Generalised multilevel Picard approximations

M. Giles and A. Jentzen and T. Welti

Research Report No. 2019-59
November 2019

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

Generalised multilevel Picard approximations

Michael B. Giles[1], Arnulf Jentzen[2,3], and Timo Welti[4]

[1] Mathematical Institute, University of Oxford, Oxford, United Kingdom, e-mail: Mike.Giles@maths.ox.ac.uk

[2] SAM, Department of Mathematics, ETH Zürich, Zürich, Switzerland, e-mail: arnulf.jentzen@sam.math.ethz.ch

[3] Faculty of Mathematics and Computer Science, University of Münster, Münster, Germany, e-mail: ajentzen@uni-muenster.de

[4] SAM, Department of Mathematics, ETH Zürich, Zürich, Switzerland, e-mail: timo.welti@sam.math.ethz.ch

November 13, 2019

Abstract

It is one of the most challenging problems in applied mathematics to approximatively solve high-dimensional partial differential equations (PDEs). In particular, most of the numerical approximation schemes studied in the scientific literature suffer under the curse of dimensionality in the sense that the number of computational operations needed to compute an approximation with an error of size at most $\varepsilon > 0$ grows at least exponentially in the PDE dimension $d \in \mathbb{N}$ or in the reciprocal of ε. Recently, so-called full-history recursive multilevel Picard (MLP) approximation methods have been introduced to tackle the problem of approximately solving high-dimensional PDEs. MLP approximation methods currently are, to the best of our knowledge, the only methods for parabolic semi-linear PDEs with general time horizons and general initial conditions for which there is a rigorous proof that they are indeed able to beat the curse of dimensionality. The main purpose of this work is to investigate MLP approximation methods in more depth, to reveal more clearly how these methods can overcome the curse of dimensionality, and to propose a generalised class of MLP approximation schemes, which covers previously analysed MLP approximation schemes as special cases. In particular, we develop an abstract framework in which this class of generalised MLP approximations can be formulated and analysed and, thereafter, apply this abstract framework to derive a computational complexity result for suitable MLP approximations for semi-linear heat equations. These resulting MLP approximations for semi-linear heat equations essentially are generalisations of previously introduced MLP approximations for semi-linear heat equations.

Keywords: Full-history recursive multilevel Picard approximations, MLP, curse of dimensionality, semi-linear partial differential equations, PDEs, semi-linear heat equations

Contents

1 Introduction

It is one of the most challenging problems in applied mathematics to approximatively solve high-dimensional partial differential equations (PDEs). In particular, most of the numerical approximation schemes studied in the scientific literature, such as finite differences, finite elements, and sparse grids, suffer under the curse of dimensionality (cf. Bellman [13]) in the sense that the number of computational operations needed to compute an approximation with an error of size at most $\varepsilon > 0$ grows at least exponentially in the PDE dimension $d \in \mathbb{N} = \{1, 2, 3, \ldots\}$ or in the reciprocal of ε. Computing such an approximation with reasonably small error thus becomes unfeasible in dimension greater than, say, ten. Therefore, a fundamental goal of current research activities is to propose and analyse numerical methods with the power to beat the curse of dimensionality in such way that the number of computational operations needed to compute an approximation with an error of size at most $\varepsilon > 0$ grows at most polynomially in both the PDE dimension $d \in \mathbb{N}$ and the reciprocal of ε (cf., e.g., Novak & Woźniakowski [88, Chapter 1] and [89, Chapter 9]). In the recent years a number of numerical schemes have been proposed to tackle the problem of approximately solving high-dimensional PDEs, which include deep learning based approximation methods (cf., e.g., [5–7, 10, 11, 16, 24, 29, 37, 39, 42, 44, 45, 53, 58, 59, 61, 66, 71, 79, 80, 84, 91, 92, 98] and the references mentioned therein), branching diffusion approximation methods (cf. [1, 12, 17, 19, 20, 23, 60, 62–64, 85, 93, 99, 101, 102]), approximation methods based on discretising a corresponding backward stochastic differential equation using iterative regression on function Hamel bases (cf., e.g., [4, 14, 15, 21, 25–28, 31–34, 36, 38, 47–52, 65, 76–78, 81–83, 86, 87, 90, 95–97, 100, 103] and the references mentioned therein) or using Wiener chaos expansions (cf. Briand & Labart [22] and Geiss & Labart [46]), and full-history recursive multilevel Picard (MLP) approximation methods (cf. [9, 40, 41, 68–70]). So far, deep learning based approximation methods for PDEs seem to work very well in the case of high-dimensional PDEs judging from a large number of numerical experiments. However, there exist only partial results (cf. [18, 43, 54–56, 67, 73, 75, 94]) and no full error analysis in the scientific literature rigorously justifying their effectiveness in the numerical approximation of high-dimensional PDEs. Moreover, while for branching diffusion methods there is a full error analysis available proving that the curse of dimensionality can be beaten for instances of PDEs with small time horizon and small initial condition, respectively, numerical simulations suggest that such methods fail to work if the time horizon or the initial condition, respectively, are not small anymore. To sum it up, MLP approximation methods currently are, to the best of our knowledge, the only methods for parabolic semi-linear PDEs with general time horizons and general initial conditions for which there is a rigorous proof that they are indeed able to beat the curse of dimensionality.

The main purpose of this work is to investigate MLP methods in more depth, to reveal more clearly how these methods can overcome the curse of dimensionality, and to generalise the MLP scheme proposed in Hutzenthaler et al. [68]. In particular, in the main result of this article, Theorem 2.14 in Subsection 2.6 below, we develop an abstract framework in which suitably generalised MLP approximations can be formulated (cf. (1) in Theorem 1.1 below) and analysed (cf. (i)–(iii) in Theorem 1.1 below) and, thereafter, apply this abstract framework to derive a computational complexity result for suitable MLP approximations for semi-linear heat equations (cf. Corollary 1.2 below). These resulting MLP approximations for semi-linear heat equations essentially are generalisations of the MLP approximations introduced in [68]. To make the reader more familiar with the contributions of this article, we now illustrate in Theorem 1.1 below the findings of the main result of this article, Theorem 2.14 in Subsection 2.6 below, in a simplified situation.

Theorem 1.1. *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, let $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be a separable $\mathbb{R}$-Banach space, let $\mathfrak{z}, \mathfrak{B}, \kappa, C, c \in [1,\infty)$, $\Theta = \cup_{n=1}^{\infty}\mathbb{Z}^n$, $(M_j)_{j\in\mathbb{N}} \subseteq \mathbb{N}$, $y \in \mathcal{Y}$ satisfy $\liminf_{j\to\infty} M_j = \infty$, $\sup_{j\in\mathbb{N}} {}^{M_{j+1}}/_{M_j} \leq \mathfrak{B}$, and $\sup_{j\in\mathbb{N}} {}^{M_j}/_j \leq \kappa$, let $(\mathcal{Z}, \mathscr{Z})$ be a measurable space, let $Z^\theta\colon \Omega \to \mathcal{Z}$, $\theta\in\Theta$, be i.i.d. $\mathscr{F}/\mathscr{Z}$-measurable functions, let $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}}, \|\cdot\|_{\mathcal{H}})$ be a separable $\mathbb{R}$-Hilbert space, let $\mathscr{S}$ be the strong σ-algebra on $L(\mathcal{Y},\mathcal{H})$, let $\psi_k\colon \Omega\to L(\mathcal{Y},\mathcal{H})$, $k\in\mathbb{N}_0 = \{0,1,2,\ldots\}$, be $\mathscr{F}/\mathscr{S}$-measurable functions, assume that $(Z^\theta)_{\theta\in\Theta}$ and $(\psi_k)_{k\in\mathbb{N}_0}$ are independent, let $\Phi_l\colon \mathcal{Y}\times\mathcal{Y}\times\mathcal{Z}\to\mathcal{Y}$, $l\in\mathbb{N}_0$, be $(\mathscr{B}(\mathcal{Y})\otimes\mathscr{B}(\mathcal{Y})\otimes\mathscr{Z})/\mathscr{B}(\mathcal{Y})$-measurable functions, let $Y^\theta_{n,j}\colon \Omega\to\mathcal{Y}$, $\theta\in\Theta$, $j\in\mathbb{N}$, $n\in(\mathbb{N}_0\cup\{-1\})$, satisfy for all $n,j\in\mathbb{N}$, $\theta\in\Theta$ that $Y^\theta_{-1,j} = Y^\theta_{0,j} = 0$ and*

$$Y^\theta_{n,j} = \sum_{l=0}^{n-1} \tfrac{1}{(M_j)^{n-l}} \left[\sum_{i=1}^{(M_j)^{n-l}} \Phi_l\big(Y^{(\theta,l,i)}_{l,j}, Y^{(\theta,-l,i)}_{l-1,j}, Z^{(\theta,l,i)}\big)\right], \tag{1}$$

let $(\mathrm{Cost}_{n,j})_{(n,j)\in(\mathbb{N}_0\cup\{-1\})\times\mathbb{N}} \subseteq [0,\infty)$ satisfy for all $n,j\in\mathbb{N}$ that $\mathrm{Cost}_{-1,j} = \mathrm{Cost}_{0,j} = 0$ and

