})} = \big(Y(\mathbb{P})_{\mathcal{S}}\big) \otimes \big(Z(\mathbb{P})_{\mathcal{G}}\big). \tag{9}$$

Fubini's theorem and the integral transformation theorem hence demonstrate that

$$\begin{aligned}
\mathbb{E}[U(Y)] &= \int_\Omega U\big(Y(\omega), \omega\big) \, \mathbb{P}(\omega) = \int_\Omega U\big(Y(\omega), Z(\omega)\big) \, \mathbb{P}(\omega) \\
&= \int_{S \times \Omega} U(y, z) \, \Big(\big((Y, Z)\big)(\mathbb{P})_{(\mathcal{S} \otimes \mathcal{G})}\Big)(dy, dz) \\
&= \int_{S \times \Omega} U(y, z) \, \Big(\big(Y(\mathbb{P})_{\mathcal{S}}\big) \otimes \big(Z(\mathbb{P})_{\mathcal{G}}\big)\Big)(dy, dz) \\
&= \int_S \int_\Omega U(y, z) \, \big(Z(\mathbb{P})_{\mathcal{G}}\big)(dz) \, \big(Y(\mathbb{P})_{\mathcal{S}}\big)(dy) = \int_S \int_\Omega U(y, \omega) \, \mathbb{P}(d\omega) \, \big(Y(\mathbb{P})_{\mathcal{S}}\big)(dy) \\
&= \int_S \mathbb{E}\big[U(y)\big] \, \big(Y(\mathbb{P})_{\mathcal{S}}\big)(dy).
\end{aligned} \tag{10}$$

This establishes Item (ii). The proof of Lemma 2.2 is thus completed. □

Lemma 2.3. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let (S, δ) be a separable metric space, let $U = (U(s))_{s \in S} = (U(s, \omega))_{s \in S, \omega \in \Omega} \colon S \times \Omega \to [0, \infty)$ be a continuous random field, let $Y : \Omega \to S$ be a random variable, assume that U and Y are independent, and let $\Phi \colon S \to [0, \infty]$ be the function which satisfies for all $s \in S$ that $\Phi(s) = \mathbb{E}[U(s)]$. Then*

(i) it holds that $U(Y) = (\Omega \ni \omega \mapsto U(Y(\omega), \omega) \in [0, \infty))$ is an $\mathcal{F}/\mathcal{B}([0, \infty))$-measurable function and

(ii) it holds that

$$\mathbb{E}[U(Y)] = \mathbb{E}[\Phi(Y)] = \int_S \mathbb{E}[U(s)] \, (Y(\mathbb{P})_{\mathcal{B}(S)})(ds). \tag{11}$$

Proof of Lemma 2.3. First, observe that the hypothesis that U is a continuous random field and [2, Lemma 2.4] (with $(E,d)=(S,\delta)$, $\mathcal{E}=[0,\infty)$, $(\Omega,\mathcal{F})=(\Omega,\sigma_\Omega(U))$, $X=U$ in the notation of [2, Lemma 2.4]) assure that U is an $(\mathcal{S}\otimes\sigma_\Omega(U))/\mathcal{B}([0,\infty))$-measurable function. Combining this and the hypothesis that $\sigma_\Omega(U)$ and Y are independent with Lemma 2.2 establishes Items (i)–(ii). The proof of Lemma 2.3 is thus completed. ☐

Lemma 2.4. *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let (S,δ) be a separable metric space, let $U=(U(s))_{s\in S}=(U(s,\omega))_{s\in S,\omega\in\Omega}\colon S\times\Omega\to\mathbb{R}$ be a continuous random field, let $Y\colon\Omega\to S$ be a random variable, assume that U and Y are independent, and assume that $\int_S\mathbb{E}[|U(s)|]\,(Y(\mathbb{P})_{\mathcal{B}(S)})(ds)<\infty$. Then*

(i) it holds that $U(Y)=(\Omega\ni\omega\mapsto U(Y(\omega),\omega)\in\mathbb{R})$ is an $\mathcal{F}/\mathcal{B}(\mathbb{R})$-measurable function and

(ii) it holds that $\mathbb{E}[|U(Y)|]<\infty$ and

$$\mathbb{E}[U(Y)]=\int_S\mathbb{1}_{\{\mathfrak{s}\in S\colon\mathbb{E}[|U(\mathfrak{s})|]<\infty\}}(s)\,\mathbb{E}[U(s)]\,(Y(\mathbb{P})_{\mathcal{B}(S)})(ds).\tag{12}$$

Proof of Lemma 2.4. Throughout this proof let $\mathcal{U},\mathfrak{U}\colon S\times\Omega\to[0,\infty)$ be the functions which satisfy for all $s\in S$, $\omega\in\Omega$ that

$$\mathcal{U}(s,\omega)=\max\{U(s,\omega),0\}\qquad\text{and}\qquad\mathfrak{U}(s,\omega)=\max\{-U(s,\omega),0\}.\tag{13}$$

Observe that $U=\mathcal{U}-\mathfrak{U}$. Item (i) in Lemma 2.3 hence implies that $U(Y)=\mathcal{U}(Y)-\mathfrak{U}(Y)$ is $\mathcal{F}/\mathcal{B}(\mathbb{R})$-measurable. This proves Item (i). In addition, note that Item (ii) in Lemma 2.3 and the fact that $|U|=\mathcal{U}+\mathfrak{U}$ assures that

$$\mathbb{E}[\mathcal{U}(Y)]+\mathbb{E}[\mathfrak{U}(Y)]=\mathbb{E}[|U(Y)|]=\int_S\mathbb{E}[|U(s)|]\,(Y(\mathbb{P})_{\mathcal{B}(S)})(ds)<\infty.\tag{14}$$

Moreover, note that the hypothesis that $\int_S\mathbb{E}[|U(s)|]\,(Y(\mathbb{P})_{\mathcal{B}(S)})(ds)<\infty$ ensures that

$$(Y(\mathbb{P})_{\mathcal{B}(S)})(\{s\in S\colon\mathbb{E}[|U(s)|]=\infty\})=0.\tag{15}$$

Item (ii) in Lemma 2.3 and (14) therefore demonstrate that

$$\begin{aligned}\mathbb{E}[U(Y)]&=\mathbb{E}[\mathcal{U}(Y)-\mathfrak{U}(Y)]=\mathbb{E}[\mathcal{U}(Y)]-\mathbb{E}[\mathfrak{U}(Y)]\\&=\int_S\mathbb{E}[\mathcal{U}(s)]\,(Y(\mathbb{P})_{\mathcal{B}(S)})(ds)-\int_S\mathbb{E}[\mathfrak{U}(s)]\,(Y(\mathbb{P})_{\mathcal{B}(S)})(ds).\\&=\int_S\mathbb{1}_{\{\mathfrak{s}\in S\colon\mathbb{E}[|U(\mathfrak{s})|]<\infty\}}(s)\,\mathbb{E}[\mathcal{U}(s)]\,(Y(\mathbb{P})_{\mathcal{B}(S)})(ds)\\&\quad-\int_S\mathbb{1}_{\{\mathfrak{s}\in S\colon\mathbb{E}[|U(\mathfrak{s})|]<\infty\}}(s)\,\mathbb{E}[\mathfrak{U}(s)]\,(Y(\mathbb{P})_{\mathcal{B}(S)})(ds)\\&=\int_S\mathbb{1}_{\{\mathfrak{s}\in S\colon\mathbb{E}[|U(\mathfrak{s})|]<\infty\}}(s)\,(\mathbb{E}[\mathcal{U}(s)]-\mathbb{E}[\mathfrak{U}(s)])(Y(\mathbb{P})_{\mathcal{B}(S)})(ds)\\&=\int_S\mathbb{1}_{\{\mathfrak{s}\in S\colon\mathbb{E}[|U(\mathfrak{s})|]<\infty\}}(s)\,\mathbb{E}[U(s)]\,(Y(\mathbb{P})_{\mathcal{B}(S)})(ds).\end{aligned}\tag{16}$$

This establishes Item (ii). The proof of Lemma 2.4 is thus completed. ☐

Corollary 2.5. *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let (S,δ) be a separable metric space, let $(E,\mathcal{E})$ be measurable space, let $U_1,U_2\colon S\times\Omega\to\mathbb{R}$ be continuous random fields, let $Y_1,Y_2\colon E\times\Omega\to S$ be random fields, assume for all $i\in\{1,2\}$ that U_i and Y_i are independent, assume that U_1 and U_2 are identically distributed, and assume that Y_1 and Y_2 are identically distributed. Then it holds that $U_1(Y_1)=(E\times\Omega\ni(e,\omega)\mapsto U_1(Y_1(e),\omega)\in\mathbb{R})$ and $U_2(Y_2)=(E\times\Omega\ni(e,\omega)\mapsto U_2(Y_2(e),\omega)\in\mathbb{R})$ are identically distributed random fields.*

Proof of Corollary 2.5. Throughout this proof let $n \in \mathbb{N}$, $e_1, e_2, \dots, e_n \in E$, $B_1, B_2, \dots, B_n \in \mathcal{B}(\mathbb{R})$, let $\mathbb{U}_i \colon S^n \times \Omega \to \mathbb{R}^n$, $i \in \{1,2\}$, be the functions which satisfy for all $i \in \{1,2\}$, $s_1, s_2, \dots, s_n \in S$, $\omega \in \Omega$ that

$$\mathbb{U}_i(s_1, \dots, s_n, \omega) = \big(U_i(s_1, \omega), \dots, U_i(s_n, \omega)\big), \tag{17}$$

let $\mathbb{Y}_i \colon \Omega \to S^n$, $i \in \{1,2\}$, be the functions which satisfy for all $i \in \{1,2\}$, $\omega \in \Omega$ that

$$\mathbb{Y}_i(\omega) = \big(Y_i(e_1, \omega), \dots, Y_i(e_n, \omega)\big), \tag{18}$$

and let $\mathbb{I} \colon \mathbb{R}^n \to [0, \infty)$ be the function which satisfies for all $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ that

$$\mathbb{I}(x) = \mathbb{1}_{B_1}(x_1) \cdot \ldots \cdot \mathbb{1}_{B_n}(x_n). \tag{19}$$

Note that the fact that for all $i \in \{1,2\}$ it holds that $\mathbb{U}_i$ is a continuous random field and Beck et al. [2, Lemma 2.4] assure that for every $i \in \{1,2\}$ it holds that $\mathbb{U}_i$ is $(\mathcal{B}(S^n) \otimes \sigma_\Omega(\mathbb{U}_i))/\mathcal{B}(\mathbb{R}^n)$-measurable. The fact that $\mathbb{I}$ is $\mathcal{B}(\mathbb{R}^n)/\mathcal{B}([0,\infty))$-measurable hence assures that for every $i \in \{1,2\}$ it holds that $\mathbb{I} \circ \mathbb{U}_i$ is $(\mathcal{B}(S^n) \otimes \sigma_\Omega(\mathbb{U}_i))/\mathcal{B}([0,\infty))$-measurable. Combining this, the fact for every $i \in \{1,2\}$ it holds that $\mathbb{Y}_i$ is $\mathcal{F}/\mathcal{B}(S^n)$-measurable, and the hypothesis that for every $i \in \{1,2\}$ it holds that $\sigma_\Omega(U_i) = \sigma_\Omega(\mathbb{U}_i)$ and Y_i are independent with Lemma 2.2 (with $(\Omega, \mathcal{F}, \mathbb{P}) = (\Omega, \mathcal{F}, \mathbb{P})$, $\mathcal{G} = \sigma_\Omega(\mathbb{U}_i)$, $(S, \mathcal{S}) = (S^n, \mathcal{B}(S^n))$, $U = \mathbb{I} \circ \mathbb{U}_i$, $Y = \mathbb{Y}_i$ for $i \in \{1,2\}$ in the notation of Lemma 2.2) demonstrate that for all $i \in \{1,2\}$ it holds that

$$\begin{aligned}
&\mathbb{P}(U_i(Y_i(e_1)) \in B_1, \dots, U_i(Y_i(e_n)) \in B_n) \\
&= \mathbb{E}[\mathbb{1}_{B_1}(U_i(Y_i(e_1))) \cdot \ldots \cdot \mathbb{1}_{B_n}(U_i(Y_i(e_n)))] = \mathbb{E}[\mathbb{I}(\mathbb{U}_i(\mathbb{Y}_i))] \\
&= \int_{S^n} \mathbb{E}[\mathbb{I}(\mathbb{U}_i(s_1, \dots, s_n))] \; (\mathbb{Y}_i(\mathbb{P}))_{\mathcal{B}(S^n)}(d(s_1, \dots, s_n)).
\end{aligned} \tag{20}$$

In addition, observe that the hypothesis that U_1 and U_2 are identically distributed assures that for all $s_1, \dots, s_n \in S$ it holds that

$$\begin{aligned}
\mathbb{E}[\mathbb{I}(\mathbb{U}_1(s_1, \dots, s_n))] &= \mathbb{P}(U_1(s_1) \in B_1, \dots, U_1(s_n) \in B_n) \\
&= \mathbb{P}(U_2(s_1) \in B_1, \dots, U_2(s_n) \in B_n) = \mathbb{E}[\mathbb{I}(\mathbb{U}_2(s_1, \dots, s_n))] \,.
\end{aligned} \tag{21}$$

Moreover, note that the hypothesis that Y_1 and Y_2 are identically distributed ensures that

$$(\mathbb{Y}_1(\mathbb{P}))_{\mathcal{B}(S^n)} = (\mathbb{Y}_2(\mathbb{P}))_{\mathcal{B}(S^n)}. \tag{22}$$

Combining this, (20), and (21) demonstrates that

$$\begin{aligned}
&\mathbb{P}(U_1(Y_1(e_1)) \in B_1, \dots, U_1(Y_1(e_n)) \in B_n) \\
&= \int_{S^n} \mathbb{E}[\mathbb{I}(\mathbb{U}_1(s_1, \dots, s_n))] \; (\mathbb{Y}_1(\mathbb{P}))_{\mathcal{B}(S^n)}(d(s_1, \dots, s_n)) \\
&= \int_{S^n} \mathbb{E}[\mathbb{I}(\mathbb{U}_2(s_1, \dots, s_n))] \; (\mathbb{Y}_2(\mathbb{P}))_{\mathcal{B}(S^n)}(d(s_1, \dots, s_n)) \\
&= \mathbb{P}(U_2(Y_2(e_1)) \in B_1, \dots, U_2(Y_2(e_n)) \in B_n).
\end{aligned} \tag{23}$$

Hence, we obtain that $U_1(Y_1)$ and $U_2(Y_2)$ are identically distributed random fields. The proof of Corollary 2.5 is thus completed. ☐

2.3 Properties of the semi-norms

In this subsection we establish in Lemma 2.6, Lemma 2.7, Lemma 2.8, Lemma 2.9, Lemma 2.10, and Lemma 2.11 a few basic properties for the quantities in (7) in Setting 2.1 above. The proof of Lemma 2.6 is clear and therefore omitted.

Lemma 2.6 (Semi-norm property)**.** *Assume Setting 2.1, let $k \in \mathbb{N}_0$, $\lambda \in \mathbb{R}$, and let $U, V\colon [0,T]\times \mathbb{R}^d \times \Omega \to \mathbb{R}$ be $(\mathcal{B}([0,T]\times\mathbb{R}^d)\otimes\mathcal{F})/\mathcal{B}(\mathbb{R})$-measurable functions. Then*

(i) it holds that $\|U+V\|_k \leq \|U\|_k + \|V\|_k$ and

(ii) it holds that $\|\lambda U\|_k = |\lambda|\|U\|_k$.