$$\mathrm{Cost}_{n,j} \leq (M_j)^n\mathfrak{z} + \textstyle\sum_{l=0}^{n-1}\big[(M_j)^{n-l}(\mathrm{Cost}_{l,j} + \mathrm{Cost}_{l-1,j} + 2\mathfrak{z})\big], \tag{2}$$

and assume for all $k \in \mathbb{N}_0$, $n, j \in \mathbb{N}$, $u, v \in \mathcal{Y}$ *that* $\mathbb{E}\big[\|\Phi_k(Y^0_{k,j}, Y^1_{k-1,j}, Z^0)\|_{\mathcal{Y}}\big] < \infty$ *and*

$$\max\big\{\mathbb{E}\big[\|\psi_k(\Phi_0(0,0,Z^0))\|^2_{\mathcal{H}}\big], \mathbb{E}\big[\|\psi_k(y)\|^2_{\mathcal{H}}\big]\big\} \leq \tfrac{C^2}{k!}, \tag{3}$$

$$\mathbb{E}\big[\|\psi_k(\Phi_n(u,v,Z^0))\|^2_{\mathcal{H}}\big] \leq c\,\mathbb{E}\big[\|\psi_{k+1}(u-v)\|^2_{\mathcal{H}}\big], \tag{4}$$

$$\mathbb{E}\Big[\big\|\psi_k\big(y - \textstyle\sum_{l=0}^{n-1}\mathbb{E}\big[\Phi_l\big(Y^0_{l,j}, Y^1_{l-1,j}, Z^0\big)\big]\big)\big\|^2_{\mathcal{H}}\Big] \leq c\,\mathbb{E}\Big[\big\|\psi_{k+1}\big(Y^0_{n-1,j} - y\big)\big\|^2_{\mathcal{H}}\Big]. \tag{5}$$

Then

(i) it holds for all $n \in \mathbb{N}$ *that* $\big(\mathbb{E}\big[\|\psi_0(Y^0_{n,n} - y)\|^2_{\mathcal{H}}\big]\big)^{1/2} \leq C\big[\tfrac{8ce^{\kappa}}{M_n}\big]^{n/2} < \infty$,

(ii) it holds for all $n \in \mathbb{N}$ *that* $\mathrm{Cost}_{n,n} \leq (5M_n)^n \mathfrak{z}$, *and*

(iii) there exists $(N_\varepsilon)_{\varepsilon\in(0,1]} \subseteq \mathbb{N}$ *such that it holds for all* $\varepsilon \in (0,1]$, $\delta \in (0,\infty)$ *that* $\sup_{n\in\{N_\varepsilon, N_\varepsilon+1,\ldots\}}\big(\mathbb{E}\big[\|\psi_0(Y^0_{n,n} - y)\|^2_{\mathcal{H}}\big]\big)^{1/2} \leq \varepsilon$ *and*

$$\mathrm{Cost}_{N_\varepsilon,N_\varepsilon} \leq 5\mathfrak{z}e^{\kappa}C^{2(1+\delta)}\big(1 + \sup_{n\in\mathbb{N}}\big[(M_n)^{-\delta}(40ce^{2\kappa}\mathfrak{B})^{(1+\delta)}\big]^n\big)\varepsilon^{-2(1+\delta)} < \infty. \tag{6}$$

Theorem 1.1 follows directly from the more general result in Corollary 2.15 in Subsection 2.6 below, which, in turn, is a consequence of the main result of this work, Theorem 2.14 in Subsection 2.6 below. In the following we provide some intuitions and further explanations for Theorem 1.1 and illustrate how it is applied in the context of numerically approximating semi-linear heat equations (cf. Corollary 1.2 below and Setting 3.1 in Section 3 below). Theorem 1.1 establishes an upper error bound (cf. (i) in Theorem 1.1) and an upper cost bound (cf. (ii) in Theorem 1.1) for the generalised MLP approximations in (1) as well as an abstract complexity result (cf. (iii) in Theorem 1.1), which states that for an approximation accuracy ε in a suitable root mean square sense the computational cost is essentially of order ε^{-2}. The separable $\mathbb{R}$-Banach space $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ is a set which the exact solution y is an element of and where the generalised MLP approximations $Y^\theta_{n,j}\colon \Omega \to \mathcal{Y}$, $\theta \in \Theta$, $j \in \mathbb{N}$, $n \in (\mathbb{N}_0 \cup \{-1\})$, which are random variables approximating $y \in \mathcal{Y}$ in an appropriate sense, take values in. When $y \in \mathcal{Y}$ is the solution of a suitable semi-linear heat equation (cf. (7) below), elements of $\mathcal{Y}$ are at most polynomially growing functions in $C([0,T] \times \mathbb{R}^d, \mathbb{R})$, where $T \in (0,\infty)$, $d \in \mathbb{N}$ (cf. (138) in Subsection 3.2.1 below). The randomness of the generalised MLP approximations $Y^\theta_{n,j}$, $\theta \in \Theta$, $j \in \mathbb{N}$, $n \in (\mathbb{N}_0 \cup \{-1\})$, stems from the i.i.d. random variables $Z^\theta\colon \Omega \to \mathcal{Z}$, $\theta \in \Theta$, taking values in a measurable space $(\mathcal{Z}, \mathscr{Z})$, which in our example about semi-linear heat equations correspond to standard Brownian motions and on $[0,1]$ uniformly distributed random variables (cf. (8) below). Observe that the generalised MLP approximations in (1) are full-history recursive since each iterate depends on all previous iterates. Together with the random variables Z^θ, $\theta \in \Theta$, the previous iterates enter through the functions $\Phi_l\colon \mathcal{Y} \times \mathcal{Y} \times \mathcal{Z} \to \mathcal{Y}$, $l \in \mathbb{N}_0$, which thus govern the dynamics of the generalised MLP approximations. This recursive dependence, the consequential nesting of the generalised MLP approximations, and the Monte Carlo sums in (1) necessitate a large number of i.i.d. samples indexed by $\theta \in \Theta = \cup_{n=1}^\infty \mathbb{Z}^n$ in order to formulate the generalised MLP approximations. In connection with this note that it holds for every $n \in (\mathbb{N}_0 \cup \{-1\})$, $j \in \mathbb{N}$ that $Y^\theta_{n,j}$, $\theta \in \Theta$, are identically distributed (cf. (v) in Proposition 2.8 below). On the other hand, the parameter $j \in \mathbb{N}$ of the generalised MLP approximations $Y^\theta_{n,j}$, $\theta \in \Theta$, $j \in \mathbb{N}$, $n \in (\mathbb{N}_0 \cup \{-1\})$, specifies the respective element of the sequence of Monte Carlo numbers $(M_j)_{j\in\mathbb{N}} \subseteq \mathbb{N}$ (which are assumed to grow to infinity not faster than linearly) and thereby determines the numbers of Monte Carlo samples to be used in (1). Thus for every $j \in \mathbb{N}$ we can consider the family $(Y^0_{n,j})_{n\in(\mathbb{N}_0\cup\{-1\})}$ of generalised MLP approximations with Monte Carlo sample numbers based on M_j, of which we pick the j-th element $Y^0_{j,j}$ to approximate $y \in \mathcal{Y}$ (cf. (iii) in Theorem 1.1). More precisely, for every $n \in \mathbb{N}$ the approximation error for $Y^0_{n,n}$ is measured in the root mean square sense in a separable $\mathbb{R}$-Hilbert space $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}}, \|\cdot\|_{\mathcal{H}})$, after linearly mapping it from $\mathcal{Y}$ to $\mathcal{H}$ using the possibly random function $\psi_0\colon \Omega \to L(\mathcal{Y},\mathcal{H})$ (cf. (i) and (iii) in Theorem 1.1 above). In our example about semi-linear heat equations, $\mathcal{H}$ is nothing but the set of real numbers $\mathbb{R}$ and ψ_0 is the deterministic evaluation of a function in $\mathcal{Y} \subseteq C([0,T]\times\mathbb{R}^d,\mathbb{R})$ at a deterministic approximation point in $[0,T]\times\mathbb{R}^d$ (cf. (141) in Subsection 3.2.1 below). Conversely, the functions $\psi_k\colon \Omega \to L(\mathcal{Y},\mathcal{H})$, $k \in \mathbb{N}$, correspond in our example to evaluations at suitable random points in $[0,T]\times\mathbb{R}^d$ multiplied with random factors that diminish quickly as $k \in \mathbb{N}$ increases (cf. (141) in Subsection 3.2.1 below). Indeed assumption (3) essentially demands that mean square norms of point evaluations of the functions ψ_k, $k \in \mathbb{N}_0$, diminish at least as fast as the reciprocal of the factorial of their index. Due to this, the functions ψ_k, $k \in \mathbb{N}_0$, can be thought of encoding magnitude in an appropriate randomised sense. Assumption (4) hence essentially requires for every $k \in \mathbb{N}_0$, $n \in \mathbb{N}$ that the k-magnitude of the dynamics function Φ_n can be bounded (up to a constant) by the $(k+1)$-magnitude of the difference of its first two arguments, while assumption (5), roughly speaking, calls for suitable telescopic cancellations (cf. (203) in Subsection 3.2.5 below) such that for every $k \in \mathbb{N}_0$ the k-magnitude of the probabilistically weak approximation error of a given MLP iterate (cf. (53)–(54) in Subsection 2.4 below) can be bounded (up to a constant) by the $(k+1)$-magnitude of the approximation error of the previous MLP iterate. Furthermore, we think of the real number $\mathfrak{z} \in [1,\infty)$ as a parameter associated to the computational cost of one realisation of Z^0 and for every $n \in (\mathbb{N}_0 \cup \{-1\})$, $j \in \mathbb{N}$ we think of the real number $\mathrm{Cost}_{n,j} \in [0,\infty)$ as an upper bound for the computational cost associated to one realisation of $\psi_0(Y^0_{n,j})$ (cf. (2) above). In our application of the abstract framework outlined above, we have that $\mathfrak{z}$ corresponds to the spacial dimension of the considered semi-linear heat equation and we have for every $n \in (\mathbb{N}_0 \cup \{-1\})$, $j \in \mathbb{N}$ that the number $\mathrm{Cost}_{n,j}$ corresponds to an upper bound for the sum of the number of realisations of standard normal random variables and the number of realisations of on $[0,1]$ uniformly distributed random variables used to compute one realisation of $\psi_0(Y^0_{n,j})$ (cf. (213) in Subsection 3.3.1 below).