Lemma 2.7 (Expectations)**.** *Assume Setting 2.1, let $k \in \mathbb{N}_0$, let $U\colon [0,T]\times\mathbb{R}^d\times\Omega\to\mathbb{R}$ be a continuous random field, assume that U and $\mathbf{W}$ are independent, and assume for all $t\in[0,T]$, $x\in\mathbb{R}^d$ that $\mathbb{E}[|U(t,x)|]<\infty$. Then it holds that*

$$\big\|[0,T]\times\mathbb{R}^d\times\Omega \ni (t,x,\omega)\mapsto \mathbb{E}[U(t,x)]\in\mathbb{R}\big\|_k = \|\mathbb{E}[U]\|_k \leq \|U\|_k\,. \tag{24}$$

Proof of Lemma 2.7. Throughout this proof let $v\colon [0,T]\times\mathbb{R}^d\to\mathbb{R}$ be the function which satisfies for all $t\in[0,T]$, $x\in\mathbb{R}^d$ that

$$v(t,x) = \mathbb{E}[U(t,x)] \tag{25}$$

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

$$\mu_t(B) = \mathbb{P}(\xi+\mathbf{W}_t\in B). \tag{26}$$

Note that Jensen's inequality and (7) assure that

$$\|\mathbb{E}[U]\|_0^2 = \|v\|_0^2 = \mathbb{E}\big[|v(0,\xi)|^2\big] = |v(0,\xi)|^2 = |\mathbb{E}[U(0,\xi)]|^2 \leq \mathbb{E}\big[|U(0,\xi)|^2\big] = \|U\|_0^2\,. \tag{27}$$

Next observe that (7) ensures that for all $l\in\mathbb{N}$ it holds that

$$\|\mathbb{E}[U]\|_l^2 = \|v\|_l^2 = \frac{1}{T^l}\int_0^T \tfrac{t^{l-1}}{(l-1)!}\,\mathbb{E}\big[|v(t,\xi+\mathbf{W}_t)|^2\big]\,dt. \tag{28}$$

Moreover, note that the integral transformation theorem, Jensen's inequality, Lemma 2.3, the hypothesis that U is a continuous random field, and the hypothesis that U and $\mathbf{W}$ are independent ensure that for all $t\in[0,T]$ it holds that

$$\begin{aligned}\mathbb{E}\big[|v(t,\xi+\mathbf{W}_t)|^2\big] &= \int_{\mathbb{R}^d}|v(t,x)|^2\,\mu_t(dx) = \int_{\mathbb{R}^d}|\mathbb{E}[U(t,x)]|^2\,\mu_t(dx)\\ &\leq \int_{\mathbb{R}^d}\mathbb{E}\big[|U(t,x)|^2\big]\,\mu_t(dx) = \mathbb{E}\big[|U(t,\xi+\mathbf{W}_t)|^2\big]\,.\end{aligned} \tag{29}$$

This and (28) imply that for all $l\in\mathbb{N}$ it holds that

$$\|\mathbb{E}[U]\|_l^2 \leq \frac{1}{T^l}\int_0^T \tfrac{t^{l-1}}{(l-1)!}\,\mathbb{E}\big[|U(t,\xi+\mathbf{W}_t)|^2\big]\,dt = \|U\|_l^2\,. \tag{30}$$

Combining this and (27) establishes (24). The proof of Lemma 2.7 is thus completed. ☐

Lemma 2.8 (Linear combinations of i.i.d. random variables)**.** *Assume Setting 2.1, let $k\in\mathbb{N}_0$, $n\in\mathbb{N}$, $r_1,\ldots,r_n\in\mathbb{R}$, let $U_1,\ldots,U_n\colon[0,T]\times\mathbb{R}^d\times\Omega\to\mathbb{R}$ be continuous i.i.d. random fields, assume that $(U_i)_{i\in\{1,2,\ldots,n\}}$ and $\mathbf{W}$ are independent, and assume for all $t\in[0,T]$, $x\in\mathbb{R}^d$ that $\mathbb{E}[|U_1(t,x)|]<\infty$. Then it holds that*

$$\left\|\sum_{i=1}^n r_i\,(U_i-\mathbb{E}[U_i])\right\|_k = \|U_1-\mathbb{E}[U_1]\|_k\left[\sum_{i=1}^n|r_i|^2\right]^{1/2} \leq \|U_1\|_k\left[\sum_{i=1}^n|r_i|^2\right]^{1/2}. \tag{31}$$

Proof of Lemma 2.8. Throughout this proof let $\mathcal{G} \subseteq \mathcal{F}$ satisfy that $\mathcal{G} = \sigma_\Omega((U_i)_{i\in\{1,2,\dots,n\}})$, let $v_i\colon [0,T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}$, $i \in \{1,2,\dots,n\}$, satisfy for all $i \in \{1,2,\dots,n\}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$v_i(t,x) = U_i(t,x) - \mathbb{E}[U_i(t,x)], \tag{32}$$

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

$$\mu_t(B) = \mathbb{P}(\xi + \mathbf{W}_t \in B). \tag{33}$$

Note that the fact that $v_1,\dots,v_n$ are continuous random fields, Beck et al. [2, Lemma 2.4], and Fubini's theorem imply that for every $i \in \{1,2,\dots,n\}$ it holds that v_i is a $(\mathcal{B}([0,T]\times\mathbb{R}^d)\otimes\mathcal{G})/\mathcal{B}(\mathbb{R})$-measurable function. The hypothesis that $\mathcal{G}$ and $\mathbf{W}$ are independent, Lemma 2.2, the fact that for all $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that $v_1(t,x), v_2(t,x), \dots, v_n(t,x)$ are i.i.d. random variables with $\mathbb{E}[|v_1(t,x)|] < \infty$ and $\mathbb{E}[v_1(t,x)] = 0$, and Klenke [26, Theorem 5.4] therefore demonstrate that for all $t \in [0,T]$ it holds that

$$\begin{aligned}
\mathbb{E}\Big[\big|\textstyle\sum_{i=1}^n r_i v_i(t,\xi+\mathbf{W}_t)\big|^2\Big] &= \int_{\mathbb{R}^d} \mathbb{E}\Big[\big|\textstyle\sum_{i=1}^n r_i v_i(t,x)\big|^2\Big]\,\mu_t(dx)\\
&= \int_{\mathbb{R}^d} \textstyle\sum_{i,j=1}^n \mathbb{E}[\,r_i r_j v_i(t,x) v_j(t,x)]\,\mu_t(dx)\\
&= \int_{\mathbb{R}^d} \textstyle\sum_{i=1}^n |r_i|^2\,\mathbb{E}[|v_i(t,x)|^2]\,\mu_t(dx)\\
&= \textstyle\sum_{i=1}^n \big(|r_i|^2 \int_{\mathbb{R}^d} \mathbb{E}[|v_i(t,x)|^2]\,\mu_t(dx)\big)\\
&= \Big[\textstyle\sum_{i=1}^n |r_i|^2\Big] \int_{\mathbb{R}^d} \mathbb{E}[|v_1(t,x)|^2]\,\mu_t(dx)\\
&= \Big[\textstyle\sum_{i=1}^n |r_i|^2\Big]\mathbb{E}[|v_1(t,\xi+\mathbf{W}_t)|^2].
\end{aligned}\tag{34}$$

This and (7) imply that

$$\begin{aligned}
\big\|\textstyle\sum_{i=1}^n r_i\,(U_i - \mathbb{E}[U_i])\big\|_0^2 &= \big\|\textstyle\sum_{i=1}^n r_i v_i\big\|_0^2 = \mathbb{E}\Big[\big|\textstyle\sum_{i=1}^n r_i v_i(0,\xi)\big|^2\Big]\\
&= \Big[\textstyle\sum_{i=1}^n |r_i|^2\Big]\mathbb{E}[|v_1(0,\xi)|^2] = \Big[\textstyle\sum_{i=1}^n |r_i|^2\Big]\|v_1\|_0^2\\
&= \Big[\textstyle\sum_{i=1}^n |r_i|^2\Big]\|U_1 - \mathbb{E}[U_1]\|_0^2.
\end{aligned}\tag{35}$$

Moreover, observe that (7) and (34) show that for all $l \in \mathbb{N}$ it holds that

$$\begin{aligned}
\big\|\textstyle\sum_{i=1}^n r_i\,(U_i - \mathbb{E}[U_i])\big\|_l^2 &= \big\|\textstyle\sum_{i=1}^n r_i v_i\big\|_l^2\\
&= \frac{1}{T^l}\int_0^T \tfrac{t^{l-1}}{(l-1)!}\,\mathbb{E}\Big[\big|\textstyle\sum_{i=1}^n r_i v_i(t,\xi+\mathbf{W}_t)\big|^2\Big]\,dt\\
&= \frac{1}{T^l}\int_0^T \tfrac{t^{l-1}}{(l-1)!}\,\Big[\textstyle\sum_{i=1}^n |r_i|^2\Big]\mathbb{E}\big[|v_1(t,\xi+\mathbf{W}_t)|^2\big]\,dt\\
&= \Big[\textstyle\sum_{i=1}^n |r_i|^2\Big]\|v_1\|_l^2 = \Big[\textstyle\sum_{i=1}^n |r_i|^2\Big]\|U_1 - \mathbb{E}[U_1]\|_l^2.
\end{aligned}\tag{36}$$

Next observe that (7) assures that

$$\begin{aligned}
\|U_1 - \mathbb{E}[U_1]\|_0^2 = \|v_1\|_0^2 &= \mathbb{E}\big[|v_1(0,\xi)|^2\big] = \mathbb{E}\big[|U_1(0,\xi) - \mathbb{E}[U_1(0,\xi)]\,|^2\big]\\
&= \mathbb{E}\big[|U_1(0,\xi)|^2\big] - |\mathbb{E}[U_1(0,\xi)]|^2 \le \mathbb{E}\big[|U_1(0,\xi)|^2\big] = \|U_1\|_0^2.
\end{aligned}\tag{37}$$

Furthermore, note that the hypothesis that $\mathcal{G}$ and $\mathbf{W}$ are independent and Lemma 2.2 assure that for all $t \in [0,T]$ it holds that

$$\begin{aligned} \mathbb{E}\big[|v_1(t,\xi+\mathbf{W}_t)|^2\big] &= \int_{\mathbb{R}^d} \mathbb{E}\big[|v_1(t,x)|^2\big]\ \mu_t(dx) \\ &= \int_{\mathbb{R}^d} \mathbb{E}\big[|U_1(t,x) - \mathbb{E}[U_1(t,x)]\,|^2\big]\ \mu_t(dx) \\ &= \int_{\mathbb{R}^d} \mathbb{E}\big[|U_1(t,x)|^2\big] - |\mathbb{E}[U_1(t,x)]|^2\, \mu_t(dx) \\ &\leq \int_{\mathbb{R}^d} \mathbb{E}\big[|U_1(t,x)|^2\big]\ \mu_t(dx) = \mathbb{E}\big[|U_1(t,\xi+\mathbf{W}_t)|^2\big]. \end{aligned} \tag{38}$$

This and (7) demonstrate that for all $l \in \mathbb{N}$ it holds that

$$\begin{aligned} \|U_1 - \mathbb{E}[U_1]\|_l^2 = \|v_1\|_l^2 &= \frac{1}{T^l}\int_0^T \tfrac{t^{l-1}}{(l-1)!}\, \mathbb{E}\big[|v_1(t,\xi+\mathbf{W}_t)|^2\big]\,dt \\ &\leq \frac{1}{T^l}\int_0^T \tfrac{t^{l-1}}{(l-1)!}\, \mathbb{E}\big[|U_1(t,\xi+\mathbf{W}_t)|^2\big]\,dt = \|U_1\|_l^2. \end{aligned} \tag{39}$$

Combining this, (35), (36), and (37) establishes that

$$\big\|\textstyle\sum_{i=1}^n r_i\,(U_i - \mathbb{E}[U_i])\big\|_k^2 = \Big[\sum_{i=1}^n |r_i|^2\Big]\, \|U_1 - \mathbb{E}[U_1]\|_k^2 \leq \Big[\sum_{i=1}^n |r_i|^2\Big]\, \|U_1\|_k^2. \tag{40}$$

This completes the proof of Lemma 2.8. □

Lemma 2.9 (Lipschitz property of F)**.** *Assume Setting 2.1, let $k \in \mathbb{N}_0$, and let $U, V\colon [0,T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}$ be continuous random fields. Then*

(i) it holds that $F(U) = \big([0,T]\times\mathbb{R}^d\times\Omega \ni (t,x,\omega) \mapsto \big[F\big([0,T]\times\mathbb{R}^d \ni (s,z) \mapsto U(s,z,\omega) \in \mathbb{R}\big)\big](t,x) \in \mathbb{R}\big)$ is a continuous random field and

(ii) it holds that $\|F(U) - F(V)\|_k \leq L\|U - V\|_k$.

Proof of Lemma 2.9. Throughout this proof let $\pi_{t,x}\colon C([0,T]\times\mathbb{R}^d,\mathbb{R}) \to \mathbb{R}$, $t \in [0,T]$, $x \in \mathbb{R}^d$, satisfy for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $v \in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ that

$$\pi_{t,x}(v) = v(t,x) \tag{41}$$

and let $\mathfrak{U}\colon \Omega \to C([0,T]\times\mathbb{R}^d,\mathbb{R})$ be the function which satisfies for all $\omega \in \Omega$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that $(\mathfrak{U}(\omega))(t,x) = U(t,x,\omega)$. Note that for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $\omega \in \Omega$ it holds that

$$(F(U))(t,x,\omega) = \big[F(\mathfrak{U}(\omega))\big](t,x) = \pi_{t,x}\big[F(\mathfrak{U}(\omega))\big]. \tag{42}$$

Hence, we obtain for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$\big(\Omega \ni \omega \mapsto (F(U))(t,x,\omega)\big) = \pi_{t,x} \circ F \circ \mathfrak{U}. \tag{43}$$

The fact that $\mathfrak{U}$ is $\mathcal{F}/\mathcal{B}(C([0,T]\times\mathbb{R}^d,\mathbb{R}))$-measurable, the fact that for all $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that $\pi_{t,x}$ is $\mathcal{B}(C([0,T]\times\mathbb{R}^d,\mathbb{R}))/\mathcal{B}(\mathbb{R})$-measurable, and the fact that F is $\mathcal{B}(C([0,T]\times\mathbb{R}^d,\mathbb{R}))/\mathcal{B}(C([0,T]\times\mathbb{R}^d,\mathbb{R}))$-measurable (cf. (6)) hence assure that for all $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that $(\Omega \ni \omega \mapsto (F(U))(t,x,\omega))$ is $\mathcal{F}/\mathcal{B}(\mathbb{R})$-measurable. Combining this with the fact that for all $v \in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ it holds that $F(v) \in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ demonstrates that $F(U)$ is a continuous random field. This establishes Item (i). Next observe that (6) and (7) show that

$$\begin{aligned} \|F(U) - F(V)\|_0^2 &= \mathbb{E}\big[|(F(U) - F(V))(0,\xi)|^2\big] \\ &\leq \mathbb{E}\big[L^2\,|(U - V)(0,\xi)|^2\big] = L^2\,\|U - V\|_0^2. \end{aligned} \tag{44}$$

Moreover, note that (6) and (7) imply that for all $l \in \mathbb{N}$ it holds that

$$
\begin{aligned}
\|F(U) - F(V)\|_l^2 &= \frac{1}{T^l} \int_0^T \tfrac{t^{l-1}}{(l-1)!} \, \mathbb{E}\Big[\big|\big(F(U) - F(V)\big)(t, \xi + \mathbf{W}_t)\big)\big|^2\Big] \, dt \\
&\leq \frac{1}{T^l} \int_0^T \tfrac{t^{l-1}}{(l-1)!} \, \mathbb{E}\Big[L^2 \big|(U - V)(t, \xi + \mathbf{W}_t)\big)\big|^2\Big] \, dt \\
&= L^2 \|U - V\|_l^2.
\end{aligned} \tag{45}
$$

Combining this and (44) establishes Item (ii). The proof of Lemma 2.9 is thus completed. □

Lemma 2.10 (Monte Carlo time integrals)**.** *Assume Setting 2.1, let $k \in \mathbb{N}_0$, let $U\colon [0,T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}$ be a continuous random field, let $\mathfrak{r}\colon \Omega \to [0,1]$ be a $\mathcal{U}_{[0,1]}$-distributed random variable, let $\mathcal{R}\colon [0,T] \times \Omega \to \mathbb{R}$ satisfy for all $t \in [0,T]$ that $\mathcal{R}_t = t + (T-t)\mathfrak{r}$, let $\mathbb{W}\colon [0,T] \times \Omega \to \mathbb{R}^d$ be a standard Brownian motion with continuous sample paths, and assume that $U, \mathbf{W}, \mathfrak{r}$, and $\mathbb{W}$ are independent. Then it holds that*

$$
\big\|[0,T] \times \mathbb{R}^d \times \Omega \ni (t,x,\omega) \mapsto (T-t)\big[U(\mathcal{R}_t, x + \mathbb{W}_{\mathcal{R}_t} - \mathbb{W}_t)\big](\omega) \in \mathbb{R}\big\|_k \leq T\|U\|_{k+1}. \tag{46}
$$

Proof of Lemma 2.10. Throughout this proof let $V^{(t)} = (V_s^{(t)}(\omega))_{s \in [t,T], \omega \in \Omega}\colon [t,T] \times \Omega \to \mathbb{R}$, $t \in [0,T]$, be the random fields which satisfy for all $t \in [0,T]$, $s \in [t,T]$ that

$$
V_s^{(t)} = U(s, \xi + \mathbf{W}_t + \mathbb{W}_s - \mathbb{W}_t). \tag{47}
$$

Observe that the fact that $\mathbf{W}$, $\mathbb{W}$, and U are independent, the hypothesis that U is a continuous random field, Lemma 2.3, and the fact that for all $t \in [0,T]$, $s \in [t,T]$ it holds that $\mathbf{W}_t + \mathbb{W}_s - \mathbb{W}_t$ and $\mathbf{W}_s$ are identically distributed ensure that for all $t \in [0,T]$, $s \in [t,T]$ it holds that

$$
\begin{aligned}
\mathbb{E}\big[|V_s^{(t)}|^2\big] &= \mathbb{E}\big[|U(s, \xi + \mathbf{W}_t + \mathbb{W}_s - \mathbb{W}_t)|^2\big] \\
&= \int_{\mathbb{R}^d} \mathbb{E}\big[|U(s, \xi + x)|^2\big] \, ((\mathbf{W}_t + \mathbb{W}_s - \mathbb{W}_t)(\mathbb{P})_{\mathcal{B}(\mathbb{R}^d)})(dx) \\
&= \int_{\mathbb{R}^d} \mathbb{E}\big[|U(s, \xi + x)|^2\big] \, ((\mathbf{W}_s)(\mathbb{P})_{\mathcal{B}(\mathbb{R}^d)})(dx) = \mathbb{E}\big[|U(s, \xi + \mathbf{W}_s)|^2\big] .
\end{aligned} \tag{48}
$$

The fact that $V^{(0)}$ is a continuous random field, the fact that $V^{(0)}$ and $\mathcal{R}_0$ are independent, Lemma 2.3, the fact that $\mathcal{R}_0$ is uniformly distributed on $[0,T]$, and (7) hence establish that

$$
\begin{aligned}
&\big\|[0,T] \times \mathbb{R}^d \times \Omega \ni (t,x,\omega) \mapsto (T-t)\big[U(\mathcal{R}_t, x + \mathbb{W}_{\mathcal{R}_t} - \mathbb{W}_t)\big](\omega) \in \mathbb{R}\big\|_0^2 \\
&= \mathbb{E}\big[|TU(\mathcal{R}_0, \xi + \mathbb{W}_{\mathcal{R}_0})|^2\big] = T^2\,\mathbb{E}\big[|V_{\mathcal{R}_0}^{(0)}|^2\big] = \frac{T^2}{T} \int_0^T \mathbb{E}\big[|V_t^{(0)}|^2\big] \, dt \\
&= \frac{T^2}{T} \int_0^T \mathbb{E}\big[|U(t, \xi + \mathbf{W}_t)|^2\big] \, dt = T^2 \|U\|_1^2.
\end{aligned} \tag{49}
$$

In addition, observe that the fact that $(V^{(t)})_{t \in [0,T]}$ and $\mathcal{R}$ are independent, the fact that $V^{(t)}$, $t \in [0,T]$, are continuous random fields, the fact that for all $t \in [0,T]$ it holds that $\mathcal{R}_t$ is uniformely distributed on $[t,T]$, Lemma 2.3, Tonelli's theorem, and (48) demonstrate that for all $l \in \mathbb{N}$ it holds that

$$
\begin{aligned}
&\big\|[0,T]\times\mathbb{R}^d\times\Omega \ni (t,x,\omega)\mapsto (T-t)\,[U(\mathcal{R}_t, x+\mathbb{W}_{\mathcal{R}_t}-\mathbb{W}_t)]\,(\omega)\in\mathbb{R}\big\|_l^2\\
&=\frac{1}{T^l}\int_0^T \tfrac{t^{l-1}}{(l-1)!}\,\mathbb{E}\big[|(T-t)U(\mathcal{R}_t,\xi+\mathbf{W}_t+\mathbb{W}_{\mathcal{R}_t}-\mathbb{W}_t)|^2\big]\,dt\\
&=\frac{1}{T^l}\int_0^T \tfrac{t^{l-1}}{(l-1)!}\,(T-t)^2\,\mathbb{E}\big[\big|V^{(t)}_{\mathcal{R}_t}\big|^2\big]\,dt\\
&=\frac{1}{T^l}\int_0^T \tfrac{t^{l-1}}{(l-1)!}\,(T-t)^2\,\tfrac{1}{(T-t)}\int_t^T \mathbb{E}\big[\big|V^{(t)}_s\big|^2\big]\,ds\,dt\\
&=\frac{1}{T^l}\int_0^T\int_0^T \mathbb{1}_{\{(\mathfrak{t},\mathfrak{s})\in[0,T]^2\colon \mathfrak{t}\le\mathfrak{s}\}}(t,s)\tfrac{t^{l-1}}{(l-1)!}\,(T-t)\,\mathbb{E}\big[|U(s,\xi+\mathbf{W}_s)|^2\big]\,dt\,ds\\
&\le\frac{T}{T^l}\int_0^T\int_0^s \tfrac{t^{l-1}}{(l-1)!}\,dt\,\mathbb{E}\big[|U(s,\xi+\mathbf{W}_s)|^2\big]\,ds\\
&=\frac{T^2}{T^{l+1}}\int_0^T \tfrac{s^l}{l!}\,\mathbb{E}\big[|U(s,\xi+\mathbf{W}_s)|^2\big]\,ds = T^2\|U\|_{l+1}.
\end{aligned}
\tag{50}
$$

Combining this and (49) establishes (46). The proof of Lemma 2.10 is thus completed. ☐

Lemma 2.11. *Assume Setting 2.1, let $k\in\mathbb{N}_0$, let $g\colon\mathbb{R}^d\to\mathbb{R}$ be a $\mathcal{B}(\mathbb{R}^d)/\mathcal{B}(\mathbb{R})$-measurable function, let $v\colon[0,T]\times\mathbb{R}^d\to\mathbb{R}$ be a $\mathcal{B}([0,T]\times\mathbb{R}^d)/\mathcal{B}(\mathbb{R})$-measurable function, let $\mathbb{W}\colon[0,T]\times\Omega\to\mathbb{R}^d$ be a standard Brownian motion with continuous sample paths, and assume that $\mathbb{W}$ and $\mathbf{W}$ are independent. Then it holds that*

(i) $\big\|[0,T]\times\mathbb{R}^d\times\Omega\ni(t,x,\omega)\mapsto g(x+\mathbb{W}_T(\omega)-\mathbb{W}_t(\omega))\in\mathbb{R}\big\|_k^2=\frac{1}{k!}\mathbb{E}\big[|g(\xi+\mathbf{W}_T)|^2\big]$ *and*

(ii) $\|v\|_k\le\frac{1}{\sqrt{k!}}\Big(\sup_{t\in[0,T]}\big(\mathbb{E}\big[|v(t,\xi+\mathbf{W}_t)|^2\big]\big)^{1/2}\Big)$.

Proof of Lemma 2.11. First, observe that (7) and the fact that $\mathbb{W}_T-\mathbb{W}_0=\mathbb{W}_T$ and $\mathbf{W}_T$ are identically distributed ensure that

$$
\begin{aligned}
&\big\|[0,T]\times\mathbb{R}^d\times\Omega\ni(t,x,\omega)\mapsto g(x+\mathbb{W}_T(\omega)-\mathbb{W}_t(\omega))\in\mathbb{R}\big\|_0^2\\
&=\mathbb{E}\big[|g(\xi+\mathbb{W}_T-\mathbb{W}_0)|^2\big]=\mathbb{E}\big[|g(\xi+\mathbf{W}_T)|^2\big].
\end{aligned}
\tag{51}
$$

Next note that the fact that $\mathbf{W}$ and $\mathbb{W}$ are independent standard Brownian motions assures that for all $t\in[0,T]$ the random variables $\mathbf{W}_T=\mathbf{W}_t+\mathbf{W}_T-\mathbf{W}_t$ and $\mathbf{W}_t+\mathbb{W}_T-\mathbb{W}_t$ are identically distributed. The definition of the semi-norm in (7) therefore shows that for all $l\in\mathbb{N}$ it holds that

$$
\begin{aligned}
&\big\|[0,T]\times\mathbb{R}^d\times\Omega\ni(t,x,\omega)\mapsto g(x+\mathbb{W}_T(\omega)-\mathbb{W}_t(\omega))\in\mathbb{R}\big\|_l^2\\
&=\frac{1}{T^l}\int_0^T\tfrac{t^{l-1}}{(l-1)!}\,\mathbb{E}\big[|g(\xi+\mathbf{W}_t+\mathbb{W}_T-\mathbb{W}_t)|^2\big]\,dt\\
&=\left[\frac{1}{T^l}\int_0^T\tfrac{t^{l-1}}{(l-1)!}\,dt\right]\mathbb{E}\big[|g(\xi+\mathbf{W}_T)|^2\big]\\
&=\left[\frac{T^l}{T^l l!}\right]\mathbb{E}\big[|g(\xi+\mathbf{W}_T)|^2\big]=\frac{\mathbb{E}\big[|g(\xi+\mathbf{W}_T)|^2\big]}{l!}.
\end{aligned}
\tag{52}
$$

Combining this and (51) proves Item (i). Next note that (7) implies that

$$
\|v\|_0^2=\mathbb{E}\big[|v(0,\xi)|^2\big]=\mathbb{E}\big[|v(0,\xi+\mathbf{W}_0)|^2\big]\le\sup_{t\in[0,T]}\mathbb{E}\big[|v(t,\xi+\mathbf{W}_t)|^2\big]. \tag{53}
$$

Furthermore, observe that (7) ensures that for all $l \in \mathbb{N}$ it holds that

$$\begin{aligned} \|v\|_l^2 &= \frac{1}{T^l} \int_0^T \frac{t^{l-1}}{(l-1)!} \, \mathbb{E}\big[|v(t,\xi+\mathbf{W}_t)|^2\big] \, dt \\ &\leq \left[\frac{1}{T^l} \int_0^T \frac{t^{l-1}}{(l-1)!} \, dt\right] \sup_{t\in[0,T]} \mathbb{E}\big[|v(t,\xi+\mathbf{W}_t)|^2\big] \\ &= \left[\frac{T^l}{T^l l!}\right] \sup_{t\in[0,T]} \mathbb{E}\big[|v(t,\xi+\mathbf{W}_t)|^2\big] = \frac{1}{l!}\left(\sup_{t\in[0,T]} \mathbb{E}\big[|v(t,\xi+\mathbf{W}_t)|^2\big]\right). \end{aligned} \tag{54}$$

This and (53) establish Item (ii). The proof of Lemma 2.11 is thus completed. □

3 Convergence rates for multilevel Picard approximations for semilinear heat equations

In this section we establish positive convergence rates for certain multilevel Picard approximations in the case where the nonlinearity is independent of the gradient of the solution and satisfies the Lipschitz condition (6).

3.1 Setting

Setting 3.1. *Assume Setting 2.1, let $g \in C(\mathbb{R}^d,\mathbb{R})$, $u \in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ satisfy for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that*

$$\begin{aligned} &\mathbb{E}\big[|g(x+\mathbf{W}_t)|\big] + \int_0^T \big(\mathbb{E}\big[|u(s,\xi+\mathbf{W}_s)|^2\big]\big)^{1/2} \, ds \\ &\quad + \int_t^T \mathbb{E}\big[\,|(F(u))(s,x+\mathbf{W}_{s-t})| + |(F(0))(s,x+\mathbf{W}_{s-t})|\,\big] \, ds < \infty \end{aligned} \tag{55}$$

and
$$u(t,x) = \mathbb{E}\left[g(x+\mathbf{W}_{T-t}) + \int_t^T (F(u))(s,x+\mathbf{W}_{s-t}) \, ds\right], \tag{56}$$

let $\Theta = \cup_{n\in\mathbb{N}}\mathbb{Z}^n$, let $W^\theta \colon [0,T]\times\Omega \to \mathbb{R}^d$, $\theta \in \Theta$, be independent standard Brownian motions with continuous sample paths, let $\mathfrak{r}^\theta \colon \Omega \to [0,1]$, $\theta \in \Theta$, be independent $\mathcal{U}_{[0,1]}$-distributed random variables, assume that $(W^\theta)_{\theta\in\Theta}$, $(\mathfrak{r}^\theta)_{\theta\in\Theta}$, and $\mathbf{W}$ are independent, let $\mathcal{R}^\theta \colon [0,T]\times\Omega \to [0,T]$, $\theta \in \Theta$, satisfy for all $t \in [0,T]$, $\theta \in \Theta$ that $\mathcal{R}^\theta_t = t + (T-t)\mathfrak{r}^\theta$, and let $U^\theta_{n,M} \colon [0,T]\times\mathbb{R}^d\times\Omega \to \mathbb{R}$, $n, M \in \mathbb{Z}$, $\theta \in \Theta$, satisfy for all $n, M \in \mathbb{N}$, $\theta \in \Theta$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that $U^\theta_{-1,M}(t,x) = U^\theta_{0,M}(t,x) = 0$ and

$$\begin{aligned} U^\theta_{n,M}(t,x) &= \frac{1}{M^n}\left[\sum_{i=1}^{M^n} g\big(x + W^{(\theta,0,-i)}_T - W^{(\theta,0,-i)}_t\big)\right] \\ &+ \sum_{l=0}^{n-1} \frac{(T-t)}{M^{n-l}}\left[\sum_{i=1}^{M^{n-l}} \Big(F\big(U^{(\theta,l,i)}_{l,M}\big) - \mathbb{1}_{\mathbb{N}}(l)F\big(U^{(\theta,-l,i)}_{l-1,M}\big)\Big)\Big(\mathcal{R}^{(\theta,l,i)}_t, x + W^{(\theta,l,i)}_{\mathcal{R}^{(\theta,l,i)}_t} - W^{(\theta,l,i)}_t\Big)\right]. \end{aligned} \tag{57}$$

3.2 Properties of the approximations

Lemma 3.2. *Assume Setting 3.1. Then*

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

(ii) it holds for all $n \in \mathbb{N}_0$, $M \in \mathbb{N}$, $\theta \in \Theta$ that $\sigma_\Omega(U^\theta_{n,M}) \subseteq \sigma_\Omega((\mathfrak{r}^{(\theta,\vartheta)})_{\vartheta\in\Theta}, (W^{(\theta,\vartheta)})_{\vartheta\in\Theta}))$,

(iii) it holds for all $n \in \mathbb{N}_0$, $M \in \mathbb{N}$, $\theta \in \Theta$ that $U^\theta_{n,M}$, W^θ, and $\mathfrak{r}^\theta$ are independent,

(iv) it holds for all $n, m \in \mathbb{N}_0$, $M \in \mathbb{N}$, $i, j, k, l, \in \mathbb{Z}$, $\theta \in \Theta$ with $(i, j) \neq (k, l)$ that $U^{(\theta,i,j)}_{n,M}$ and $U^{(\theta,k,l)}_{m,M}$ are independent, and

(v) it holds for all $n \in \mathbb{N}_0$, $M \in \mathbb{N}$ that $(U^\theta_{n,M})_{\theta\in\Theta}$ are identically distributed.

Proof of Lemma 3.2. First, observe that the hypothesis that for all $M \in \mathbb{N}$, $\theta \in \Theta$ it holds that $U^\theta_{0,M} = 0$, (57), Item (i) in Lemma 2.9, the fact for all $\theta \in \Theta$ it holds that W^θ and $\mathcal{R}^\theta$ are continuous random fields, the hypothesis that g is continuous, and induction on $\mathbb{N}_0$ establish Item (i). Next note that Item (i) in Lemma 2.9, Beck et al. [2, Lemma 2.4], and Item (i) assure that for all $n \in \mathbb{N}_0$, $M \in \mathbb{N}$, $\theta \in \Theta$ it holds that $F(U^\theta_{n,M})$ is $(\mathcal{B}([0, T] \times \mathbb{R}^d) \otimes \sigma_\Omega(U^\theta_{n,M}))/\mathcal{B}(\mathbb{R})$-measurable. The hypothesis that for all $M \in \mathbb{N}$, $\theta \in \Theta$ it holds that $U^\theta_{0,M} = 0$, (57), the fact that for all $\theta \in \Theta$ it holds that W^θ is $(\mathcal{B}([0, T]) \otimes \sigma_\Omega(W^\theta))/\mathcal{B}(\mathbb{R}^d)$-measurable, the fact that for all $\theta \in \Theta$ it holds that $\mathcal{R}^\theta$ is $(\mathcal{B}([0, T]) \otimes \sigma_\Omega(\mathfrak{r}^\theta))/\mathcal{B}([0, T])$-measurable, and induction on $\mathbb{N}_0$ prove Item (ii). Furthermore, observe that Item (ii) and the fact that for all $\theta \in \Theta$ it holds that $(\mathfrak{r}^{(\theta,\vartheta)})_{\vartheta\in\Theta}, (W^{(\theta,\vartheta)})_{\vartheta\in\Theta}$, W^θ, and $\mathfrak{r}^\theta$ are independent establish Item (iii). In addition, note that Item (ii) and the fact that for all $i, j, k, l, \in \mathbb{Z}$, $\theta \in \Theta$ with $(i, j) \neq (k, l)$ it holds that $((\mathfrak{r}^{(\theta,i,j,\vartheta)})_{\vartheta\in\Theta}, (W^{(\theta,i,j,\vartheta)})_{\vartheta\in\Theta})$ and $((\mathfrak{r}^{(\theta,k,l,\vartheta)})_{\vartheta\in\Theta}, (W^{(\theta,k,l,\vartheta)})_{\vartheta\in\Theta})$ are independent prove Item (iv). Finally, observe that the hypothesis that for all $M \in \mathbb{N}$, $\theta \in \Theta$ it holds that $U^\theta_{0,M} = 0$, the hypothesis that $(W^\theta)_{\theta\in\Theta}$ are i.i.d., the hypothesis that $(\mathcal{R}^\theta)_{\theta\in\Theta}$ are i.i.d., Items (i)–(iv), Corollary 2.5, and induction on $\mathbb{N}_0$ establish Item (v). The proof of Lemma 3.2 is thus completed. ☐

Lemma 3.3 (Approximations are integrable)**.** *Assume Setting 3.1. Then it holds for all $n \in \mathbb{N}_0$, $M \in \mathbb{N}$, $\theta \in \Theta$, $t \in [0, T]$, $s \in [t, T]$, $x \in \mathbb{R}^d$ that*

$$\mathbb{E}\bigg[\big|U^\theta_{n,M}(s, x + W^\theta_{s-t})\big| + \int_t^T \big|U^\theta_{n,M}(r, x + W^\theta_{r-t})\big| + \big|(F(U^\theta_{n,M}))(r, x + W^\theta_{r-t})\big|\, dr\bigg] < \infty. \tag{58}$$