The abstract framework in Theorem 1.1 can be applied to prove convergence and computational complexity results for MLP approximations in more specific settings. We demonstrate this for the example of MLP approximations for semi-linear heat equations. In particular, Corollary 1.2 below establishes that the MLP approximations in (8), which essentially are generalisations of the MLP approximations introduced in [68], approximate solutions of semi-linear heat equations (7) at the origin without the curse of dimensionality (cf. [68, Theorem 1.1] and [69, Theorem 1.1 and Theorem 4.1]).

Corollary 1.2. *Let $T \in (0,\infty)$, $p \in [0,\infty)$, $\Theta = \cup_{n=1}^\infty \mathbb{Z}^n$, $(M_j)_{j\in\mathbb{N}} \subseteq \mathbb{N}$ satisfy $\sup_{j\in\mathbb{N}} ({}^{M_{j+1}}/_{M_j} + {}^{M_j}/_{j}) < \infty = \liminf_{j\to\infty} M_j$, let $f\colon \mathbb{R}\to\mathbb{R}$ be a Lipschitz continuous function, let $g_d \in C(\mathbb{R}^d,\mathbb{R})$, $d \in \mathbb{N}$, satisfy $\sup_{d\in\mathbb{N}}\sup_{x\in\mathbb{R}^d} {}^{|g_d(x)|}/_{\max\{1,\|x\|^p_{\mathbb{R}^d}\}} < \infty$, for every $d \in \mathbb{N}$ let $y_d \in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ be an at most polynomially growing viscosity solution of*

$$\big(\tfrac{\partial y_d}{\partial t}\big)(t,x) + \tfrac{1}{2}(\Delta_x y_d)(t,x) + f(y_d(t,x)) = 0 \tag{7}$$

with $y_d(T,x) = g_d(x)$ for $(t,x) \in (0,T)\times\mathbb{R}^d$, let $(\Omega,\mathscr{F},\mathbb{P})$ be a probability space, let $U^\theta\colon \Omega \to [0,1]$, $\theta \in \Theta$, be independent on $[0,1]$ uniformly distributed random variables, *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, assume that* $(U^\theta)_{\theta\in\Theta}$ *and* $(W^{d,\theta})_{(d,\theta)\in\mathbb{N}\times\Theta}$ *are independent, let* $Y^{d,\theta}_{n,j}\colon [0,T]\times\mathbb{R}^d\times\Omega\to\mathbb{R}$, $\theta\in\Theta$, $d,j\in\mathbb{N}$, $n\in(\mathbb{N}_0\cup\{-1\})$, *satisfy for all* $n,j,d\in\mathbb{N}$, $\theta\in\Theta$, $t\in[0,T]$, $x\in\mathbb{R}^d$ *that* $Y^{d,\theta}_{-1,j}(t,x)=Y^{d,\theta}_{0,j}(t,x)=0$ *and*

$$Y^{d,\theta}_{n,j}(T-t,x)=\sum_{l=0}^{n-1}\tfrac{t}{(M_j)^{n-l}}\Bigg[\sum_{i=1}^{(M_j)^{n-l}}\Big[f\Big(Y^{d,(\theta,l,i)}_{l,j}\big(T-t+U^{(\theta,l,i)}t,x+W^{d,(\theta,l,i)}_{U^{(\theta,l,i)}t}\big)\Big) \tag{8}$$

$$-\mathbb{1}_{\mathbb{N}}(l)f\Big(Y^{d,(\theta,-l,i)}_{l-1,j}\big(T-t+U^{(\theta,l,i)}t,x+W^{d,(\theta,l,i)}_{U^{(\theta,l,i)}t}\big)\Big)\Big]\Bigg]+\tfrac{1}{(M_j)^n}\Bigg[\sum_{i=1}^{(M_j)^n}g_d\big(x+W^{d,(\theta,0,i)}_t\big)\Bigg],$$

and for every $d,n\in\mathbb{N}$ *let* $\mathrm{Cost}_{d,n}\in\mathbb{N}_0$ *be the number of realisations of standard normal random variables used to compute one realisation of* $Y^{d,0}_{n,n}(0,0)$ *(cf.* (238) *below for a precise definition). Then there exist* $(N_{d,\varepsilon})_{(d,\varepsilon)\in\mathbb{N}\times(0,1]}\subseteq\mathbb{N}$ *and* $(C_\delta)_{\delta\in(0,\infty)}\subseteq(0,\infty)$ *such that it holds for all* $d\in\mathbb{N}$, $\varepsilon\in(0,1]$, $\delta\in(0,\infty)$ *that* $\mathrm{Cost}_{d,N_{d,\varepsilon}}\le C_\delta d^{1+p(1+\delta)}\varepsilon^{-2(1+\delta)}$ *and*

$$\sup\nolimits_{n\in\{N_{d,\varepsilon},N_{d,\varepsilon}+1,\dots\}}\big(\mathbb{E}\big[|Y^{d,0}_{n,n}(0,0)-y_d(0,0)|^2\big]\big)^{1/2}\le\varepsilon. \tag{9}$$

Corollary 1.2 is a direct consequence of Corollary 3.18 in Subsection 3.3.2 below, which itself is a direct consequence of Theorem 3.17 in Subsection 3.3.2 below. Theorem 3.17, in turn, follows from Corollary 2.15 in Subsection 2.6 below and Theorem 1.1 above, respectively, and essentially is a slight generalisation of [68, Theorem 1.1]. More specifically, the MLP approximations in (229) in Theorem 3.17 and in (8) above, respectively, allow for general sequences of Monte Carlo numbers $(M_j)_{j\in\mathbb{N}}\subseteq\mathbb{N}$ satisfying $\sup_{j\in\mathbb{N}}({}^{M_{j+1}}/{}_{M_j}+{}^{M_j}/{}_j)<\infty=\liminf_{j\to\infty}M_j$. This includes, in particular, the special case where $\forall\, j\in\mathbb{N}\colon M_j=j$, which essentially corresponds to the MLP approximations in [68] (cf. [68, (1) in Theorem 1.1]).

This work is structured as follows. Section 2 is devoted to the abstract framework of generalised MLP approximations. In particular, we study several elementary but crucial properties of such approximations in Proposition 2.8 in Subsection 2.3. Moreover, we derive an error analysis for generalised MLP approximations (see Subsection 2.4), which relies on suitably generalised versions of well-known identities involving bias and variance in Hilbert spaces (see Corollary 2.5 and Corollary 2.7 in Subsection 2.2). The error analysis in Subsection 2.4 is subsequently combined with the cost analysis in Subsection 2.5 to establish in Subsection 2.6 a complexity analysis for generalised MLP approximations (see Theorem 2.14 and Corollary 2.15 in Subsection 2.6). Throughout Section 2 the measurability results in Subsection 2.1 are used. In Section 3 we employ the abstract framework for generalised MLP approximations from Section 2 to analyse numerical approximations for semi-linear heat equations. Subsection 3.1 collects several elementary and well-known auxiliary results, which are used in Subsection 3.2 to verify that the main assumptions of the abstract complexity result in Corollary 2.15 are fulfilled in the case of the example setting for numerical approximations for semi-linear heat equations. Finally, in Subsection 3.3 we combine the results from Subsection 3.2 with Corollary 2.15 to obtain a complexity analysis for MLP approximations for semi-linear heat equations (see Proposition 3.16, Theorem 3.17, and Corollary 3.18 in Subsection 3.3).

2 Generalised full-history recursive multilevel Picard (MLP)

In this section we introduce generalised full-history recursive multilevel Picard (MLP) approximations and provide an error analysis (see Subsection 2.4), cost analysis (see Subsection 2.5), and complexity analysis (see Subsection 2.6) for such approximations.

For the formulation of the error analysis for generalised MLP approximations we require random variables which take values in the Banach space $L(\mathcal{Y}, \mathcal{H})$ of continuous linear functions between a separable Banach space $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ and a separable Hilbert space $(\mathcal{H}, \langle \cdot, \cdot \rangle_{\mathcal{H}}, \|\cdot\|_{\mathcal{H}})$ equipped with the strong σ-algebra. Let us recall that the strong σ-algebra on $L(\mathcal{Y}, \mathcal{H})$ is nothing but the trace σ-algebra of the product σ-algebra on the set $\mathcal{H}^{\mathcal{Y}}$ of all functions from $\mathcal{Y}$ to $\mathcal{H}$. Having this in mind, Subsection 2.1 collects three elementary measurability results (see Lemmas 2.1–2.2 and Corollary 2.3) about functions whose domains or codomains involve a set of functions equipped with the trace σ-algebra of the product σ-algebra.

In Subsection 2.2 we first recall the elementary and well-known bias–variance decomposition of the mean square error for random variables that take values in a separable Hilbert space (see Lemma 2.4). Thereafter, we present in Corollary 2.5 a generalised bias–variance decomposition, where the mean square error, the bias, and the variance are measured in a certain randomised sense. Analogously, we also recall the elementary and well-known result that the mean square norm of the sum of independent zero mean random variables in a separable Hilbert space is equal to the sum of the individual mean square norms (see Lemma 2.6) and prove a randomised generalisation thereof (see Corollary 2.7). This generalisation as well as the generalised bias–variance decomposition in Corollary 2.5 are used in our error analysis for generalised MLP approximations (see (47), (48), and (51) in the proof of Proposition 2.9 below).

Subsequently, Proposition 2.8 in Subsection 2.3 establishes several elementary but crucial properties of generalised MLP approximations, which are a consequence of their definition.

Subsection 2.4 is devoted to the error analysis for generalised MLP approximations. Proposition 2.9 specifies the most general hypotheses in this article (see (38)–(41) below) under which we prove an error estimate for generalised MLP approximations. The upper bound for the error in Proposition 2.9 (see (42) below) can be much shortened by choosing a natural number $\mathfrak{M} \in \mathbb{N}$ such that for every $n \in \mathbb{N}$, $l \in \{0, 1, \ldots, n-1\}$ the Monte Carlo sample number $M_{n,l} \in \mathbb{N}$ in the generalised MLP approximations (see (38) in Proposition 2.9) is equal to $\mathfrak{M}^{n-l}$, which is the assertion of Corollary 2.10 (see (63) below).

The subject of Subsection 2.5 is the cost analysis for generalised MLP approximations. The cost estimate in Proposition 2.11 follows from an application of the discrete Gronwall-type inequality in Agarwal [2, Theorem 4.1.1]. The second cost estimate in Subsection 2.5 (see Corollary 2.13), in turn, is a consequence of Proposition 2.11 and the elementary and well-known estimate in Lemma 2.12.

In Subsection 2.6 the error analysis from Subsection 2.4 and the cost analysis from Subsection 2.5 are combined to derive a complexity analysis for generalised MLP approximations. More precisely, the main result of this article, Theorem 2.14, relates the error estimate in Corollary 2.10 to the cost estimate in Corollary 2.13 in order to arrive at a complexity estimate (see (90) below). The subsequent result, Corollary 2.15, is obtained by replacing assumption (87) in Theorem 2.14 by the simpler assumption (102) and choosing for every $k \in \mathbb{N}_0$ the coefficient $\mathfrak{c}_k \in (0, \infty)$ in Theorem 2.14 to be equal to $k!$. Finally, the elementary result in Lemma 2.16 shows that a strictly increasing and at most linearly growing sequence of natural numbers automatically fulfils the hypotheses on the sequence $(M_j)_{j\in\mathbb{N}} \subseteq \mathbb{N}$ in Corollary 2.15.

2.1 Measurability involving the strong σsigma-algebra

Lemma 2.1. *Let $\mathcal{E}$ be a set, let $(\mathcal{F}, \mathscr{F})$ and $(\mathcal{G}, \mathscr{G})$ be measurable spaces, let $\mathcal{S} \subseteq \mathcal{F}^{\mathcal{E}}$, let $\mathscr{S} = \sigma_{\mathcal{S}}\big(\{\{\varphi \in \mathcal{S}\colon \varphi(e) \in \mathcal{A}\} \subseteq \mathcal{S}\colon e \in \mathcal{E},\, \mathcal{A} \in \mathscr{F}\}\big)$, and let $\psi\colon \mathcal{G} \to \mathcal{S}$ be a function. Then it holds that ψ is $\mathscr{G}/\mathscr{S}$-measurable if and only if it holds for all $e \in \mathcal{E}$ that $\mathcal{G} \ni \omega \mapsto [\psi(\omega)](e) \in \mathcal{F}$ is $\mathscr{G}/\mathscr{F}$-measurable.*

Proof of Lemma 2.1. Throughout this proof let $P_e\colon \mathcal{S} \to \mathcal{F}$, $e \in \mathcal{E}$, satisfy for all $e \in \mathcal{E}$, $\varphi \in \mathcal{S}$ that $P_e(\varphi) = \varphi(e)$. Observe that it holds that

$$\mathscr{S} = \{(\mathcal{A} \cap \mathcal{S}) \subseteq \mathcal{S}\colon \mathcal{A} \in (\otimes_{e\in\mathcal{E}}\mathscr{F})\} = \sigma_{\mathcal{S}}\big((P_e)_{e\in\mathcal{E}}\big). \tag{10}$$

This in ensures for all $e \in \mathcal{E}$ that $P_e\colon \mathcal{S} \to \mathcal{F}$ is an $\mathscr{S}/\mathscr{F}$-measurable function. Equation (10) hence shows that $\psi\colon \mathcal{G} \to \mathcal{S}$ is $\mathscr{G}/\mathscr{S}$-measurable if and only if it holds for all $e \in \mathcal{E}$ that $P_e \circ \psi\colon \mathcal{G} \to \mathcal{F}$ is $\mathscr{G}/\mathscr{F}$-measurable. The proof of Lemma 2.1 is thus complete. ☐

Lemma 2.2. *Let $(\mathcal{E}, d_{\mathcal{E}})$ be a separable metric space, let $(\mathcal{F}, d_{\mathcal{F}})$ be a metric space, let $\mathcal{S} \subseteq C(\mathcal{E}, \mathcal{F})$, and let $\mathscr{S} = \sigma_{\mathcal{S}}\big(\{\{\varphi \in \mathcal{S}\colon \varphi(e) \in \mathcal{A}\} \subseteq \mathcal{S}\colon e \in \mathcal{E},\, \mathcal{A} \in \mathscr{B}(\mathcal{F})\}\big)$. Then it holds that $\mathcal{S} \times \mathcal{E} \ni (\varphi, e) \mapsto \varphi(e) \in \mathcal{F}$ is an $(\mathscr{S} \otimes \mathscr{B}(\mathcal{E}))/\mathscr{B}(\mathcal{F})$-measurable function.*

Proof of Lemma 2.2. Throughout this proof let $f\colon \mathcal{S}\times\mathcal{E} \to \mathcal{F}$ satisfy for all $\varphi \in \mathcal{S}$, $e \in \mathcal{E}$ that $f(\varphi, e) = \varphi(e)$. Note that it holds for all $\varphi \in \mathcal{S}$ that

$$(\mathcal{E} \ni e \mapsto f(\varphi, e) \in \mathcal{F}) = \varphi \in \mathcal{S} \subseteq C(\mathcal{E}, \mathcal{F}). \tag{11}$$

In addition, observe that it holds for all $e \in \mathcal{E}$, $\mathcal{A} \in \mathscr{B}(\mathcal{F})$ that

$$\{\varphi \in \mathcal{S}\colon f(\varphi, e) \in \mathcal{A}\} = \{\varphi \in \mathcal{S}\colon \varphi(e) \in \mathcal{A}\} \in \mathscr{S}. \tag{12}$$

This proves for all $e \in \mathcal{E}$ that $\mathcal{S} \ni \varphi \mapsto f(\varphi, e) \in \mathcal{F}$ is an $\mathscr{S}/\mathscr{B}(\mathcal{F})$-measurable function. Combining this and (11) with Aliprantis & Border [3, Lemma 4.51] (also see, e.g., [6, Lemma 2.4]) establishes that $f\colon \mathcal{S}\times\mathcal{E} \to \mathcal{F}$ is an $(\mathscr{S}\otimes\mathscr{B}(\mathcal{E}))/\mathscr{B}(\mathcal{F})$-measurable function. The proof of Lemma 2.2 is thus complete. ☐