Proof of Lemma 3.3. Throughout this proof let $M \in \mathbb{N}$, $\theta \in \Theta$, $x \in \mathbb{R}^d$. We claim that for all $n \in \mathbb{N}_0$, $t \in [0, T]$, $s \in [t, T]$ it holds that

$$\mathbb{E}\bigg[\big|U^\theta_{n,M}(s, x + W^\theta_{s-t})\big| + \int_t^T \big|U^\theta_{n,M}(r, x + W^\theta_{r-t})\big| + \big|(F(U^\theta_{n,M}))(r, x + W^\theta_{r-t})\big|\, dr\bigg] < \infty. \tag{59}$$

We now prove (59) by induction on $n \in \mathbb{N}_0$. For the base case $n = 0$, note that (55) and the fact that $U^\theta_{0,M} = 0$ ensure that for all $t \in [0, T]$, $s \in [t, T]$ it holds that

$$\begin{aligned}&\mathbb{E}\bigg[\big|U^\theta_{0,M}(s, x + W^\theta_{s-t})\big| + \int_t^T \big|U^\theta_{0,M}(r, x + W^\theta_{r-t})\big| + \big|(F(U^\theta_{0,M}))(r, x + W^\theta_{r-t})\big|\, dr\bigg]\\ &= \mathbb{E}\bigg[\int_t^T \big|(F(0))(r, x + W^\theta_{r-t})\big|\, dr\bigg] < \infty.\end{aligned} \tag{60}$$

This establishes (59) 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 $k \in \mathbb{N}_0 \cap [0, n)$, $t \in [0, T]$, $s \in [t, T]$ it holds that

$$\mathbb{E}\bigg[\big|U^\theta_{k,M}(s, x + W^\theta_{s-t})\big| + \int_t^T \big|U^\theta_{k,M}(r, x + W^\theta_{r-t})\big| + \big|(F(U^\theta_{k,M}))(r, x + W^\theta_{r-t})\big|\, dr\bigg] < \infty. \tag{61}$$

Observe that the triangle inequality and (57) ensure that for all $t \in [0,T]$, $s \in [t,T]$ it holds that

$$\begin{aligned}
&\mathbb{E}\big[\big|U^{\theta}_{n,M}(s, x+W^{\theta}_{s-t})\big|\big] \\
&\leq \frac{1}{M^n}\sum_{i=1}^{M^n}\mathbb{E}\Big[\Big|g(x+W^{\theta}_{s-t}+W^{(\theta,0,-i)}_T-W^{(\theta,0,-i)}_s)\Big|\Big] \\
&\quad+\sum_{l=0}^{n-1}\tfrac{(T-s)}{M^{n-l}}\sum_{i=1}^{M^{n-l}}\mathbb{E}\Big[\Big|\Big(F(U^{(\theta,l,i)}_{l,M})-\mathbb{1}_{\mathbb{N}}(l)F(U^{(\theta,-l,i)}_{l-1,M})\Big)(\mathcal{R}^{(\theta,l,i)}_s, x+W^{\theta}_{s-t}+W^{(\theta,l,i)}_{\mathcal{R}^{(\theta,l,i)}_s}-W^{(\theta,l,i)}_s)\Big|\Big].
\end{aligned} \tag{62}$$

In addition, note that the fact that for all $i \in \mathbb{Z}$ it holds that W^{θ} and $W^{(\theta,0,i)}$ are independent Brownian motions assures that for all $t \in [0,T]$, $s \in [t,T]$, $i \in \mathbb{Z}$ it holds that

$$\mathbb{E}\Big[\Big|g(x+W^{\theta}_{s-t}+W^{(\theta,0,i)}_T-W^{(\theta,0,i)}_s)\Big|\Big]=\mathbb{E}\big[\big|g(x+W^{\theta}_{(s-t)+(T-s)})\big|\big]=\mathbb{E}\big[\big|g(x+W^{\theta}_{T-t})\big|\big]. \tag{63}$$

Moreover, note that Lemma 3.2, the hypothesis that $(W^{\theta})_{\theta\in\Theta}$ are i.i.d., the hypothesis that $(\mathcal{R}^{\theta})_{\theta\in\Theta}$ are i.i.d., the hypothesis that $(W^{\theta})_{\theta\in\Theta}$ and $(\mathcal{R}^{\theta})_{\theta\in\Theta}$ are independent, Lemma 2.3, and the triangle inequality assure that for all $t \in [0,T]$, $s \in [t,T]$ it holds that

$$\begin{aligned}
&\sum_{l=0}^{n-1}\tfrac{(T-s)}{M^{n-l}}\sum_{i=1}^{M^{n-l}}\mathbb{E}\Big[\Big|\Big(F(U^{(\theta,l,i)}_{l,M})-\mathbb{1}_{\mathbb{N}}(l)F(U^{(\theta,-l,i)}_{l-1,M})\Big)(\mathcal{R}^{(\theta,l,i)}_s, x+W^{\theta}_{s-t}+W^{(\theta,l,i)}_{\mathcal{R}^{(\theta,l,i)}_s}-W^{(\theta,l,i)}_s)\Big|\Big] \\
&=\sum_{l=0}^{n-1}(T-s)\mathbb{E}\Big[\Big|\Big(F(U^{(\theta,l,0)}_{l,M})-\mathbb{1}_{\mathbb{N}}(l)F(U^{(\theta,-l,0)}_{l-1,M})\Big)(\mathcal{R}^{(\theta,l,0)}_s, x+W^{\theta}_{s-t}+W^{(\theta,l,0)}_{\mathcal{R}^{(\theta,l,0)}_s}-W^{(\theta,l,0)}_s)\Big|\Big] \\
&\leq 2\sum_{l=0}^{n-1}(T-s)\mathbb{E}\Big[\Big|\big(F(U^{(\theta,l,0)}_{l,M})\big)(\mathcal{R}^{(\theta,l,0)}_s, x+W^{\theta}_{s-t}+W^{(\theta,l,0)}_{\mathcal{R}^{(\theta,l,0)}_s}-W^{(\theta,l,0)}_s)\Big|\Big].
\end{aligned} \tag{64}$$

Furthermore, observe that Lemma 3.2, the fact that for all $l \in \mathbb{Z}$ it holds that W^{θ}, $W^{(\theta,l,0)}$, $\mathcal{R}^{(\theta,l,0)}$, and $U^{(\theta,l,0)}$ are independent, and Lemma 2.3 demonstrate that for all $t \in [0,T]$, $s \in [t,T]$, $l \in \mathbb{N}_0 \cap [0,n)$ it holds that

$$\begin{aligned}
&(T-s)\mathbb{E}\Big[\Big|\big(F(U^{(\theta,l,0)}_{l,M})\big)(\mathcal{R}^{(\theta,l,0)}_s, x+W^{\theta}_{s-t}+W^{(\theta,l,0)}_{\mathcal{R}^{(\theta,l,0)}_s}-W^{(\theta,l,0)}_s)\Big|\Big] \\
&=\int_s^T\mathbb{E}\Big[\big|\big(F(U^{(\theta,l,0)}_{l,M})\big)(r, x+W^{\theta}_{s-t}+W^{(\theta,l,0)}_r-W^{(\theta,l,0)}_s)\big|\Big]\,dr \\
&=\int_s^T\mathbb{E}\Big[\big|\big(F(U^{(\theta,l,0)}_{l,M})\big)(r, x+W^{(\theta,l,0)}_{(s-t)+(r-s)})\big|\Big]\,dr=\int_s^T\mathbb{E}\big[\big|\big(F(U^{\theta}_{l,M})\big)(r, x+W^{\theta}_{r-t})\big|\big]\,dr.
\end{aligned} \tag{65}$$

Combining this, (62), (63), and (64) with (55), (61), and Tonelli's theorem establishes that for all $t \in [0,T]$, $s \in [t,T]$ it holds that

$$\begin{aligned}
&\mathbb{E}\big[\big|U^{\theta}_{n,M}(s, x+W^{\theta}_{s-t})\big|\big] \\
&\leq \frac{1}{M^n}\sum_{i=1}^{M^n}\mathbb{E}\big[\big|g(x+W^{\theta}_{T-t})\big|\big]+2\sum_{l=0}^{n-1}\int_t^T\mathbb{E}\big[\big|\big(F(U^{\theta}_{l,M})\big)(r, x+W^{\theta}_{r-t})\big|\big]\,dr \\
&=\mathbb{E}\big[\big|g(x+W^{\theta}_{T-t})\big|\big]+2\sum_{l=0}^{n-1}\mathbb{E}\Big[\int_t^T\big|\big(F(U^{\theta}_{l,M})\big)(r, x+W^{\theta}_{t-r})\big|\,dr\Big]<\infty.
\end{aligned} \tag{66}$$

This, Tonelli's theorem, and (61) imply that for all $t \in [0,T]$ it holds that

$$\begin{aligned} &\mathbb{E}\bigg[\int_t^T \big|U^\theta_{n,M}(s,x+W^\theta_{s-t})\big|\,ds\bigg] = \int_t^T \mathbb{E}\big[\big|U^\theta_{n,M}(s,x+W^\theta_{s-t})\big|\big]\,ds \\ &\leq (T-t)\left[\mathbb{E}\big[\big|g(x+W^\theta_{T-t})\big|\big] + 2\sum_{l=0}^{n-1}\int_t^T \mathbb{E}\big[\big|\big(F(U^\theta_{l,M})\big)(r,x+W^\theta_{r-t})\big|\big]\,dr\right] < \infty. \end{aligned} \tag{67}$$

The triangle inequality, Tonelli's theorem, (6), and (55) hence prove that for all $t \in [0,T]$ it holds that

$$\begin{aligned} &\mathbb{E}\left[\int_t^T \big|(F(U^\theta_{n,M}))(s,x+W^\theta_{s-t})\big|\,ds\right] = \int_t^T \mathbb{E}\Big[\big|(F(U^\theta_{n,M}))(s,x+W^\theta_{s-t})\big|\Big]\,ds \\ &\leq \int_t^T \mathbb{E}\Big[\,\big|\big(F(U^\theta_{n,M})-F(0)\big)(s,x+W^\theta_{s-t})\big|\,\Big]\,ds + \int_t^T \mathbb{E}\Big[\big|(F(0))(s,x+W^\theta_{s-t})\big|\Big]\,ds \\ &\leq \int_t^T \mathbb{E}\Big[L\big|U^\theta_{n,M}(s,x+W^\theta_{s-t})\big|\Big]\,ds + \int_t^T \mathbb{E}\Big[\big|(F(0))(s,x+W^\theta_{s-t})\big|\Big]\,ds < \infty. \end{aligned} \tag{68}$$

Induction, (66), and (67) hence establish (59). The proof of Lemma 3.3 is thus completed. ☐

3.3 Upper bound for the exact solution

In this subsection we establish the upper bound (69) below for the exact solution which is well-known in the literature and included here for the reason of being self-contained.

Lemma 3.4 (Upper bound for exact solution)**.** *Assume Setting 3.1. Then it holds that*

$$\sup_{t\in[0,T]} \big(\mathbb{E}\big[|u(t,\xi+\mathbf{W}_t)|^2\big]\big)^{1/2} \leq e^{LT}\Big[\big(\mathbb{E}\big[|g(\xi+\mathbf{W}_T)|^2\big]\big)^{1/2} + T\|F(0)\|_1\Big]. \tag{69}$$

Proof of Lemma 3.4. Throughout this proof let $\mathbb{W}\colon [0,T]\times\Omega \to \mathbb{R}^d$ be a standard Brownian motion with continuous sample paths, assume that $\mathbf{W}$ and $\mathbb{W}$ are independent, 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

$$\mu_t(B) = \mathbb{P}(\xi + \mathbb{W}_t \in B), \tag{70}$$

and assume w.lo.g. that $\mathbb{E}\big[|g(\xi+\mathbf{W}_T)|^2\big] + \|F(0)\|_1 < \infty$. Observe that the integral transformation theorem, (56), and the triangle inequality assure that for all $t \in [0,T]$ it holds that

$$\begin{aligned} &\big(\mathbb{E}\big[|u(t,\xi+\mathbf{W}_t)|^2\big]\big)^{1/2} = \big(\mathbb{E}\big[|u(t,\xi+\mathbb{W}_t)|^2\big]\big)^{1/2} = \left(\int_{\mathbb{R}^d} |u(t,x)|^2\,\mu_t(dx)\right)^{1/2} \\ &= \left(\int_{\mathbb{R}^d}\left|\mathbb{E}\left[g(x+\mathbf{W}_{T-t}) + \int_t^T (F(u))(s,x+\mathbf{W}_{s-t})\,ds\right]\right|^2 \mu_t(dx)\right)^{1/2} \\ &\leq \left(\int_{\mathbb{R}^d}\big|\mathbb{E}\big[g(x+\mathbf{W}_{T-t})\big]\big|^2\,\mu_t(dx)\right)^{1/2} + \left(\int_{\mathbb{R}^d}\left|\mathbb{E}\left[\int_t^T (F(u))(s,x+\mathbf{W}_{s-t})\,ds\right]\right|^2 \mu_t(dx)\right)^{1/2}. \end{aligned} \tag{71}$$

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

$$\begin{aligned} \big(\mathbb{E}\big[|u(t,\xi+\mathbf{W}_t)|^2\big]\big)^{1/2} &\leq \left(\int_{\mathbb{R}^d}\mathbb{E}\big[|g(x+\mathbf{W}_{T-t})|^2\big]\,\mu_t(dx)\right)^{1/2} \\ &\quad + \left(\int_{\mathbb{R}^d}\mathbb{E}\left[\left(\int_t^T \big|(F(u))(s,x+\mathbf{W}_{s-t})\big|\,ds\right)^2\right]\mu_t(dx)\right)^{1/2}. \end{aligned} \tag{72}$$

Furthermore, observe that Lemma 2.3, (70), and the fact that $\mathbf{W}$ and $\mathbb{W}$ are independent Brownian motions demonstrate that for all $t \in [0,T]$ it holds that

$$\begin{aligned}\left(\int_{\mathbb{R}^d} \mathbb{E}\big[|g(x+\mathbf{W}_{T-t})|^2\big]\ \mu_t(dx)\right)^{1/2} &= \big(\mathbb{E}\big[|g(\xi+\mathbb{W}_t+\mathbf{W}_{T-t})|^2\big]\big)^{1/2} \\ &= \big(\mathbb{E}\big[|g(\xi+\mathbf{W}_T)|^2\big]\big)^{1/2}.\end{aligned} \tag{73}$$

In addition, note that Minkowski's integral inequality, Lemma 2.3, (70), and the fact that $\mathbf{W}$ and $\mathbb{W}$ are independent Brownian motions imply that for all $t \in [0,T]$ it holds that

$$\begin{aligned}&\left(\int_{\mathbb{R}^d} \mathbb{E}\left[\left(\int_t^T \big|(F(u))(s,x+\mathbf{W}_{s-t})\big|\,ds\right)^2\right]\ \mu_t(dx)\right)^{1/2} \\ &\le \int_t^T \left(\int_{\mathbb{R}^d} \mathbb{E}\big[|(F(u))(s,x+\mathbf{W}_{s-t})|^2\big]\ \mu_t(dx)\right)^{1/2} ds \\ &= \int_t^T \big(\mathbb{E}\big[|(F(u))(s,\xi+\mathbb{W}_t+\mathbf{W}_{s-t})|^2\big]\big)^{1/2} ds = \int_t^T \big(\mathbb{E}\big[|(F(u))(s,\xi+\mathbf{W}_s)|^2\big]\big)^{1/2} ds.\end{aligned} \tag{74}$$

This, the triangle inequality, and (6) assure that for all $t \in [0,T]$ it holds that

$$\begin{aligned}&\left(\int_{\mathbb{R}^d} \mathbb{E}\left[\left(\int_t^T \big|(F(u))(s,x+\mathbf{W}_{s-t})\big|\,ds\right)^2\right]\ \mu_t(dx)\right)^{1/2} \\ &\le \int_t^T \big(\mathbb{E}\big[|(F(0))(s,\xi+\mathbf{W}_s)|^2\big]\big)^{1/2}\ ds + \int_t^T \big(\mathbb{E}\big[|(F(u)-F(0))(s,\xi+\mathbf{W}_s)|^2\big]\big)^{1/2}\ ds \\ &\le \int_t^T \big(\mathbb{E}\big[|(F(0))(s,\xi+\mathbf{W}_s)|^2\big]\big)^{1/2}\ ds + \int_t^T \big(\mathbb{E}\big[L^2|u(s,\xi+\mathbf{W}_s)|^2\big]\big)^{1/2}\ ds.\end{aligned} \tag{75}$$

Furthermore, note that Jensen's inequality and (7) ensure that for all $t \in [0,T)$ it holds that

$$\begin{aligned}&\int_t^T \big(\mathbb{E}\big[|(F(0))(s,\xi+\mathbf{W}_s)|^2\big]\big)^{1/2}\ ds \\ &= (T-t)\left(\tfrac{1}{(T-t)}\int_t^T \big(\mathbb{E}\big[|(F(0))(s,\xi+\mathbf{W}_s)|^2\big]\big)^{1/2}\ ds\right) \\ &\le (T-t)\left(\tfrac{1}{(T-t)}\int_t^T \mathbb{E}\big[|(F(0))(s,\xi+\mathbf{W}_s)|^2\big]\ ds\right)^{1/2} \\ &\le \sqrt{T}\left(\int_0^T \mathbb{E}\big[|(F(0))(s,\xi+\mathbf{W}_s)|^2\big]\ ds\right)^{1/2} = T\|F(0)\|_1.\end{aligned} \tag{76}$$

Combining this with (72), (73), and (75) implies that for all $t \in [0,T]$ it holds that

$$\begin{aligned}&\big(\mathbb{E}\big[|u(t,\xi+\mathbf{W}_t)|^2\big]\big)^{1/2} \\ &\le \big(\mathbb{E}\big[|g(\xi+\mathbf{W}_T)|^2\big]\big)^{1/2} + T\|F(0)\|_1 + L\int_t^T \big(\mathbb{E}\big[|u(s,\xi+\mathbf{W}_s)|^2\big]\big)^{1/2}\ ds.\end{aligned} \tag{77}$$

The hypothesis that $\int_0^T (\mathbb{E}[|u(t,\xi+\mathbf{W}_t)|^2])^{1/2}\,dt < \infty$ and Gronwall's integral inequality hence establish that for all $t \in [0,T]$ it holds that

$$\begin{aligned}\big(\mathbb{E}\big[|u(t,\xi+\mathbf{W}_t)|^2\big]\big)^{1/2} &\le e^{L(T-t)}\left[\big(\mathbb{E}\big[|g(\xi+\mathbf{W}_T)|^2\big]\big)^{1/2} + T\|F(0)\|_1\right] \\ &\le e^{LT}\left[\big(\mathbb{E}\big[|g(\xi+\mathbf{W}_T)|^2\big]\big)^{1/2} + T\|F(0)\|_1\right].\end{aligned} \tag{78}$$

The proof of Lemma 3.4 is thus completed. □

3.4 Error analysis for multilevel Picard approximations

Theorem 3.5. *Assume Setting 3.1 and let $N, M \in \mathbb{N}$. Then it holds that*

$$\Big(\mathbb{E}\Big[\big|U^0_{N,M}(0,\xi) - u(0,\xi)\big|^2\Big]\Big)^{1/2} \leq e^{LT}\Big[\big(\mathbb{E}\big[|g(\xi + \mathbf{W}_T)|^2\big]\big)^{1/2} + T\,\|F(0)\|_1\Big]\frac{e^{M/2}(1+2LT)^N}{M^{N/2}}. \tag{79}$$

Proof of Theorem 3.5. Throughout this proof assume w.l.o.g. that $\mathbb{E}\big[|g(\xi+\mathbf{W}_T)|^2\big] + \|F(0)\|_1 < \infty$. Note that Item (i) in Lemma 2.6 and Lemma 3.3 assure that for all $n \in \mathbb{N}$, $k \in \mathbb{N}_0$ it holds that

$$\big\|U^0_{n,M} - u\big\|_k \leq \big\|U^0_{n,M} - \mathbb{E}\big[U^0_{n,M}\big]\big\|_k + \big\|\mathbb{E}\big[U^0_{n,M}\big] - u\big\|_k. \tag{80}$$

Next observe that Lemma 2.3, Item (i) in Lemma 2.9, Lemma 3.2, Lemma 3.3, Corollary 2.5, and the fact that for all $\theta \in \Theta$, $t \in [0,T]$ it holds that $\mathcal{R}^\theta_t$ is uniformly distributed on $[t,T]$ assure that for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $n \in \mathbb{N}$, $i,j,k \in \mathbb{Z}$, $\theta \in \Theta$ it holds that

$$\begin{aligned}
&(T-t)\mathbb{E}\Big[\big|(F(U^{(\theta,k,i)}_{n,M}))(\mathcal{R}^{(\theta,j,i)}_t, x + W^{(\theta,j,i)}_{\mathcal{R}^{(\theta,j,i)}_t} - W^{(\theta,j,i)}_t\big|\Big] \\
&= \int_t^T \mathbb{E}\Big[\big|(F(U^{(\theta,k,i)}_{n,M}))(s, x + W^{(\theta,j,i)}_s - W^{(\theta,j,i)}_t)\big|\Big]\,ds \\
&= \int_t^T \mathbb{E}\big[\big|(F(U^{\theta}_{n,M}))(s, x + W^{\theta}_{s-t})\big|\big]\,ds < \infty.
\end{aligned} \tag{81}$$

Combining this with the fact that for all $t \in [0,T]$, $\theta \in \Theta$ it holds that $\mathbb{E}\big[\big|g(x + W^\theta_T - W^\theta_t)\big|\big] = \mathbb{E}\big[\big|g(x + W^\theta_{T-t})\big|\big] < \infty$ and (57) ensures that for all $n \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
\mathbb{E}\big[U^0_{n,M}(t,x)\big] &= \frac{1}{M^n}\left[\sum_{i=1}^{M^n}\mathbb{E}\Big[g\big(x + W^{(0,0,-i)}_T - W^{(0,0,-i)}_t\big)\Big]\right] \\
&+ \sum_{l=0}^{n-1}\frac{(T-t)}{M^{n-l}}\left[\sum_{i=1}^{M^{n-l}}\mathbb{E}\Big[\Big(F\big(U^{(0,l,i)}_{l,M}\big) - \mathbb{1}_{\mathbb{N}}(l)F\big(U^{(0,-l,i)}_{l-1,M}\big)\Big)\Big(\mathcal{R}^{(0,l,i)}_t, x + W^{(0,l,i)}_{\mathcal{R}^{(0,l,i)}_t} - W^{(0,l,i)}_t\Big)\Big]\right].
\end{aligned} \tag{82}$$

This and (57) imply that for all $n \in \mathbb{N}$, $k \in \mathbb{N}_0$ it holds that

$$\begin{aligned}
&\big\|U^0_{n,M} - \mathbb{E}\big[U^0_{n,M}\big]\big\|_k \\
&= \big\|[0,T]\times\mathbb{R}^d\times\Omega \ni (t,x,\omega) \mapsto U^0_{n,M}(t,x,\omega) - \mathbb{E}\big[U^0_{n,M}(t,x)\big]\big\|_k \\
&\leq \Bigg\|[0,T]\times\mathbb{R}^d\times\Omega \ni (t,x,\omega) \\
&\qquad\mapsto \left[\sum_{i=1}^{M^n}\tfrac{1}{M^n}\Big(\Big[g\big(x + W^{(0,0,-i)}_T - W^{(0,0,-i)}_t\big)\Big](\omega) - \mathbb{E}\Big[g\big(x + W^{(0,0,-i)}_T - W^{(0,0,-i)}_t\big)\Big]\Big)\right] \in \mathbb{R}\Bigg\|_k \\
&\quad+ \sum_{l=0}^{n-1}\Bigg\|[0,T]\times\mathbb{R}^d\times\Omega \ni (t,x,\omega) \\
&\qquad\mapsto \left[\sum_{i=1}^{M^{n-l}}\tfrac{T-t}{M^{n-l}}\Bigg(\Big[\big(F(U^{(0,l,i)}_{l,M}) - \mathbb{1}_{\mathbb{N}}(l)F(U^{(0,-l,i)}_{l-1,M})\big)\big(\mathcal{R}^{(0,l,i)}_t, x + W^{(0,l,i)}_{\mathcal{R}^{(0,l,i)}_t} - W^{(0,l,i)}_t\big)\Big](\omega)\right. \\
&\qquad\left. - \mathbb{E}\Big[\big(F(U^{(0,l,i)}_{l,M}) - \mathbb{1}_{\mathbb{N}}(l)F(U^{(0,-l,i)}_{l-1,M})\big)\big(\mathcal{R}^{(0,l,i)}_t, x + W^{(0,l,i)}_{\mathcal{R}^{(0,l,i)}_t} - W^{(0,l,i)}_t\big)\Big]\Bigg)\right] \in \mathbb{R}\Bigg\|_k.
\end{aligned} \tag{83}$$

Moreover, note that Lemma 3.2, the hypothesis that $(W^\theta)_{\theta\in\Theta}$ are i.i.d., the hypothesis that $(\mathcal{R}^\theta)_{\theta\in\Theta}$ are i.i.d., Item (i) in Lemma 2.6, and Corollary 2.5 ensure that for all $l \in \mathbb{N}_0$ it holds that

$$\begin{aligned}\Big([0,T]\times\mathbb{R}^d\times\Omega \ni (t,x,\omega)\\ \mapsto \Big[\big(F(U^{(0,l,i)}_{l,M}) - \mathbb{1}_{\mathbb{N}}(l)F(U^{(0,-l,i)}_{l-1,M})\big)\big(\mathcal{R}^{(0,l,i)}_t, x + W^{(0,l,i)}_{\mathcal{R}^{(0,l,i)}_t} - W^{(0,l,i)}_t\big)\Big](\omega)\in\mathbb{R}\Big)_{i\in\mathbb{Z}}\end{aligned} \tag{84}$$

are continuous i.i.d. random fields. Lemma 3.2, the hypothesis that $(W^\theta)_{\theta\in\Theta}$ are i.i.d., (83), and Lemma 2.8 therefore show that for all $n \in \mathbb{N}$, $k \in \mathbb{N}_0$ it holds that

$$\begin{aligned}&\big\|U^0_{n,M} - \mathbb{E}\big[U^0_{n,M}\big]\big\|_k\\ &\leq \Bigg[\sum_{i=1}^{M^n}\big|\tfrac{1}{M^n}\big|^2\Bigg]^{1/2}\Big\|[0,T]\times\mathbb{R}^d\times\Omega\ni(t,x,\omega)\mapsto\Big[g(x+W^{(0,0,-1)}_T - W^{(0,0,-1)}_t)\Big](\omega)\in\mathbb{R}\Big\|_k\\ &\quad+\sum_{l=0}^{n-1}\Big[\textstyle\sum_{i=1}^{M^{n-l}}\big|\tfrac{1}{M^{n-l}}\big|^2\Big]^{1/2}\Big\|[0,T]\times\mathbb{R}^d\times\Omega\ni(t,x,\omega)\\ &\qquad\mapsto(T-t)\Big[\big(F(U^{(0,l,1)}_{l,M}) - \mathbb{1}_{\mathbb{N}}(l)F(U^{(0,-l,1)}_{l-1,M})\big)\big(\mathcal{R}^{(0,l,1)}_t, x+W^{(0,l,1)}_{\mathcal{R}^{(0,l,1)}_t} - W^{(0,l,1)}_t\big)\Big](\omega)\in\mathbb{R}\Big\|_k\\ &=\tfrac{1}{\sqrt{M^n}}\Big\|[0,T]\times\mathbb{R}^d\times\Omega\ni(t,x,\omega)\mapsto\Big[g(x+W^{(0,0,-1)}_T - W^{(0,0,-1)}_t)\Big](\omega)\in\mathbb{R}\Big\|_k\\ &\quad+\sum_{l=0}^{n-1}\tfrac{1}{\sqrt{M^{(n-l)}}}\Big\|[0,T]\times\mathbb{R}^d\times\Omega\ni(t,x,\omega)\\ &\qquad\mapsto(T-t)\Big[\big(F(U^{(0,l,1)}_{l,M}) - \mathbb{1}_{\mathbb{N}}(l)F(U^{(0,-l,1)}_{l-1,M})\big)\big(\mathcal{R}^{(0,l,1)}_t, x+W^{(0,l,1)}_{\mathcal{R}^{(0,l,1)}_t} - W^{(0,l,1)}_t\big)\Big](\omega)\in\mathbb{R}\Big\|_k.\end{aligned} \tag{85}$$

Moreover, observe that Item (i) in Lemma 2.11 and the hypothesis that $(W^\theta)_{\theta\in\Theta}$ and $\mathbf{W}$ are independent assure that for all $k \in \mathbb{N}_0$ it holds that

$$\begin{aligned}&\Big\|[0,T]\times\mathbb{R}^d\times\Omega\ni(t,x,\omega)\mapsto\Big[g(x+W^{(0,0,-1)}_T - W^{(0,0,-1)}_t)\Big](\omega)\in\mathbb{R}\Big\|_k\\ &=\tfrac{1}{\sqrt{k!}}\big(\mathbb{E}\big[|g(\xi+\mathbf{W}_T)|^2\big]\big)^{1/2}.\end{aligned} \tag{86}$$

Furthermore, note that the hypothesis that $(W^\theta)_{\theta\in\Theta}$ are i.i.d., the hypothesis that $(\mathcal{R}^\theta)_{\theta\in\Theta}$ are i.i.d., the hypothesis that $(W^\theta)_{\theta\in\Theta}$, $(\mathcal{R}^\theta)_{\theta\in\Theta}$, and $\mathbf{W}$ are independent, Lemma 3.2, and Lemma 2.10 imply that for all $n \in \mathbb{N}$, $k \in \mathbb{N}_0$ it holds that

$$\begin{aligned}&\sum_{l=0}^{n-1}\tfrac{1}{\sqrt{M^{(n-l)}}}\Big\|[0,T]\times\mathbb{R}^d\times\Omega\ni(t,x,\omega)\\ &\quad\mapsto(T-t)\Big[\big(F(U^{(0,l,1)}_{l,M}) - \mathbb{1}_{\mathbb{N}}(l)F(U^{(0,-l,1)}_{l-1,M})\big)\big(\mathcal{R}^{(0,l,1)}_t, x+W^{(0,l,1)}_{\mathcal{R}^{(0,l,1)}_t} - W^{(0,l,1)}_t\big)\Big](\omega)\in\mathbb{R}\Big\|_k\\ &\leq\sum_{l=0}^{n-1}\tfrac{T}{\sqrt{M^{(n-l)}}}\Big\|F(U^{(0,l,1)}_{l,M}) - \mathbb{1}_{\mathbb{N}}(l)F(U^{(0,-l,1)}_{l-1,M})\Big\|_{k+1}\end{aligned} \tag{87}$$

Item (i) in Lemma 2.6, the hypothesis that $U^0_{0,M} = 0$, and Lemma 2.9 therefore demonstrate that for all $n \in \mathbb{N}$, $k \in \mathbb{N}_0$ it holds that

$$\begin{aligned}
&\sum_{l=0}^{n-1} \tfrac{1}{\sqrt{M^{(n-l)}}} \bigg\| [0,T] \times \mathbb{R}^d \times \Omega \ni (t,x,\omega) \\
&\quad \mapsto (T-t) \Big[\big(F(U^{(0,l,1)}_{l,M}) - \mathbb{1}_{\mathbb{N}}(l) F(U^{(0,-l,1)}_{l-1,M})\big)(\mathcal{R}^{(0,l,1)}_t, x + W^{(0,l,1)}_{\mathcal{R}^{(0,l,1)}_t} - W^{(0,l,1)}_t) \Big](\omega) \in \mathbb{R} \bigg\|_k \\
&\leq \tfrac{T}{\sqrt{M^n}} \big\| F(U^0_{0,M}) \big\|_{k+1} + \sum_{l=1}^{n-1} \tfrac{T}{\sqrt{M^{(n-l)}}} \Big(\big\| F(U^0_{l,M}) - F(u) \big\|_{k+1} + \big\| F(u) - F(U^0_{l-1,M}) \big\|_{k+1} \Big) \\
&\leq \tfrac{T}{\sqrt{M^n}} \| F(0) \|_{k+1} + \left[\sum_{l=1}^{n-1} \tfrac{TL}{\sqrt{M^{(n-l)}}} \big\| U^0_{l,M} - u \big\|_{k+1} \right] + \left[\sum_{l=1}^{n-1} \tfrac{TL}{\sqrt{M^{(n-l)}}} \big\| U^0_{l-1,M} - u \big\|_{k+1} \right] \\
&\leq \tfrac{T}{\sqrt{M^n}} \| F(0) \|_{k+1} + \sum_{l=0}^{n-1} \tfrac{(2 - \mathbb{1}_{\{n-1\}}(l)) LT}{\sqrt{M^{(n-l-1)}}} \big\| U^0_{l,M} - u \big\|_{k+1}.
\end{aligned} \tag{88}$$

In addition, observe that (7) ensures that for all $k \in \mathbb{N}_0$ it holds that

$$\begin{aligned}
\|F(0)\|^2_{k+1} &= \tfrac{1}{T^{k+1}} \int_0^T \tfrac{t^k}{k!} \mathbb{E}\big[|F(0)(t, \xi + \mathbf{W}_t)|^2 \big] \, dt \\
&\leq \tfrac{T^k}{T^{k+1} k!} \int_0^T \mathbb{E}\big[|F(0)(t, \xi + \mathbf{W}_t)|^2 \big] \, dt = \tfrac{1}{k!} \|F(0)\|^2_1.
\end{aligned} \tag{89}$$

Combining this (85), (86), and (88) establishes that for all $n \in \mathbb{N}$, $k \in \mathbb{N}_0$ it holds that

$$\begin{aligned}
&\big\| U^0_{n,M} - \mathbb{E}[U^0_{n,M}] \big\|_k \\
&\leq \tfrac{1}{\sqrt{k! M^n}} \big(\mathbb{E}\big[|g(\xi + \mathbf{W}_T)|^2 \big] \big)^{1/2} + \tfrac{T}{\sqrt{k! M^n}} \|F(0)\|_1 + \sum_{l=0}^{n-1} \tfrac{(2 - \mathbb{1}_{\{n-1\}}(l)) LT}{\sqrt{M^{(n-l-1)}}} \big\| U^0_{l,M} - u \big\|_{k+1} \\
&= \tfrac{1}{\sqrt{k! M^n}} \Big[\big(\mathbb{E}\big[|g(\xi + \mathbf{W}_T)|^2 \big] \big)^{1/2} + T \|F(0)\|_1 \Big] + \sum_{l=0}^{n-1} \tfrac{(2 - \mathbb{1}_{\{n-1\}}(l)) LT}{\sqrt{M^{(n-l-1)}}} \big\| U^0_{l,M} - u \big\|_{k+1}.
\end{aligned} \tag{90}$$