Corollary 2.3. *Let $(\Omega, \mathscr{F})$ be a measurable space, let $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$ be a separable normed $\mathbb{R}$-vector space, let $Y\colon \Omega \to \mathcal{V}$ be an $\mathscr{F}/\mathscr{B}(\mathcal{V})$-measurable function, let $(\mathcal{W}, \|\cdot\|_{\mathcal{W}})$ be a normed $\mathbb{R}$-vector space, let $\mathscr{S} = \sigma_{L(\mathcal{V},\mathcal{W})}\big(\{\{\varphi \in L(\mathcal{V}, \mathcal{W})\colon \varphi(v) \in \mathcal{B}\} \subseteq L(\mathcal{V}, \mathcal{W})\colon v \in \mathcal{V},\, \mathcal{B} \in \mathscr{B}(\mathcal{W})\}\big)$, and let $\psi\colon \Omega \to L(\mathcal{V}, \mathcal{W})$ be an $\mathscr{F}/\mathscr{S}$-measurable function. Then*

(i) it holds that $L(\mathcal{V}, \mathcal{W})\times\mathcal{V} \ni (\varphi, v) \mapsto \varphi(v) \in \mathcal{W}$ is an $(\mathscr{S}\otimes\mathscr{B}(\mathcal{V}))/\mathscr{B}(\mathcal{W})$-measurable function and

(ii) it holds that $\psi(Y) = (\Omega \ni \omega \mapsto [\psi(\omega)](Y(\omega)) \in \mathcal{W})$ is an $\mathscr{F}/\mathscr{B}(\mathcal{W})$-measurable function.

Proof of Corollary 2.3. Observe that Lemma 2.2 (with $\mathcal{E} = \mathcal{V}$, $\mathcal{F} = \mathcal{W}$, $\mathcal{S} = L(\mathcal{V}, \mathcal{W})$ in the notation of Lemma 2.2) implies (i). In addition, the fact that $\Omega \ni \omega \mapsto (\psi(\omega), Y(\omega)) \in L(\mathcal{V}, \mathcal{W}) \times \mathcal{V}$ is an $\mathscr{F}/(\mathscr{S} \otimes \mathscr{B}(\mathcal{V}))$-measurable function and (i) show (ii). The proof of Corollary 2.3 is thus complete. ☐

2.2 Identities involving bias and variance in Hilbert spaces

2.2.1 Bias–variance decomposition

Lemma 2.4. *Let* $(\Omega, \mathscr{F}, \mathbb{P})$ *be a probability space, let* $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}}, \|\cdot\|_{\mathcal{H}})$ *be a separable* $\mathbb{R}$*-Hilbert space, let* $h \in \mathcal{H}$*, and let* $X\colon \Omega \to \mathcal{H}$ *be an* $\mathscr{F}/\mathscr{B}(\mathcal{H})$*-measurable function which satisfies* $\mathbb{E}[\|X\|_{\mathcal{H}}] < \infty$*. Then*

$$\mathbb{E}\big[\|X - h\|_{\mathcal{H}}^2\big] = \mathbb{E}\big[\|X - \mathbb{E}[X]\|_{\mathcal{H}}^2\big] + \|\mathbb{E}[X] - h\|_{\mathcal{H}}^2. \tag{13}$$

Proof of Lemma 2.4. Note that the Cauchy–Schwarz inequality implies that

$$\mathbb{E}\big[|\langle X - \mathbb{E}[X], \mathbb{E}[X] - h\rangle_{\mathcal{H}}|\big] \le \mathbb{E}\big[\|X - \mathbb{E}[X]\|_{\mathcal{H}}\big]\|\mathbb{E}[X] - h\|_{\mathcal{H}} < \infty. \tag{14}$$

This ensures that

$$\begin{aligned}\mathbb{E}\big[\|X - h\|_{\mathcal{H}}^2\big] &= \mathbb{E}\big[\|X - \mathbb{E}[X] + \mathbb{E}[X] - h\|_{\mathcal{H}}^2\big] \\ &= \mathbb{E}\big[\|X - \mathbb{E}[X]\|_{\mathcal{H}}^2 + \|\mathbb{E}[X] - h\|_{\mathcal{H}}^2 + 2\langle X - \mathbb{E}[X], \mathbb{E}[X] - h\rangle_{\mathcal{H}}\big] \\ &= \mathbb{E}\big[\|X - \mathbb{E}[X]\|_{\mathcal{H}}^2\big] + \|\mathbb{E}[X] - h\|_{\mathcal{H}}^2 + 2\langle \mathbb{E}[X] - \mathbb{E}[X], \mathbb{E}[X] - h\rangle_{\mathcal{H}} \\ &= \mathbb{E}\big[\|X - \mathbb{E}[X]\|_{\mathcal{H}}^2\big] + \|\mathbb{E}[X] - h\|_{\mathcal{H}}^2.\end{aligned} \tag{15}$$

The proof of Lemma 2.4 is thus complete. □

2.2.2 Generalised bias–variance decomposition

Corollary 2.5. *Let* $(\Omega, \mathscr{F}, \mathbb{P})$ *be a probability space, let* $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ *be a separable* $\mathbb{R}$*-Banach space, let* $y \in \mathcal{Y}$*, let* $Y\colon \Omega \to \mathcal{Y}$ *be an* $\mathscr{F}/\mathscr{B}(\mathcal{Y})$*-measurable function which satisfies* $\mathbb{E}[\|Y\|_{\mathcal{Y}}] < \infty$*, let* $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}}, \|\cdot\|_{\mathcal{H}})$ *be a separable* $\mathbb{R}$*-Hilbert space, let* $\mathscr{S} = \sigma_{L(\mathcal{Y},\mathcal{H})}\big(\big\{\{\varphi \in L(\mathcal{Y},\mathcal{H})\colon \varphi(x) \in \mathcal{B}\} \subseteq L(\mathcal{Y},\mathcal{H})\colon x \in \mathcal{Y}, \mathcal{B} \in \mathscr{B}(\mathcal{H})\big\}\big)$*, let* $\psi\colon \Omega \to L(\mathcal{Y},\mathcal{H})$ *be an* $\mathscr{F}/\mathscr{S}$*-measurable function, and assume that* Y *and* ψ *are independent. Then*

$$\mathbb{E}\big[\|\psi(Y - y)\|_{\mathcal{H}}^2\big] = \mathbb{E}\big[\|\psi(Y - \mathbb{E}[Y])\|_{\mathcal{H}}^2\big] + \mathbb{E}\big[\|\psi(\mathbb{E}[Y] - y)\|_{\mathcal{H}}^2\big] \tag{16}$$

(cf. (ii) in Corollary 2.3).

Proof of Corollary 2.5. The fact that it holds for all $x \in \mathcal{Y}$ that $L(\mathcal{Y},\mathcal{H}) \times \Omega \ni (\varphi,\omega) \mapsto (\varphi, Y(\omega) - x) \in L(\mathcal{Y},\mathcal{H}) \times \mathcal{Y}$ is an $(\mathscr{S} \otimes \sigma_\Omega(Y))/(\mathscr{S} \otimes \mathscr{B}(\mathcal{Y}))$-measurable function and the fact that $L(\mathcal{Y},\mathcal{H}) \times \mathcal{Y} \ni (\varphi, x) \mapsto \varphi(x) \in \mathcal{H}$ is an $(\mathscr{S} \otimes \mathscr{B}(\mathcal{Y}))/\mathscr{B}(\mathcal{H})$-measurable function (cf. (i) in Corollary 2.3) imply for all $x \in \mathcal{Y}$ that

$$L(\mathcal{Y},\mathcal{H}) \times \Omega \ni (\varphi,\omega) \mapsto \varphi(Y(\omega) - x) \in \mathcal{H} \tag{17}$$

is an $(\mathscr{S} \otimes \sigma_\Omega(Y))/\mathscr{B}(\mathcal{H})$-measurable function. Lemma 2.2 in Hutzenthaler et al. [68] (with $\mathcal{G} = \sigma_\Omega(Y)$, $(S, \mathcal{S}) = (L(\mathcal{Y},\mathcal{H}), \mathscr{S})$, $U = (L(\mathcal{Y},\mathcal{H}) \times \Omega \ni (\varphi,\omega) \mapsto \|\varphi(Y(\omega) - y)\|_{\mathcal{H}}^2 \in [0,\infty))$, $Y = \psi$ in the notation of [68, Lemma 2.2]) and Lemma 2.4 hence yield that

$$\begin{aligned}&\mathbb{E}\big[\|\psi(Y - y)\|_{\mathcal{H}}^2\big] = \int_{L(\mathcal{Y},\mathcal{H})} \mathbb{E}\big[\|\varphi(Y - y)\|_{\mathcal{H}}^2\big]\,(\psi(\mathbb{P})_{\mathscr{S}})(\mathrm{d}\varphi) \\ &= \int_{L(\mathcal{Y},\mathcal{H})} \mathbb{E}\big[\|\varphi(Y) - \varphi(y)\|_{\mathcal{H}}^2\big]\,(\psi(\mathbb{P})_{\mathscr{S}})(\mathrm{d}\varphi) \\ &= \int_{L(\mathcal{Y},\mathcal{H})} \mathbb{E}\big[\|\varphi(Y) - \mathbb{E}[\varphi(Y)]\|_{\mathcal{H}}^2\big] + \|\mathbb{E}[\varphi(Y)] - \varphi(y)\|_{\mathcal{H}}^2\,(\psi(\mathbb{P})_{\mathscr{S}})(\mathrm{d}\varphi).\end{aligned} \tag{18}$$

This, the fact that $L(\mathcal{Y},\mathcal{H}) \ni \varphi \mapsto \varphi(\mathbb{E}[Y]-y) \in \mathcal{H}$ is an $\mathscr{S}/\mathscr{B}(\mathcal{H})$-measurable function (cf. Lemma 2.1), (17), and again [68, Lemma 2.2] (with $\mathcal{G} = \sigma_\Omega(Y)$, $(S,\mathcal{S}) = (L(\mathcal{Y},\mathcal{H}),\mathscr{S})$, $U = (L(\mathcal{Y},\mathcal{H}) \times \Omega \ni (\varphi,\omega) \mapsto \|\varphi(Y(\omega)-\mathbb{E}[Y])\|_{\mathcal{H}}^2 \in [0,\infty))$, $Y = \psi$ in the notation of [68, Lemma 2.2]) establish that

$$\begin{aligned}\mathbb{E}\big[\|\psi(Y-y)\|_{\mathcal{H}}^2\big] &= \int_{L(\mathcal{Y},\mathcal{H})} \mathbb{E}\big[\|\varphi(Y-\mathbb{E}[Y])\|_{\mathcal{H}}^2\big]\,(\psi(\mathbb{P})_{\mathscr{S}})(\mathrm{d}\varphi) \\ &\quad + \int_{L(\mathcal{Y},\mathcal{H})} \|\varphi(\mathbb{E}[Y]-y)\|_{\mathcal{H}}^2\,(\psi(\mathbb{P})_{\mathscr{S}})(\mathrm{d}\varphi) \\ &= \mathbb{E}\big[\|\psi(Y-\mathbb{E}[Y])\|_{\mathcal{H}}^2\big] + \mathbb{E}\big[\|\psi(\mathbb{E}[Y]-y)\|_{\mathcal{H}}^2\big].\end{aligned} \tag{19}$$

The proof of Corollary 2.5 is thus complete. □

2.2.3 Variance identity

Lemma 2.6. *Let $n \in \mathbb{N}$, let $(\Omega,\mathscr{F},\mathbb{P})$ be a probability space, let $(\mathcal{H},\langle\cdot,\cdot\rangle_{\mathcal{H}},\|\cdot\|_{\mathcal{H}})$ be a separable $\mathbb{R}$-Hilbert space, and let $X_1,X_2,\ldots,X_n\colon\Omega\to\mathcal{H}$ be independent $\mathscr{F}/\mathscr{B}(\mathcal{H})$-measurable functions which satisfy for all $i\in\{1,2,\ldots,n\}$ that $\mathbb{E}[\|X_i\|_{\mathcal{H}}]<\infty$. Then*

$$\mathbb{E}\left[\left\|\sum_{i=1}^n (X_i - \mathbb{E}[X_i])\right\|_{\mathcal{H}}^2\right] = \sum_{i=1}^n \mathbb{E}\big[\|X_i-\mathbb{E}[X_i]\|_{\mathcal{H}}^2\big]. \tag{20}$$

Proof of Lemma 2.6. Observe that the Cauchy–Schwarz inequality, the fact that $X_1, X_2, \ldots, X_n$ are independent, and the fact that it holds for all independent random variables $Y,Z\colon\Omega\to\mathbb{R}$ with $\mathbb{E}[|Y|+|Z|]<\infty$ that $\mathbb{E}[|YZ|]<\infty$ and $\mathbb{E}[YZ]=\mathbb{E}[Y]\mathbb{E}[Z]$ (cf., e.g., Klenke [74, Theorem 5.4]) demonstrate for all $i,j\in\{1,2,\ldots,n\}$ with $i\neq j$ that

$$\begin{aligned}\mathbb{E}\big[|\langle X_i-\mathbb{E}[X_i],X_j-\mathbb{E}[X_j]\rangle_{\mathcal{H}}|\big] &\le \mathbb{E}\big[\|X_i-\mathbb{E}[X_i]\|_{\mathcal{H}}\|X_j-\mathbb{E}[X_j]\|_{\mathcal{H}}\big] \\ &= \mathbb{E}\big[\|X_i-\mathbb{E}[X_i]\|_{\mathcal{H}}\big]\mathbb{E}\big[\|X_j-\mathbb{E}[X_j]\|_{\mathcal{H}}\big] < \infty.\end{aligned} \tag{21}$$

Moreover, the fact that $X_1,X_2,\ldots,X_n$ are independent ensures for all $i,j\in\{1,2,\ldots,n\}$ with $i\neq j$ that

$$(X_i,X_j)(\mathbb{P})_{\mathscr{B}(\mathcal{H})\otimes\mathscr{B}(\mathcal{H})} = [(X_i)(\mathbb{P})_{\mathscr{B}(\mathcal{H})}]\otimes[(X_j)(\mathbb{P})_{\mathscr{B}(\mathcal{H})}]. \tag{22}$$

Fubini's theorem and (21) hence show for all $i,j\in\{1,2,\ldots,n\}$ with $i\neq j$ that

$$\begin{aligned}&\mathbb{E}\big[\langle X_i-\mathbb{E}[X_i],X_j-\mathbb{E}[X_j]\rangle_{\mathcal{H}}\big] \\ &= \int_{\mathcal{H}\times\mathcal{H}} \langle x-\mathbb{E}[X_i],y-\mathbb{E}[X_j]\rangle_{\mathcal{H}}\,\big((X_i,X_j)(\mathbb{P})_{\mathscr{B}(\mathcal{H})\otimes\mathscr{B}(\mathcal{H})}\big)(\mathrm{d}x,\mathrm{d}y) \\ &= \int_{\mathcal{H}\times\mathcal{H}} \langle x-\mathbb{E}[X_i],y-\mathbb{E}[X_j]\rangle_{\mathcal{H}}\,\big([(X_i)(\mathbb{P})_{\mathscr{B}(\mathcal{H})}]\otimes[(X_j)(\mathbb{P})_{\mathscr{B}(\mathcal{H})}]\big)(\mathrm{d}x,\mathrm{d}y) \\ &= \int_{\mathcal{H}}\int_{\mathcal{H}} \langle x-\mathbb{E}[X_i],y-\mathbb{E}[X_j]\rangle_{\mathcal{H}}\,\big((X_i)(\mathbb{P})_{\mathscr{B}(\mathcal{H})}\big)(\mathrm{d}x)\big((X_j)(\mathbb{P})_{\mathscr{B}(\mathcal{H})}\big)(\mathrm{d}y) \\ &= \int_{\mathcal{H}} \mathbb{E}\big[\langle X_i-\mathbb{E}[X_i],y-\mathbb{E}[X_j]\rangle_{\mathcal{H}}\big]\,\big((X_j)(\mathbb{P})_{\mathscr{B}(\mathcal{H})}\big)(\mathrm{d}y) \\ &= \int_{\mathcal{H}} \langle\mathbb{E}[X_i]-\mathbb{E}[X_i],y-\mathbb{E}[X_j]\rangle_{\mathcal{H}}\,\big((X_j)(\mathbb{P})_{\mathscr{B}(\mathcal{H})}\big)(\mathrm{d}y) \\ &= \langle\mathbb{E}[X_i]-\mathbb{E}[X_i],\mathbb{E}[X_j]-\mathbb{E}[X_j]\rangle_{\mathcal{H}} = 0.\end{aligned} \tag{23}$$

This and again (21) prove that

$$\begin{aligned}
&\mathbb{E}\Bigg[\Bigg\|\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i])\Bigg\|_{\mathcal{H}}^2\Bigg] = \mathbb{E}\Bigg[\Bigg\langle \sum_{i=1}^{n}(X_i - \mathbb{E}[X_i]), \sum_{j=1}^{n}(X_j - \mathbb{E}[X_j])\Bigg\rangle_{\mathcal{H}}\Bigg] \\
&= \mathbb{E}\Bigg[\sum_{i,j=1}^{n}\langle X_i - \mathbb{E}[X_i], X_j - \mathbb{E}[X_j]\rangle_{\mathcal{H}}\Bigg] \\
&= \Bigg[\sum_{i=1}^{n}\mathbb{E}\big[\langle X_i - \mathbb{E}[X_i], X_i - \mathbb{E}[X_i]\rangle_{\mathcal{H}}\big]\Bigg] + \sum_{i,j=1,\, i\neq j}^{n}\mathbb{E}\big[\langle X_i - \mathbb{E}[X_i], X_j - \mathbb{E}[X_j]\rangle_{\mathcal{H}}\big] \\
&= \sum_{i=1}^{n}\mathbb{E}\big[\|X_i - \mathbb{E}[X_i]\|_{\mathcal{H}}^2\big].
\end{aligned} \tag{24}$$