Next observe that, (81), (82), and (84) demonstrate that for all $n \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
&\mathbb{E}\big[U^0_{n,M}(t,x) \big] \\
&= \frac{1}{M^n} \left[\sum_{i=1}^{M^n} \mathbb{E}\big[g\big(x + W^0_T - W^0_t\big) \big] \right] \\
&\quad + \sum_{l=0}^{n-1} \frac{(T-t)}{M^{n-l}} \left[\sum_{i=1}^{M^{n-l}} \mathbb{E}\Big[\big(F(U^0_{l,M}) - \mathbb{1}_{\mathbb{N}}(l) F(U^0_{l-1,M}) \big)(\mathcal{R}^0_t, x + W^0_{\mathcal{R}^0_t} - W^0_t) \Big] \right] \\
&= \mathbb{E}\big[g(x + W^0_T - W^0_t) \big] + (T-t) \Bigg(\sum_{l=0}^{n-1} \mathbb{E}\Big[F(U^0_{l,M})(\mathcal{R}^0_t, x + W^0_{\mathcal{R}^0_t} - W^0_t) \Big] \\
&\qquad - \mathbb{1}_{\mathbb{N}}(l) \, \mathbb{E}\Big[F(U^0_{l-1,M})(\mathcal{R}^0_t, x + W^0_{\mathcal{R}^0_t} - W^0_t) \Big] \Bigg) \\
&= \mathbb{E}\big[g(x + W^0_{T-t}) \big] + (T-t) \mathbb{E}\Big[\big(F(U^0_{n-1,M}) \big)(\mathcal{R}^0_t, x + W^0_{\mathcal{R}^0_t} - W^0_t) \Big].
\end{aligned} \tag{91}$$

In addition, note that (55), (56), Fubini's theorem, and Lemma 2.4 assure that for all $t \in [0,T]$,

$x \in \mathbb{R}^d$ it holds that

$$\begin{aligned} u(t,x) &= \mathbb{E}[g(x+\mathbf{W}_{T-t})] + \int_t^T \mathbb{E}[(F(u))(s, x+\mathbf{W}_s - \mathbf{W}_t)]\, ds \\ &= \mathbb{E}[g(x+\mathbf{W}_{T-t})] + (T-t)\mathbb{E}\big[(F(u))(\mathcal{R}_t^0, x+\mathbf{W}_{\mathcal{R}_t^0} - \mathbf{W}_t)\big] \\ &= \mathbb{E}\big[g(x+W_{T-t}^0)\big] + (T-t)\mathbb{E}\Big[(F(u))(\mathcal{R}_t^0, x+W_{\mathcal{R}_t^0}^0 - W_t^0)\Big]. \end{aligned} \tag{92}$$

Combining this with (91) yields that for all $n \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned} &\mathbb{E}\big[U_{n,M}^0(t,x)\big] - u(t,x) \\ &= (T-t)\left(\mathbb{E}\Big[\big(F(U_{n-1,M}^0)\big)(\mathcal{R}_t^0, x+W_{\mathcal{R}_t^0}^0 - W_t^0)\Big] - \mathbb{E}\Big[(F(u))(\mathcal{R}_t^0, x+W_{\mathcal{R}_t^0}^0 - W_t^0)\Big]\right) \\ &= \mathbb{E}\Big[(T-t)\left(F(U_{n-1,M}^0) - F(u)\right)(\mathcal{R}_t^0, x+W_{\mathcal{R}_t^0}^0 - W_t^0)\Big]. \end{aligned} \tag{93}$$

Lemma 2.7, Lemma 2.10, Lemma 2.9, and Lemma 3.2 hence show that for all $n \in \mathbb{N}$, $k \in \mathbb{N}_0$ it holds that

$$\begin{aligned} \big\|\mathbb{E}\big[U_{n,M}^0\big] - u\big\|_k &= \Big\|[0,T]\times\mathbb{R}^d\times\Omega \ni (t,x,\omega) \\ &\qquad \mapsto \mathbb{E}\Big[(T-t)\left(F(U_{n-1,M}^0) - F(u)\right)(\mathcal{R}_t^0, x+W_{\mathcal{R}_t^0}^0 - W_t^0)\Big] \in \mathbb{R}\Big\|_k \\ &\le \Big\|[0,T]\times\mathbb{R}^d\times\Omega \ni (t,x,\omega) \\ &\qquad \mapsto (T-t)\left[\left(F(U_{n-1,M}^0) - F(u)\right)(\mathcal{R}_t^0, x+W_{\mathcal{R}_t^0}^0 - W_t^0)\right](\omega) \in \mathbb{R}\Big\|_k \\ &\le T\left\|F(U_{n-1,M}^0) - F(u)\right\|_{k+1} \\ &\le LT\left\|U_{n-1,M}^0 - u\right\|_{k+1}. \end{aligned} \tag{94}$$

This, (80), and (90) demonstrate that for all $n \in \mathbb{N}$, $k \in \mathbb{N}_0$ it holds that

$$\begin{aligned} \left\|U_{n,M}^0 - u\right\|_k &\le \tfrac{1}{\sqrt{k!M^n}}\left[\left(\mathbb{E}\big[|g(\xi+\mathbf{W}_T)|^2\big]\right)^{1/2} + T\|F(0)\|_1\right] \\ &\quad + \left[\sum_{l=0}^{n-1} \tfrac{(2-\mathbb{1}_{\{n-1\}}(l))LT}{\sqrt{M^{(n-l-1)}}}\left\|U_{l,M}^0 - u\right\|_{k+1}\right] + LT\left\|U_{n-1,M}^0 - u\right\|_{k+1} \\ &\le \tfrac{1}{\sqrt{k!M^n}}\left[\left(\mathbb{E}\big[|g(\xi+\mathbf{W}_T)|^2\big]\right)^{1/2} + T\|F(0)\|_1\right] + \sum_{l=0}^{n-1}\tfrac{2LT}{\sqrt{M^{(n-l-1)}}}\left\|U_{l,M}^0 - u\right\|_{k+1}. \end{aligned} \tag{95}$$

For the next step let $\varepsilon_n \in [0,\infty]$, $n \in [0,N]\cap\mathbb{N}_0$, satisfy for all $n \in [0,N]\cap\mathbb{N}_0$ that

$$\varepsilon_n = \sup\left\{\tfrac{1}{\sqrt{M^j}}\left\|U_{n,M}^0 - u\right\|_k : j,k \in \mathbb{N}_0, j+n+k = N\right\} \tag{96}$$

and let $a_1, a_2 \in [0,\infty)$ be given by

$$a_1 = \sup_{k\in\{0,\ldots,N\}} \tfrac{1}{\sqrt{k!M^{N-k}}}\left[\left(\mathbb{E}\big[|g(\xi+\mathbf{W}_T)|^2\big]\right)^{1/2} + T\|F(0)\|_1\right] \qquad \text{and} \qquad a_2 = 2LT. \tag{97}$$

Observe that (95) implies that for all $n \in [1,N]\cap\mathbb{N}$, $j,k \in \mathbb{N}_0$ with $j+n+k = N$ it holds that

$$\begin{aligned} &\tfrac{1}{\sqrt{M^j}}\left\|U_{n,M}^0 - u\right\|_k \\ &\le \tfrac{1}{\sqrt{k!M^{n+j}}}\left[\left(\mathbb{E}\big[|g(\xi+\mathbf{W}_T)|^2\big]\right)^{1/2} + T\|F(0)\|_1\right] + \sum_{l=0}^{n-1}\tfrac{2LT}{\sqrt{M^{(n+j-l-1)}}}\left\|U_{l,M}^0 - u\right\|_{k+1} \\ &\le \tfrac{1}{\sqrt{k!M^{N-k}}}\left[\left(\mathbb{E}\big[|g(\xi+\mathbf{W}_T)|^2\big]\right)^{1/2} + T\|F(0)\|_1\right] + \sum_{l=0}^{n-1}\tfrac{2LT}{\sqrt{M^{(N-k-l-1)}}}\left\|U_{l,M}^0 - u\right\|_{k+1} \\ &\le a_1 + a_2\sum_{l=0}^{n-1}\varepsilon_l. \end{aligned} \tag{98}$$

Hence, we obtain for all $n \in [1, N] \cap \mathbb{N}$ that

$$\varepsilon_n \leq a_1 + a_2 \sum_{l=0}^{n-1} \varepsilon_l = (a_1 + a_2\varepsilon_0) + a_2 \sum_{l=1}^{n-1} \varepsilon_l. \tag{99}$$

The discrete Gronwall-type inequality in [1, Corollary 4.1.2] hence proves that for all $n \in [1, N] \cap \mathbb{N}$ it holds that

$$\varepsilon_n \leq (a_1 + a_2\varepsilon_0)(1 + a_2)^{n-1}. \tag{100}$$

This, (7), and (96) imply that

$$\begin{aligned}\Big(\mathbb{E}\Big[\big|U^0_{N,M}(0,\xi) - u(0,\xi)\big|^2\Big]\Big)^{1/2} &= \big\|U^0_{N,M} - u\big\|_0 = \varepsilon_N \\ &\leq (a_1 + a_2\varepsilon_0)(1 + a_2)^{N-1} \leq \max\{a_1, \varepsilon_0\}(1 + a_2)^N.\end{aligned} \tag{101}$$

Moreover, observe that

$$\sup_{k\in\{0,\dots,N\}} \tfrac{1}{M^{(N-k)}k!} = \frac{1}{M^N} \sup_{k\in\{0,\dots,N\}} \tfrac{M^k}{k!} \leq \frac{1}{M^N} \sum_{k=0}^{\infty} \tfrac{M^k}{k!} = \frac{e^M}{M^N}. \tag{102}$$

Therefore, we obtain that

$$a_1 \leq \Big[\big(\mathbb{E}\big[|g(\xi + \mathbf{W}_T)|^2\big]\big)^{1/2} + T\,\|F(0)\|_1\Big] \frac{e^{M/2}}{M^{N/2}}. \tag{103}$$

In addition, note that the hypothesis that $U^0_{0,M} = 0$, Item (ii) in Lemma 2.11, (102), and Lemma 3.4 ensure that

$$\begin{aligned}\varepsilon_0 = \sup_{k\in\{0,\dots,N\}} \frac{\|u\|_k}{\sqrt{M^{(N-k)}}} &\leq \Big[\sup_{t\in[0,T]} \big(\mathbb{E}\big[|u(t,\xi + \mathbf{W}_t)|^2\big]\big)^{1/2}\Big] \Big[\sup_{k\in\{0,\dots,N\}} \tfrac{1}{\sqrt{M^{(N-k)}k!}}\Big] \\ &\leq e^{LT} \Big[\big(\mathbb{E}\big[|g(\xi + \mathbf{W}_T)|^2\big]\big)^{1/2} + T\|F(0)\|_1\Big] \frac{e^{M/2}}{M^{N/2}}.\end{aligned} \tag{104}$$

This and (103) assure that

$$\max\{a_1, \varepsilon_0\} \leq e^{LT} \Big[\big(\mathbb{E}\big[|g(\xi + \mathbf{W}_T)|^2\big]\big)^{1/2} + T\|F(0)\|_1\Big] \frac{e^{M/2}}{M^{N/2}}. \tag{105}$$

Combining this with (97) and (101) establishes that

$$\Big(\mathbb{E}\Big[\big|U^0_{N,M}(0,\xi) - u(0,\xi)\big|^2\Big]\Big)^{1/2} \leq e^{LT} \Big[\big(\mathbb{E}\big[|g(\xi + \mathbf{W}_T)|^2\big]\big)^{1/2} + T\,\|F(0)\|_1\Big] \frac{e^{M/2}(1 + 2LT)^N}{M^{N/2}}. \tag{106}$$

The proof of Theorem 3.5 is thus completed. □

3.5 Analysis of the computational effort

In Lemma 3.6 below, for every $n \in \mathbb{N}_0$ and every $M \in \mathbb{N}$ let $\mathrm{RV}_{n,M}$ be an upper bound for the number of realizations of random variables, which are scalar standard normal or uniformly distributed on $[0, 1]$ and required to compute one realization of $U^0_{n,M}(0, 0)$.

Lemma 3.6 (Computational effort)**.** *Let* $d \in \mathbb{N}$*,* $(\mathrm{RV}_{n,M})_{n,M\in\mathbb{Z}} \subseteq \mathbb{N}_0$ *satisfy for all* $n, M \in \mathbb{N}$ *that* $\mathrm{RV}_{0,M} = 0$ *and*

$$\mathrm{RV}_{n,M} \leq dM^n + \sum_{l=0}^{n-1} \big[M^{(n-l)}(d + 1 + \mathrm{RV}_{l,M} + \mathbb{1}_{\mathbb{N}}(l)\,\mathrm{RV}_{l-1,M})\big]. \tag{107}$$

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

Proof of Lemma 3.6. First, observe that (107) and the fact that for all $M \in \mathbb{N}$ it holds that $\mathrm{RV}_{0,M} = 0$ imply that for all $n \in \mathbb{N}$, $M \in \mathbb{N} \cap [2, \infty)$ it holds that

$$\begin{aligned}
(M^{-n} \mathrm{RV}_{n,M}) &\leq d + \sum_{l=0}^{n-1} \left[M^{-l} (d + 1 + \mathrm{RV}_{l,M} + \mathbb{1}_{\mathbb{N}}(l) \, \mathrm{RV}_{l-1,M}) \right] \\
&\leq d + (d+1) \left[\sum_{l=0}^{n-1} M^{-l} \right] + \left[\sum_{l=0}^{n-1} M^{-l} \, \mathrm{RV}_{l,M} \right] + \left[\sum_{l=0}^{n-2} M^{-(l+1)} \, \mathrm{RV}_{l,M} \right] \\
&= d + (d+1) \tfrac{(1-M^{-n})}{(1-M^{-1})} + \left[\sum_{l=0}^{n-1} M^{-l} \, \mathrm{RV}_{l,M} \right] + \frac{1}{M} \left[\sum_{l=0}^{n-2} M^{-l} \, \mathrm{RV}_{l,M} \right] \\
&\leq d + (d+1) \tfrac{1}{(1-\frac{1}{2})} + \left(1 + \tfrac{1}{M}\right) \left[\sum_{l=0}^{n-1} M^{-l} \, \mathrm{RV}_{l,M} \right] \\
&= 3d + 2 + \left(1 + \tfrac{1}{M}\right) \left[\sum_{l=1}^{n-1} M^{-l} \, \mathrm{RV}_{l,M} \right].
\end{aligned} \tag{108}$$

The discrete Gronwall-type inequality in [1, Corollary 4.1.2] hence ensures that for all $n \in \mathbb{N}$, $M \in \mathbb{N} \cap [2, \infty)$ it holds that

$$(M^{-n} \mathrm{RV}_{n,M}) \leq (3d+2) \left(2 + \tfrac{1}{M}\right)^{n-1}. \tag{109}$$

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

$$\mathrm{RV}_{n,M} \leq (3d+2) \left(2 + \tfrac{1}{M}\right)^{n-1} M^n \leq (5d) 3^{n-1} M^n \leq d(5M)^n. \tag{110}$$

Moreover, observe that the fact that $\mathrm{RV}_{0,1} = 0$ and (107) demonstrate that for all $n \in \mathbb{N}$ it holds that

$$\mathrm{RV}_{n,1} \leq d + \sum_{l=0}^{n-1} (d + 1 + \mathrm{RV}_{l,1} + \mathbb{1}_{\mathbb{N}}(l) \, \mathrm{RV}_{l-1,1}) \leq d + n(d+1) + 2 \sum_{l=1}^{n-1} \mathrm{RV}_{l,1}. \tag{111}$$

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

$$\mathrm{RV}_{k,1} \leq d + n(d+1) + 2 \sum_{l=1}^{k-1} \mathrm{RV}_{l,1}. \tag{112}$$

Combining this with the discrete Gronwall-type inequality in [1, Corollary 4.1.2] proves that for all $n \in \mathbb{N}$, $k \in \mathbb{N} \cap (0, n]$ it holds that

$$\mathrm{RV}_{k,1} \leq (d + n(d+1)) 3^{k-1}. \tag{113}$$

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

$$\mathrm{RV}_{n,1} \leq (d + n(d+1)) 3^{n-1} = d \left(1 + n \left(1 + \tfrac{1}{d}\right)\right) 3^{n-1} \leq d \, (1 + 2n) 3^{n-1} \leq d \, 5^n. \tag{114}$$

Combining this with (110) completes the proof of Lemma 3.6. ◻

Corollary 3.7. *Assume Setting 3.1, and let $\delta \in (0, \infty)$, $C \in (0, \infty]$, $(\mathrm{RV}_{n,M})_{n,M \in \mathbb{Z}} \subseteq \mathbb{N}_0$ satisfy for all $n, M \in \mathbb{N}$ that*

$$\mathrm{RV}_{0,M} = 0, \qquad \mathrm{RV}_{n,M} \leq dM^n + \sum_{l=0}^{n-1} \left[M^{(n-l)} (d + 1 + \mathrm{RV}_{l,M} + \mathbb{1}_{\mathbb{N}}(l) \, \mathrm{RV}_{l-1,M}) \right]. \tag{115}$$

$$\textit{and} \qquad C = \left[e^{LT}\left[\left(\mathbb{E}\big[|g(\xi+\mathbf{W}_T)|^2\big]\right)^{1/2} + T\,\|F(0)\|_1\right]\right]^{2+\delta}\left[\sup_{n\in\mathbb{N}}\tfrac{(4+8LT)^{n(2+\delta)}}{n^{((n\delta)/2)}}\right]. \tag{116}$$