The proof of Lemma 2.6 is thus complete. □

2.2.4 Generalised variance identity

Corollary 2.7. *Let $n \in \mathbb{N}$, let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, let $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be a separable $\mathbb{R}$-Banach space, let $Y_1, Y_2, \ldots, Y_n \colon \Omega \to \mathcal{Y}$ be independent $\mathscr{F}/\mathscr{B}(\mathcal{Y})$-measurable functions which satisfy for all $i \in \{1, 2, \ldots, n\}$ that $\mathbb{E}[\|Y_i\|_{\mathcal{Y}}] < \infty$, let $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}}, \|\cdot\|_{\mathcal{H}})$ be a separable $\mathbb{R}$-Hilbert space, let $\mathscr{S} = \sigma_{L(\mathcal{Y},\mathcal{H})}\big(\big\{\{\varphi \in L(\mathcal{Y}, \mathcal{H}) \colon \varphi(y) \in \mathcal{B}\} \subseteq L(\mathcal{Y}, \mathcal{H}) \colon y \in \mathcal{Y},\, \mathcal{B} \in \mathscr{B}(\mathcal{H})\big\}\big)$, let $\psi \colon \Omega \to L(\mathcal{Y}, \mathcal{H})$ be an $\mathscr{F}/\mathscr{S}$-measurable function, and assume that $(Y_i)_{i\in\{1,2,\ldots,n\}}$ and ψ are independent. Then*

$$\mathbb{E}\Bigg[\Bigg\|\sum_{i=1}^{n}\psi(Y_i - \mathbb{E}[Y_i])\Bigg\|_{\mathcal{H}}^2\Bigg] = \sum_{i=1}^{n}\mathbb{E}\big[\|\psi(Y_i - \mathbb{E}[Y_i])\|_{\mathcal{H}}^2\big] \tag{25}$$

(cf. (ii) in Corollary 2.3).

Proof of Corollary 2.7. The fact that it holds for all $i \in \{1, 2, \ldots, n\}$ that $L(\mathcal{Y}, \mathcal{H}) \times \Omega \ni (\varphi, \omega) \mapsto (\varphi, Y_i(\omega) - \mathbb{E}[Y_i]) \in L(\mathcal{Y}, \mathcal{H}) \times \mathcal{Y}$ is an $(\mathscr{S} \otimes \sigma_\Omega((Y_j)_{j\in\{1,2,\ldots,n\}}))/(\mathscr{S} \otimes \mathscr{B}(\mathcal{Y}))$-measurable function and the fact that $L(\mathcal{Y}, \mathcal{H}) \times \mathcal{Y} \ni (\varphi, y) \mapsto \varphi(y) \in \mathcal{H}$ is an $(\mathscr{S} \otimes \mathscr{B}(\mathcal{Y}))/\mathscr{B}(\mathcal{H})$-measurable function (cf. (i) in Corollary 2.3) ensure for all $i \in \{1, 2, \ldots, n\}$ that

$$L(\mathcal{Y}, \mathcal{H}) \times \Omega \ni (\varphi, \omega) \mapsto \varphi(Y_i(\omega) - \mathbb{E}[Y_i]) \in \mathcal{H} \tag{26}$$

is an $(\mathscr{S} \otimes \sigma_\Omega((Y_j)_{j\in\{1,2,\ldots,n\}}))/\mathscr{B}(\mathcal{H})$-measurable function. This and [68, Lemma 2.2] (with $\mathcal{G} = \sigma_\Omega((Y_i)_{i\in\{1,2,\ldots,n\}})$, $(S, \mathcal{S}) = (L(\mathcal{Y}, \mathcal{H}), \mathscr{S})$, $U = (L(\mathcal{Y}, \mathcal{H}) \times \Omega \ni (\varphi, \omega) \mapsto \|\sum_{i=1}^{n}\varphi(Y_i(\omega) - \mathbb{E}[Y_i])\|_{\mathcal{H}}^2 \in [0, \infty))$, $Y = \psi$ in the notation of [68, Lemma 2.2]) establish that

$$\begin{aligned}
\mathbb{E}\Bigg[\Bigg\|\sum_{i=1}^{n}\psi(Y_i - \mathbb{E}[Y_i])\Bigg\|_{\mathcal{H}}^2\Bigg] &= \int_{L(\mathcal{Y},\mathcal{H})}\mathbb{E}\Big[\big\|\textstyle\sum_{i=1}^{n}\varphi(Y_i - \mathbb{E}[Y_i])\big\|_{\mathcal{H}}^2\Big]\,(\psi(\mathbb{P})_{\mathscr{S}})(\mathrm{d}\varphi) \\
&= \int_{L(\mathcal{Y},\mathcal{H})}\mathbb{E}\Big[\big\|\textstyle\sum_{i=1}^{n}(\varphi(Y_i) - \mathbb{E}[\varphi(Y_i)])\big\|_{\mathcal{H}}^2\Big]\,(\psi(\mathbb{P})_{\mathscr{S}})(\mathrm{d}\varphi).
\end{aligned} \tag{27}$$

Lemma 2.6, (26), and [68, Lemma 2.2] (with $\mathcal{G} = \sigma_\Omega((Y_j)_{j\in\{1,2,\ldots,n\}})$, $(S, \mathcal{S}) = (L(\mathcal{Y}, \mathcal{H}), \mathscr{S})$, $U = (L(\mathcal{Y}, \mathcal{H}) \times \Omega \ni (\varphi, \omega) \mapsto \|\varphi(Y_i(\omega) - \mathbb{E}[Y_i])\|_{\mathcal{H}}^2 \in [0, \infty))$, $Y = \psi$ for $i \in$

$\{1, 2, \ldots, n\}$ in the notation of [68, Lemma 2.2]) hence show that

$$\begin{aligned}
\mathbb{E}\left[\left\|\sum_{i=1}^{n} \psi(Y_i - \mathbb{E}[Y_i])\right\|_{\mathcal{H}}^2\right] &= \int_{L(\mathcal{Y},\mathcal{H})} \sum_{i=1}^{n} \mathbb{E}\big[\|\varphi(Y_i) - \mathbb{E}[\varphi(Y_i)]\|_{\mathcal{H}}^2\big] \, (\psi(\mathbb{P})_{\mathscr{S}})(\mathrm{d}\varphi) \\
&= \sum_{i=1}^{n} \int_{L(\mathcal{Y},\mathcal{H})} \mathbb{E}\big[\|\varphi(Y_i - \mathbb{E}[Y_i])\|_{\mathcal{H}}^2\big] \, (\psi(\mathbb{P})_{\mathscr{S}})(\mathrm{d}\varphi) \\
&= \sum_{i=1}^{n} \mathbb{E}\big[\|\psi(Y_i - \mathbb{E}[Y_i])\|_{\mathcal{H}}^2\big].
\end{aligned} \tag{28}$$

The proof of Corollary 2.7 is thus complete. □

2.3 Properties of generalised MLP approximations

Proposition 2.8. *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, let $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be a separable $\mathbb{R}$-Banach space, let $\Theta = \cup_{n=1}^{\infty} \mathbb{Z}^n$, $(M_{n,l})_{(n,l)\in\mathbb{N}\times\mathbb{N}_0} \subseteq \mathbb{N}$, let $(\mathcal{Z}, \mathscr{Z})$ be a measurable space, let $Z^\theta\colon \Omega \to \mathcal{Z}$, $\theta \in \Theta$, be i.i.d. $\mathscr{F}/\mathscr{Z}$-measurable functions, let $\Phi_l\colon \mathcal{Y} \times \mathcal{Y} \times \mathcal{Z} \to \mathcal{Y}$, $l \in \mathbb{N}_0$, be $(\mathscr{B}(\mathcal{Y}) \otimes \mathscr{B}(\mathcal{Y}) \otimes \mathscr{Z})/\mathscr{B}(\mathcal{Y})$-measurable functions, let $Y_{-1}^\theta\colon \Omega \to \mathcal{Y}$, $\theta \in \Theta$, be i.i.d. $\mathscr{F}/\mathscr{B}(\mathcal{Y})$-measurable functions, let $Y_0^\theta\colon \Omega \to \mathcal{Y}$, $\theta \in \Theta$, be i.i.d. $\mathscr{F}/\mathscr{B}(\mathcal{Y})$-measurable functions, assume that $(Y_{-1}^\theta)_{\theta\in\Theta}$, $(Y_0^\theta)_{\theta\in\Theta}$, and $(Z^\theta)_{\theta\in\Theta}$ are independent, and let $Y_n^\theta\colon \Omega \to \mathcal{Y}$, $\theta \in \Theta$, $n \in \mathbb{N}$, satisfy for all $n \in \mathbb{N}$, $\theta \in \Theta$ that*

$$Y_n^\theta = \sum_{l=0}^{n-1} \tfrac{1}{M_{n,l}} \left[\sum_{i=1}^{M_{n,l}} \Phi_l\big(Y_l^{(\theta,l,i)}, Y_{l-1}^{(\theta,-l,i)}, Z^{(\theta,l,i)}\big)\right]. \tag{29}$$