Then it holds for all $N\in\mathbb{N}$ that

$$\mathrm{RV}_{N,N} \le d\,C\left[\left(\mathbb{E}\big[|U^0_{N,N}(0,\xi)-u(0,\xi)|^2\big]\right)^{1/2}\right]^{-(2+\delta)}. \tag{117}$$

Proof of Corollary 3.7. Throughout this proof let $N\in\mathbb{N}$, assume w.l.o.g. that $C<\infty$ and $\mathbb{E}\big[|U^0_{N,N}(0,\xi)-u(0,\xi)|^2\big]>0$, and let $c,\kappa,\varepsilon\in(0,\infty)$ be given by

$$c = e^{LT}\left[\left(\mathbb{E}\big[|g(\xi+\mathbf{W}_T)|^2\big]\right)^{1/2} + T\,\|F(0)\|_1\right], \qquad \kappa=\sqrt{e}(1+2LT), \tag{118}$$

$$\text{and} \qquad \varepsilon = \left(\mathbb{E}\big[|U^0_{N,N}(0,\xi)-u(0,\xi)|^2\big]\right)^{1/2}. \tag{119}$$

Observe that Theorem 3.5 (with $N=N$, $M=N$ in the notation of Theorem 3.5) proves that

$$\varepsilon \le \frac{c\,e^{N/2}(1+2LT)^N}{N^{N/2}} = \frac{c\kappa^N}{N^{N/2}}. \tag{120}$$

Lemma 3.6 hence demonstrates that

$$\begin{aligned}
\mathrm{RV}_{N,N} &\le d(5N)^N = d(5N)^N\varepsilon^{2+\delta}\varepsilon^{-(2+\delta)} \le d(5N)^N\left(\frac{c\,\kappa^N}{N^{N/2}}\right)^{2+\delta}\varepsilon^{-(2+\delta)} \\
&= \frac{d\,c^{2+\delta}(5N)^N\kappa^{N(2+\delta)}}{N^{N+\frac{N\delta}{2}}}\varepsilon^{-(2+\delta)} = d\,c^{2+\delta}\left(\frac{5^N\kappa^{N(2+\delta)}}{N^{((N\delta)/2)}}\right)\varepsilon^{-(2+\delta)} \\
&\le d\,c^{2+\delta}\left[\sup_{n\in\mathbb{N}}\frac{(5\kappa^{(2+\delta)})^n}{n^{((n\delta)/2)}}\right]\varepsilon^{-(2+\delta)}.
\end{aligned} \tag{121}$$

In addition, observe that the fact that $\sqrt{5e}\le 4$ assures that

$$5\kappa^{(2+\delta)} \le (\sqrt{5e}(1+2LT))^{(2+\delta)} \le (4(1+2LT))^{(2+\delta)} = (4+8LT)^{(2+\delta)}. \tag{122}$$

This and (121) establish that

$$\mathrm{RV}_{N,N} \le d\,c^{2+\delta}\left[\sup_{n\in\mathbb{N}}\frac{(4+8LT)^{n(2+\delta)}}{n^{((n\delta)/2)}}\right]\varepsilon^{-(2+\delta)} = d\,C\,\varepsilon^{-(2+\delta)}. \tag{123}$$

The proof of Corollary 3.7 is thus completed. ☐

Theorem 3.8. *Let $d\in\mathbb{N}$, $L,T,\delta\in(0,\infty)$, $C\in(0,\infty]$, $\xi\in\mathbb{R}^d$, $\Theta=\cup_{n=1}^\infty\mathbb{Z}^n$, let $f\in C([0,T]\times\mathbb{R}^d\times\mathbb{R},\mathbb{R})$, $g\in C(\mathbb{R}^d,\mathbb{R})$, $u\in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ be at most polynomially growing functions, let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $W^\theta\colon[0,T]\times\Omega\to\mathbb{R}^d$, $\theta\in\Theta$, be independent standard Brownian motions with continuous sample paths, let $\mathfrak{r}^\theta\colon\Omega\to[0,1]$, $\theta\in\Theta$, be independent $\mathcal{U}_{[0,1]}$-distributed random variables, assume that $(W^\theta)_{\theta\in\Theta}$ and $(\mathfrak{r}^\theta)_{\theta\in\Theta}$ are independent, assume for all $t\in[0,T]$, $x\in\mathbb{R}^d$, $v,w\in\mathbb{R}$ that*

$$u(t,x) = \mathbb{E}\Big[g(x+W^0_{T-t}) + \textstyle\int_t^T f\big(s,x+W^0_{s-t},u(s,x+W^0_{s-t})\big)\,ds\Big], \tag{124}$$

$$C = \tag{125}$$
$$\left[e^{LT}\left(\left(\mathbb{E}\big[|g(\xi+W^0_T)|^2\big]\right)^{1/2} + \sqrt{T}\,\big|\textstyle\int_0^T\mathbb{E}\big[|f(s,\xi+W^0_s,0)|^2\big]\,ds\big|^{1/2}\right)\right]^{2+\delta}\left[\sup_{n\in\mathbb{N}}\tfrac{(4+8LT)^{n(2+\delta)}}{n^{((n\delta)/2)}}\right],$$

and $|f(t,x,v)-f(t,x,w)| \leq L|v-w|$, *let* $\mathcal{R}^\theta\colon [0,T]\times\Omega\to[0,T]$, $\theta\in\Theta$, *and* $U^\theta_{n,M}\colon [0,T]\times\mathbb{R}^d\times\Omega\to\mathbb{R}$, $n,M\in\mathbb{Z}$, $\theta\in\Theta$, *satisfy for all* $n,M\in\mathbb{N}$, $\theta\in\Theta$, $t\in[0,T]$, $x\in\mathbb{R}^d$ *that* $\mathcal{R}^\theta_t = t+(T-t)\mathfrak{r}^\theta$, $U^\theta_{-1,M}(t,x)=U^\theta_{0,M}(t,x)=0$, *and*

$$\begin{aligned} U^\theta_{n,M}(t,x) = \Bigg[\sum_{l=0}^{n-1}\tfrac{(T-t)}{M^{n-l}}\sum_{i=1}^{M^{n-l}} f\Big(\mathcal{R}^{(\theta,l,i)}_t, x+W^{(\theta,l,i)}_{\mathcal{R}^{(\theta,l,i)}_t-t}, U^{(\theta,l,i)}_{l,M}(\mathcal{R}^{(\theta,l,i)}_t, x+W^{(\theta,l,i)}_{\mathcal{R}^{(\theta,l,i)}_t-t})\Big) \\ -\mathbb{1}_{\mathbb{N}}(l) f\Big(\mathcal{R}^{(\theta,l,i)}_t, x+W^{(\theta,l,i)}_{\mathcal{R}^{(\theta,l,i)}_t-t}, U^{(\theta,l,i)}_{l-1,M}(\mathcal{R}^{(\theta,l,i)}_t, x+W^{(\theta,l,i)}_{\mathcal{R}^{(\theta,l,i)}_t-t})\Big)\Bigg] + \sum_{i=1}^{M^n}\frac{g(x+W^{(\theta,0,-i)}_{T-t})}{M^n}, \end{aligned} \tag{126}$$

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

$$\mathrm{RV}_{n,M} \leq dM^n + \sum_{l=0}^{n-1}\left[M^{(n-l)}(d+1+\mathrm{RV}_{l,M}+\mathbb{1}_{\mathbb{N}}(l)\,\mathrm{RV}_{l-1,M})\right]. \tag{127}$$

Then

(i) it holds that $\limsup_{n\to\infty}\mathbb{E}\big[|u(0,\xi)-U^0_{n,n}(0,\xi)|^2\big]=0$,

(ii) it holds that $C<\infty$, *and*

(iii) it holds for all $N\in\mathbb{N}$ *that* $\mathrm{RV}_{N,N}\leq d\,C\left[\big(\mathbb{E}\big[|u(0,\xi)-U^0_{N,N}(0,\xi)|^2\big]\big)^{1/2}\right]^{-(2+\delta)}$.

Proof of Theorem 3.8. Throughout this proof let $F\colon C([0,T]\times\mathbb{R}^d,\mathbb{R})\to C([0,T]\times\mathbb{R}^d,\mathbb{R})$ be the function which satisfies for all $v\in C([0,T]\times\mathbb{R}^d,\mathbb{R})$, $t\in[0,T]$, $x\in\mathbb{R}^d$ that

$$(F(v))(t,x) = f(t,x,v(t,x)). \tag{128}$$

Observe that the hypothesis that for all $t\in[0,T]$, $x\in\mathbb{R}^d$, $v,w\in\mathbb{R}$ it holds that $|f(t,x,v)-f(t,x,w)|\leq L|v-w|$ ensures that for all $v,w\in C([0,T]\times\mathbb{R}^d,\mathbb{R})$, $t\in[0,T]$, $x\in\mathbb{R}^d$ it holds that

$$|(F(v))(t,x)-(F(w))(t,x)| = |f(t,x,v(t,x))-f(t,x,w(t,x))| \leq L\,|v(t,x)-w(t,x)|. \tag{129}$$

Moreover, note that the hypothesis that $f\colon[0,T]\times\mathbb{R}^d\times\mathbb{R}\to\mathbb{R}$, $g\colon\mathbb{R}^d\to\mathbb{R}$, and $u\colon[0,T]\times\mathbb{R}^d\to\mathbb{R}$ are at most polynomially growing functions and the fact that for all $p\in(0,\infty)$ it holds that

$$\mathbb{E}\big[\sup_{t\in[0,T]}\|W^0_t\|^p_{\mathbb{R}^d}\big]<\infty \tag{130}$$

demonstrate that

$$\begin{aligned} &\mathbb{E}\big[|g(x+W^0_t)|\big] + \int_0^T \big(\mathbb{E}\big[|u(s,\xi+W^0_s)|^2\big]\big)^{1/2}\,ds \\ &+\int_t^T \mathbb{E}\big[|(F(u))(s,x+W^0_{s-t})|+|(F(0))(s,x+W^0_{s-t})|\big]\,ds<\infty. \end{aligned} \tag{131}$$

Combining this and (129) with Theorem 3.5 establishes Item (i). In addition, observe that the hypothesis (130) and the hypothesis that $f\colon[0,T]\times\mathbb{R}^d\times\mathbb{R}\to\mathbb{R}$ and $g\colon\mathbb{R}^d\to\mathbb{R}$, are at most polynomially growing functions ensure that

$$\left[e^{LT}\Big(\big(\mathbb{E}\big[|g(\xi+W^0_T)|^2\big]\big)^{1/2}+\sqrt{T}\,\big|\textstyle\int_0^T\mathbb{E}\big[|f(s,\xi+W^0_s,0)|^2\big]\,ds\big|^{1/2}\Big)\right]^{2+\delta}<\infty. \tag{132}$$

The fact that

$$\begin{aligned} &\sup_{n\in\mathbb{N}}\left(\frac{(4+8LT)^{n(2+\delta)}}{n^{((n\delta)/2)}}\right) = \left[\sup_{n\in\mathbb{N}}\left(\frac{\big[(4+8LT)^{2(2+\delta)/\delta}\big]^n}{n^n}\right)\right]^{\delta/2} \\ &\leq\left[\sum_{n=1}^{\infty}\left(\frac{\big[(4+8LT)^{2(2+\delta)/\delta}\big]^n}{n^n}\right)\right]^{\delta/2} \leq \left[\exp\big((4+8LT)^{2(2+\delta)/\delta}\big)\right]^{\delta/2} \\ &= \exp\big(\tfrac{\delta}{2}\left[(4+8LT)^{2(2+\delta)/\delta}\right]\big)<\infty \end{aligned} \tag{133}$$

hence establishes Item (ii). In addition, observe that (129), (131), and Corollary 3.7 establish Item (iii). The proof of Theorem 3.8 is thus completed. $\square$

3.6 Existence, uniqueness, and regularity properties for solutions of certain stochastic fixed point equations

In this subsection we employ in Corollary 3.11 below the Banach fixed point theorem to establish the existence of unique solutions of certain stochastic fixed point equations (cf. (124) above and (143) below). Our proof of Corollary 3.11 uses Lemma 3.9 and Lemma 3.10 below. Lemma 3.9 and Lemma 3.10 establish basic a priori estimates for such stochastic fixed point equations.

Lemma 3.9. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $T, p, L \in [0, \infty)$, $d \in \mathbb{N}$, $f \in C([0, T] \times \mathbb{R}^d \times \mathbb{R}, \mathbb{R})$, $g \in C(\mathbb{R}^d, \mathbb{R})$, $u \in C([0, T] \times \mathbb{R}^d, \mathbb{R})$ satisfy for all $t \in [0, T]$, $x \in \mathbb{R}^d$, $v, w \in \mathbb{R}$ that $\max\{|f(t, x, 0)|, |g(x)|, |u(t, x)|\} \leq L(1 + \|x\|^p_{\mathbb{R}^d})$ and $|f(t, x, v) - f(t, x, w)| \leq L|v - w|$, and let $W : [0, T] \times \Omega \to \mathbb{R}^d$ be a standard Brownian motion with continuous sample paths. Then it holds for all $t \in [0, T]$, $x \in \mathbb{R}^d$ that*

$$
\begin{aligned}
&\mathbb{E}\Big[|g(x + W_{T-t})| + \textstyle\int_t^T \big|f\big(s, x + W_{s-t}, u(s, x + W_{s-t})\big)\big|\, ds\Big] \\
&\leq 2^{\max\{p-1,0\}} L(LT + T + 1) \left(1 + \|x\|^p_{\mathbb{R}^d} + \mathbb{E}\big[\|W_T\|^p_{\mathbb{R}^d}\big]\right) \\
&\leq \left(2^{\max\{p-1,0\}} L(LT + T + 1) \left(1 + \mathbb{E}\big[\|W_T\|^p_{\mathbb{R}^d}\big]\right)\right) \left(1 + \|x\|^p_{\mathbb{R}^d}\right) < \infty.
\end{aligned}
\tag{134}
$$

Proof of Lemma 3.9. Observe that for all $t \in [0, T]$, $x \in \mathbb{R}^d$ it holds that

$$
\begin{aligned}
\mathbb{E}\big[|g(x + W_{T-t})|\big] &\leq \mathbb{E}\big[L(1 + \|x + W_{T-t}\|^p_{\mathbb{R}^d})\big] = L\left(1 + \mathbb{E}\big[\|x + W_{T-t}\|^p_{\mathbb{R}^d}\big]\right) \\
&\leq 2^{\max\{p-1,0\}} L \left(1 + \|x\|^p_{\mathbb{R}^d} + \mathbb{E}\big[\|W_{T-t}\|^p_{\mathbb{R}^d}\big]\right) \\
&\leq 2^{\max\{p-1,0\}} L \left(1 + \|x\|^p_{\mathbb{R}^d} + \mathbb{E}\big[\|W_T\|^p_{\mathbb{R}^d}\big]\right).
\end{aligned}
\tag{135}
$$

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

$$
\begin{aligned}
&\mathbb{E}\Big[\textstyle\int_t^T \big|f\big(s, x + W_{s-t}, u(s, x + W_{s-t})\big)\big|\, ds\Big] = \displaystyle\int_t^T \mathbb{E}\Big[\big|f\big(s, x + W_{s-t}, u(s, x + W_{s-t})\big)\big|\Big]\, ds \\
&\leq \int_t^T \mathbb{E}\Big[\big|f\big(s, x + W_{s-t}, u(s, x + W_{s-t})\big) - f\big(s, x + W_{s-t}, 0\big)\big| + \big|f\big(s, x + W_{s-t}, 0\big)\big|\Big]\, ds \\
&\leq \int_t^T \mathbb{E}\big[L|u(s, x + W_{s-t})| + |f(s, x + W_{s-t}, 0)|\big]\, ds \\
&\leq \int_t^T \mathbb{E}\big[L|u(s, x + W_{s-t})| + L(1 + \|x + W_{s-t}\|^p_{\mathbb{R}^d})\big]\, ds.
\end{aligned}
\tag{136}
$$

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

$$
\begin{aligned}
&\mathbb{E}\Big[\textstyle\int_t^T \big|f\big(s, x + W_{s-t}, u(s, x + W_{s-t})\big)\big|\, ds\Big] \\
&\leq L \int_t^T \mathbb{E}\big[|u(s, x + W_{s-t})| + 1 + \|x + W_{s-t}\|^p_{\mathbb{R}^d}\big]\, ds \\
&\leq L \int_t^T \mathbb{E}\big[L(1 + \|x + W_{s-t}\|^p_{\mathbb{R}^d}) + 1 + \|x + W_{s-t}\|^p_{\mathbb{R}^d}\big]\, ds \\
&= L(L + 1) \left[\int_t^T 1 + \mathbb{E}\big[\|x + W_{s-t}\|^p_{\mathbb{R}^d}\big]\, ds\right] \\
&\leq 2^{\max\{p-1,0\}} L(L + 1) \left[\int_t^T 1 + \|x\|^p_{\mathbb{R}^d} + \mathbb{E}\big[\|W_{s-t}\|^p_{\mathbb{R}^d}\big]\, ds\right].
\end{aligned}
\tag{137}
$$

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

$$\begin{aligned}
&\mathbb{E}\Big[\textstyle\int_t^T \big|f\big(s, x+W_{s-t}, u(s,x+W_{s-t})\big)\big|\, ds\Big] \\
&\leq 2^{\max\{p-1,0\}} L(L+1) \left[\int_t^T 1 + \|x\|_{\mathbb{R}^d}^p + \mathbb{E}\big[\|W_{s-t}\|_{\mathbb{R}^d}^p\big]\, ds\right] \\
&\leq 2^{\max\{p-1,0\}} LT(L+1) \left(1 + \|x\|_{\mathbb{R}^d}^p + \mathbb{E}\big[\|W_T\|_{\mathbb{R}^d}^p\big]\right).
\end{aligned} \tag{138}$$

Combining this with (135) establishes (134). The proof of Lemma 3.9 is thus completed. □

Lemma 3.10. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $T, p, L \in [0,\infty)$, $d \in \mathbb{N}$, $u, v \in C([0,T] \times \mathbb{R}^d, \mathbb{R})$, $f \in C([0,T] \times \mathbb{R}^d \times \mathbb{R}, \mathbb{R})$ satisfy for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $a, b \in \mathbb{R}$ that $|f(t,x,a) - f(t,x,b)| \leq L|a-b|$, and let $W \colon [0,T] \times \Omega \to \mathbb{R}^d$ be a standard Brownian motion with continuous sample paths. Then it holds for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $\lambda \in (0,\infty)$ that*

$$\begin{aligned}
&e^{\lambda t}\, \mathbb{E}\Big[\textstyle\int_t^T \big|f\big(s, x+W_{s-t}, u(s,x+W_{s-t})\big) - f\big(s, x+W_{s-t}, v(s,x+W_{s-t})\big)\big|\, ds\Big] \\
&\leq \left[\frac{2^{\max\{p-1,0\}} L \left(1 + \|x\|_{\mathbb{R}^d}^p + \mathbb{E}\big[\|W_T\|_{\mathbb{R}^d}^p\big]\right)}{\lambda}\right] \left[\sup_{s\in[t,T]} \sup_{y\in\mathbb{R}^d} \left(\frac{e^{\lambda s}|u(s,y)-v(s,y)|}{(1+\|y\|_{\mathbb{R}^d}^p)}\right)\right] \\
&\leq \left[\frac{2^{\max\{p-1,0\}} L \left(1 + \mathbb{E}\big[\|W_T\|_{\mathbb{R}^d}^p\big]\right)}{\lambda}\right] \left[\sup_{s\in[t,T]} \sup_{y\in\mathbb{R}^d} \left(\frac{e^{\lambda s}|u(s,y)-v(s,y)|}{(1+\|y\|_{\mathbb{R}^d}^p)}\right)\right] \left(1 + \|x\|_{\mathbb{R}^d}^p\right).
\end{aligned} \tag{139}$$

Proof of Lemma 3.10. Observe that for all $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
&\mathbb{E}\Big[\textstyle\int_t^T \big|f\big(s, x+W_{s-t}, u(s,x+W_{s-t})\big) - f\big(s, x+W_{s-t}, v(s,x+W_{s-t})\big)\big|\, ds\Big] \\
&= \int_t^T \mathbb{E}\Big[\big|f\big(s, x+W_{s-t}, u(s,x+W_{s-t})\big) - f\big(s, x+W_{s-t}, v(s,x+W_{s-t})\big)\big|\Big]\, ds \\
&\leq L \int_t^T \mathbb{E}\Big[\big|u(s,x+W_{s-t}) - v(s,x+W_{s-t})\big|\Big]\, ds \\
&= L \int_t^T \mathbb{E}\left[\left(\frac{|u(s,x+W_{s-t}) - v(s,x+W_{s-t})|}{1 + \|x+W_{s-t}\|_{\mathbb{R}^d}^p}\right)\left(1 + \|x+W_{s-t}\|_{\mathbb{R}^d}^p\right)\right] ds.
\end{aligned} \tag{140}$$

Therefore, we obtain that for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $\lambda \in \mathbb{R}$ it holds that

$$\begin{aligned}
&\mathbb{E}\Big[\textstyle\int_t^T \big|f\big(s, x+W_{s-t}, u(s,x+W_{s-t})\big) - f\big(s, x+W_{s-t}, v(s,x+W_{s-t})\big)\big|\, ds\Big] \\
&\leq L \int_t^T \mathbb{E}\left[\left(\sup_{y\in\mathbb{R}^d}\left[\frac{|u(s,y)-v(s,y)|}{1+\|y\|_{\mathbb{R}^d}^p}\right]\right)\left(1 + \|x+W_{s-t}\|_{\mathbb{R}^d}^p\right)\right] ds \\
&\leq L \left[\int_t^T e^{-\lambda s}\, \mathbb{E}\big[\left(1 + \|x+W_{s-t}\|_{\mathbb{R}^d}^p\right)\big]\, ds\right] \left[\sup_{s\in[t,T]} \sup_{y\in\mathbb{R}^d} \left(\frac{e^{\lambda s}|u(s,y)-v(s,y)|}{(1+\|y\|_{\mathbb{R}^d}^p)}\right)\right] \\
&\leq 2^{\max\{p-1,0\}} L \left[\int_t^T e^{-\lambda s} \left(1 + \|x\|_{\mathbb{R}^d}^p + \mathbb{E}\big[\|W_T\|_{\mathbb{R}^d}^p\big]\right) ds\right] \left[\sup_{s\in[t,T]} \sup_{y\in\mathbb{R}^d} \left(\frac{e^{\lambda s}|u(s,y)-v(s,y)|}{(1+\|y\|_{\mathbb{R}^d}^p)}\right)\right].
\end{aligned} \tag{141}$$

This shows that for all $t \in [0,T]$, $x \in \mathbb{R}^d$. $\lambda \in (0,\infty)$ it holds that

$$
\begin{aligned}
&e^{\lambda t}\,\mathbb{E}\Big[\textstyle\int_t^T \big|f\big(s, x+W_{s-t}, u(s,x+W_{s-t})\big) - f\big(s,x+W_{s-t}, v(s,x+W_{s-t})\big)\big|\,ds\Big] \\
&\leq 2^{\max\{p-1,0\}} L \left[\int_t^T e^{\lambda(t-s)}\left(1+\|x\|_{\mathbb{R}^d}^p + \mathbb{E}\big[\|W_T\|_{\mathbb{R}^d}^p\big]\right) ds\right] \left[\sup_{s\in[t,T]}\sup_{y\in\mathbb{R}^d}\left(\frac{e^{\lambda s}|u(s,y)-v(s,y)|}{(1+\|y\|_{\mathbb{R}^d}^p)}\right)\right] \\
&= 2^{\max\{p-1,0\}} L \left[\int_0^{T-t} e^{-\lambda s}\,ds\right]\left(1+\|x\|_{\mathbb{R}^d}^p + \mathbb{E}\big[\|W_T\|_{\mathbb{R}^d}^p\big]\right)\left[\sup_{s\in[t,T]}\sup_{y\in\mathbb{R}^d}\left(\frac{e^{\lambda s}|u(s,y)-v(s,y)|}{(1+\|y\|_{\mathbb{R}^d}^p)}\right)\right] \\
&\leq 2^{\max\{p-1,0\}} L \left[\int_0^{\infty} e^{-\lambda s}\,ds\right]\left(1+\|x\|_{\mathbb{R}^d}^p + \mathbb{E}\big[\|W_T\|_{\mathbb{R}^d}^p\big]\right)\left[\sup_{s\in[t,T]}\sup_{y\in\mathbb{R}^d}\left(\frac{e^{\lambda s}|u(s,y)-v(s,y)|}{(1+\|y\|_{\mathbb{R}^d}^p)}\right)\right] \\
&= \left[\frac{2^{\max\{p-1,0\}} L \left(1+\|x\|_{\mathbb{R}^d}^p + \mathbb{E}\big[\|W_T\|_{\mathbb{R}^d}^p\big]\right)}{\lambda}\right]\left[\sup_{s\in[t,T]}\sup_{y\in\mathbb{R}^d}\left(\frac{e^{\lambda s}|u(s,y)-v(s,y)|}{(1+\|y\|_{\mathbb{R}^d}^p)}\right)\right].
\end{aligned} \tag{142}
$$

This establishes (139). The proof of Lemma 3.10 is thus completed. □

Corollary 3.11. *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $T,p,L \in [0,\infty)$, $d \in \mathbb{N}$, $f \in C([0,T]\times\mathbb{R}^d\times\mathbb{R},\mathbb{R})$, $g \in C(\mathbb{R}^d,\mathbb{R})$ satisfy for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $v,w \in \mathbb{R}$ that $\max\{|g(x)|,|f(t,x,0)|\} \leq L(1+\|x\|_{\mathbb{R}^d}^p)$ and $|f(t,x,v)-f(t,x,w)| \leq L|v-w|$, and let $W\colon [0,T]\times\Omega \to \mathbb{R}^d$ be a standard Brownian motion with continuous sample paths. Then there exists a unique continuous function $u\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}$ such that for all $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that $\sup_{s\in[0,T]}\sup_{y\in\mathbb{R}^d}\big(\frac{|u(s,y)|}{1+\|y\|_{\mathbb{R}^d}^p}\big) < \infty$ and*

$$
u(t,x) = \mathbb{E}\Big[g(x+W_{T-t}) + \textstyle\int_t^T f\big(s,x+W_{s-t},u(s,x+W_{s-t})\big)\,ds\Big]. \tag{143}
$$

Proof of Corollary 3.11. Throughout this proof let $\mathcal{V}$ be the set given by

$$
\mathcal{V} = \left\{u \in C([0,T]\times\mathbb{R}^d,\mathbb{R})\colon \left[\sup_{t\in[0,T]}\sup_{x\in\mathbb{R}^d}\left(\frac{|u(t,x)|}{(1+\|x\|_{\mathbb{R}^d}^p)}\right) < \infty\right]\right\} \tag{144}
$$

and let $\|\cdot\|_\lambda : \mathcal{V} \to [0,\infty)$, $\lambda \in \mathbb{R}$, be the functions which satisfy for all $\lambda \in \mathbb{R}$, $u \in \mathcal{V}$ that

$$
\|u\|_\lambda = \sup_{t\in[0,T]}\sup_{x\in\mathbb{R}^d}\left(\frac{e^{\lambda t}|u(t,x)|}{(1+\|x\|_{\mathbb{R}^d}^p)}\right). \tag{145}
$$

Observe that Lemma 3.9 ensures that there exists a unique function $\Phi\colon \mathcal{V} \to \mathcal{V}$ which satisfies for all $u \in \mathcal{V}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$
\big(\Phi(u)\big)(t,x) = \mathbb{E}\Big[g(x+W_{T-t}) + \textstyle\int_t^T f\big(s,x+W_{s-t},u(s,x+W_{s-t})\big)\,ds\Big]. \tag{146}
$$

Moreover, note that Lemma 3.10 proves that for all $\lambda \in (0,\infty)$, $u,v \in \mathcal{V}$ it holds that

$$
\|\Phi(u)-\Phi(v)\|_\lambda \leq \left[\frac{2^{\max\{p-1,0\}} L \left(1+\mathbb{E}\big[\|W_T\|_{\mathbb{R}^d}^p\big]\right)}{\lambda}\right]\|u-v\|_\lambda \tag{147}
$$

Combining this with Banach's fixed point theorem establishes (143). The proof of Corollary 3.11 is thus completed. □

Acknowledgements

This project has been partially supported through the German Research Foundation research grant HU1889/6-1. Christian Beck, Ramon Braunwarth, and Emilia Magnani are gratefully acknowledged for bringing a few typos into our notice.

References

[1] Agarwal, R. P. *Difference equations and inequalities. Theory, methods, and applications.* Marcel Dekker, Inc., 2000.

[2] 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.

[3] Beck, C., E, W., and Jentzen, A. Machine learning approximation algorithms for high-dimensional fully nonlinear partial differential equations and second-order backward stochastic differential equations. *arXiv:1709.05963* (2017), 56 pages. Revision requested from J. Nonlinear Sci.

[4] Becker, S., Cheridito, P., and Jentzen, A. Deep optimal stopping. *arXiv:1804.05394* (2018), 18 pages.

[5] Bouchard, B., Tan, X., Warin, X., and Zou, Y. Numerical approximation of bsdes using local polynomial drivers and branching processes. *Monte Carlo Methods and Applications 23*, 4 (2017), 241–263.

[6] 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.

[7] Briand, P., Labart, C., et al. Simulation of bsdes by wiener chaos expansion. *The Annals of Applied Probability 24*, 3 (2014), 1129–1171.

[8] Creutzig, J., Dereich, S., Müller-Gronbach, T., and Ritter, K. Infinite-dimensional quadrature and approximation of distributions. *Found. Comput. Math. 9*, 4 (2009), 391–429.

[9] Crisan, D., and Manolarakis, K. Probabilistic methods for semilinear partial differential equations. applications to finance. *ESAIM: Mathematical Modelling and Numerical Analysis 44*, 05 (2010), 1107–1133.

[10] E, W., Han, J., and Jentzen, A. Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. *Commun. Math. Stat. 5*, 4 (2017), 349–380.

[11] E, W., Hutzenthaler, M., Jentzen, A., and Kruse, T. Linear scaling algorithms for solving high-dimensional nonlinear parabolic differential equations. *arXiv:1605.00856* (2016).

[12] 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. *arXiv:1708.03223* (2017).

[13] Fujii, M., Takahashi, A., and Takahashi, M. Asymptotic Expansion as Prior Knowledge in Deep Learning Method for high dimensional BSDEs. *arXiv:1710.07030* (2017).

[14] Giles, M. B. Improved multilevel Monte Carlo convergence using the Milstein scheme. In *Monte Carlo and quasi-Monte Carlo methods 2006*. Springer, Berlin, 2008, pp. 343–358.

[15] Gobet, E., Lemor, J.-P., and Warin, X. A regression-based monte carlo method to solve backward stochastic differential equations. *The Annals of Applied Probability 15*, 3 (2005), 2172–2202.

[16] Graham, C., and Talay, D. *Stochastic simulation and Monte Carlo methods*, vol. 68 of *Stochastic Modelling and Applied Probability.* Springer, Heidelberg, 2013. Mathematical foundations of stochastic simulation.

[17] Han, J., Jentzen, A., and E, W. Overcoming the curse of dimensionality: Solving high-dimensional partial differential equations using deep learning. *arXiv:1707.02568* (2017), 1–13. Accepted in Proc. Natl. Acad. Sci. USA.

[18] Han, J., Jentzen, A., and Weinan, E. Solving high-dimensional partial differential equations using deep learning. *arXiv:1707.02568* (2017). To appear in Proc. Natl. Acad. Sci.

[19] Heinrich, S. Monte Carlo complexity of global solution of integral equations. *J. Complexity 14*, 2 (1998), 151–175.

[20] Heinrich, S. Multilevel Monte Carlo methods. In *Large-Scale Scientific Computing*, vol. 2179 of *Lect. Notes Comput. Sci.* Springer, Berlin, 2001, pp. 58–67.

[21] Henry-Labordère, P. Counterparty risk valuation: a marked branching diffusion approach. *arXiv:1203.2369* (2012), 17 pages.

[22] Henry-Labordère, P. Deep Primal-Dual Algorithm for BSDEs: Applications of Machine Learning to CVA and IM. *Available at SSRN: http://dx.doi.org/10.2139/ssrn.3071506* (2017).

[23] Henry-Labordère, P., Oudjane, N., Tan, X., Touzi, N., and Warin, X. Branching diffusion representation of semilinear pdes and monte carlo approximation. *arXiv:1603.01727* (2016), 30 pages.

[24] Henry-Labordère, P., Tan, X., and Touzi, N. A numerical algorithm for a class of BSDEs via the branching process. *Stochastic Process. Appl. 124*, 2 (2014), 1112–1140.

[25] Hutzenthaler, M., and Kruse, T. Multi-level Picard approximations of high-dimensional semilinear parabolic differential equations with gradient-dependent nonlinearities. *arXiv:1711.01080* (2017).

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

[27] Novak, E., and Woźniakowski, H. *Tractability of Multivariate Problems: Standard information for functionals*, vol. 12. European Mathematical Society, 2008.

[28] Pardoux, É., and Peng, S. G. Adapted solution of a backward stochastic differential equation. *Systems Control Lett. 14*, 1 (1990), 55–61.

[29] Raissi, M. Forward-Backward Stochastic Neural Networks: Deep Learning of High-dimensional Partial Differential Equations. *arXiv preprint arXiv:1804.07010* (2018).

[30] Sirignano, J., and Spiliopoulos, K. DGM: A deep learning algorithm for solving partial differential equations. *arXiv preprint arXiv:1708.07469* (2017).

[31] Warin, X. Monte carlo for high-dimensional degenerated Semi Linear and Full Non Linear PDEs. *arXiv:1805.05078* (2018), 23 pages.

[32] Warin, X. Nesting Monte Carlo for high-dimensional Non Linear PDEs. *arXiv:1804.08432* (2018), 45 pages.

[33] Weinan, E., and Yu, B. The Deep Ritz method: A deep learning-based numerical algorithm for solving variational problems. *Communications in Mathematics and Statistics 6*, 1 (2018), 1–12.