Then

(i) it holds for all $n \in (\mathbb{N}_0 \cup \{-1\})$, $\theta \in \Theta$ that $Y_n^\theta\colon \Omega \to \mathcal{Y}$ is an $\mathscr{F}/\mathscr{B}(\mathcal{Y})$-measurable function,

(ii) it holds for all $n \in \mathbb{N}$, $\theta \in \Theta$ that $\sigma_\Omega(Y_n^\theta) \subseteq \sigma_\Omega\big((Y_{-1}^{(\theta,\vartheta)})_{\vartheta\in\Theta}, (Y_0^{(\theta,\vartheta)})_{\vartheta\in\Theta}, (Z^{(\theta,\vartheta)})_{\vartheta\in\Theta}\big)$,

(iii) it holds for every $n, m \in (\mathbb{N}_0 \cup \{-1\})$, $k \in \mathbb{N}$, $\theta_1, \theta_2, \vartheta \in \mathbb{Z}^k$ with $\theta_1 \neq \theta_2$ that $Y_n^{\theta_1}$, $Y_m^{\theta_2}$, and Z^ϑ are independent,

(iv) it holds for every $\theta \in \Theta$ that $\big(Y_l^{(\theta,l,i)}, Y_{l-1}^{(\theta,-l,i)}, Z^{(\theta,l,i)}\big)$, $i \in \mathbb{N}$, $l \in \mathbb{N}_0$, are independent,

(v) it holds for every $n \in (\mathbb{N}_0 \cup \{-1\})$ that Y_n^θ, $\theta \in \Theta$, are identically distributed, and

(vi) it holds for every $\theta \in \Theta$, $l \in \mathbb{N}_0$, $i \in \mathbb{N}$ that $\Omega \ni \omega \mapsto \Phi_l\big(Y_l^{(\theta,l,i)}(\omega), Y_{l-1}^{(\theta,-l,i)}(\omega), Z^{(\theta,l,i)}(\omega)\big) \in \mathcal{Y}$ and $\Omega \ni \omega \mapsto \Phi_l\big(Y_l^0(\omega), Y_{l-1}^1(\omega), Z^0(\omega)\big) \in \mathcal{Y}$ are identically distributed.

Proof of Proposition 2.8. Throughout this proof let $R^{\theta,l,i}\colon \Omega \to \mathcal{Y}\times\mathcal{Y}\times\mathcal{Z}$, $i \in \mathbb{N}$, $l \in \mathbb{N}_0$, $\theta \in \Theta$, satisfy for all $\theta \in \Theta$, $l \in \mathbb{N}_0$, $i \in \mathbb{N}$ that

$$R^{\theta,l,i}(\omega) = \big(Y_l^{(\theta,l,i)}(\omega), Y_{l-1}^{(\theta,-l,i)}(\omega), Z^{(\theta,l,i)}(\omega)\big) \tag{30}$$

and let $\Psi_n\colon (\mathcal{Y} \times \mathcal{Y} \times \mathcal{Z})^{\{0,1,\ldots,n-1\}\times\mathbb{N}} \to \mathcal{Y}$, $n \in \mathbb{N}$, satisfy for all $n \in \mathbb{N}$, $r = (r^{l,i})_{(l,i)\in\{0,1,\ldots,n-1\}\times\mathbb{N}} \in (\mathcal{Y} \times \mathcal{Y} \times \mathcal{Z})^{\{0,1,\ldots,n-1\}\times\mathbb{N}}$ that $\Psi_n(r) = \sum_{l=0}^{n-1} \frac{1}{M_{n,l}} \big[\sum_{i=1}^{M_{n,l}} \Phi_l(r^{l,i})\big]$.

First, note that the assumption that it holds for all $\theta \in \Theta$ that $Y_{-1}^\theta\colon \Omega \to \mathcal{Y}$ and $Y_0^\theta\colon \Omega \to \mathcal{Y}$ are $\mathscr{F}/\mathscr{B}(\mathcal{Y})$-measurable functions, the assumption that it holds for all

$\theta \in \Theta$ that $Z^\theta\colon \Omega \to \mathcal{Z}$ is an $\mathscr{F}/\mathscr{Z}$-measurable function, the assumption that it holds for all $l \in \mathbb{N}_0$ that $\Phi_l\colon \mathcal{Y} \times \mathcal{Y} \times \mathcal{Z} \to \mathcal{Y}$ is an $(\mathscr{B}(\mathcal{Y}) \otimes \mathscr{B}(\mathcal{Y}) \otimes \mathscr{Z})/\mathscr{B}(\mathcal{Y})$-measurable function, the assumption that $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ is a separable $\mathbb{R}$-Banach space, and induction on $\mathbb{N}_0$ prove (i).

Second, we show (ii) by induction on $n \in \mathbb{N}$. For the base case $n = 1$ observe that it holds for all $\theta \in \Theta$ that

$$Y_1^\theta = \tfrac{1}{M_{1,0}} \sum_{i=1}^{M_{1,0}} \Phi_0\big(Y_0^{(\theta,0,i)}, Y_{-1}^{(\theta,0,i)}, Z^{(\theta,0,i)}\big). \tag{31}$$

This demonstrates for all $\theta \in \Theta$ that

$$\begin{aligned} \sigma_\Omega(Y_1^\theta) &\subseteq \sigma_\Omega\big((Y_{-1}^{(\theta,0,i)})_{i\in\mathbb{N}}, (Y_0^{(\theta,0,i)})_{i\in\mathbb{N}}, (Z^{(\theta,0,i)})_{i\in\mathbb{N}}\big) \\ &\subseteq \sigma_\Omega\big((Y_{-1}^{(\theta,\vartheta)})_{\vartheta\in\Theta}, (Y_0^{(\theta,\vartheta)})_{\vartheta\in\Theta}, (Z^{(\theta,\vartheta)})_{\vartheta\in\Theta}\big). \end{aligned} \tag{32}$$

This establishes (ii) in the base case $n = 1$. For the induction step $\mathbb{N} \ni n-1 \to n \in \{2, 3, \ldots\}$ let $n \in \{2, 3, \ldots\}$ and assume for all $l \in \{1, \ldots, n-1\}$, $\theta \in \Theta$ that

$$\sigma_\Omega(Y_l^\theta) \subseteq \sigma_\Omega\big((Y_{-1}^{(\theta,\vartheta)})_{\vartheta\in\Theta}, (Y_0^{(\theta,\vartheta)})_{\vartheta\in\Theta}, (Z^{(\theta,\vartheta)})_{\vartheta\in\Theta}\big). \tag{33}$$

This and (29) imply for all $\theta \in \Theta$ that

$$\begin{aligned} \sigma_\Omega(Y_n^\theta) &\subseteq \sigma_\Omega\big((Y_{l-1}^{(\theta,-l,i)})_{(l,i)\in\{0,1,\ldots,n-1\}\times\mathbb{N}}, (Y_l^{(\theta,l,i)})_{(l,i)\in\{0,1,\ldots,n-1\}\times\mathbb{N}}, (Z^{(\theta,l,i)})_{(l,i)\in\mathbb{N}_0\times\mathbb{N}}\big) \\ &\subseteq \sigma_\Omega\big((Y_l^{(\theta,\mathfrak{z})})_{(l,\mathfrak{z})\in\{1,\ldots,n-1\}\times\mathbb{Z}^2}, (Y_{-1}^{(\theta,\mathfrak{z})})_{\mathfrak{z}\in\mathbb{Z}^2}, (Y_0^{(\theta,\mathfrak{z})})_{\mathfrak{z}\in\mathbb{Z}^2}, (Z^{(\theta,\mathfrak{z})})_{\mathfrak{z}\in\mathbb{Z}^2}\big) \\ &\subseteq \sigma_\Omega\big((Y_{-1}^{(\theta,\mathfrak{z},\vartheta)})_{(\mathfrak{z},\vartheta)\in\mathbb{Z}^2\times\Theta}, (Y_0^{(\theta,\mathfrak{z},\vartheta)})_{(\mathfrak{z},\vartheta)\in\mathbb{Z}^2\times\Theta}, (Z^{(\theta,\mathfrak{z},\vartheta)})_{(\mathfrak{z},\vartheta)\in\mathbb{Z}^2\times\Theta}, \\ &\qquad (Y_{-1}^{(\theta,\vartheta)})_{\vartheta\in\Theta}, (Y_0^{(\theta,\vartheta)})_{\vartheta\in\Theta}, (Z^{(\theta,\vartheta)})_{\vartheta\in\Theta}\big) \\ &= \sigma_\Omega\big((Y_{-1}^{(\theta,\vartheta)})_{\vartheta\in\Theta}, (Y_0^{(\theta,\vartheta)})_{\vartheta\in\Theta}, (Z^{(\theta,\vartheta)})_{\vartheta\in\Theta}\big). \end{aligned} \tag{34}$$

Induction hence establishes (ii).

Third, observe that the assumption that $(Y_{-1}^\theta)_{\theta\in\Theta}$, $(Y_0^\theta)_{\theta\in\Theta}$, and $(Z^\theta)_{\theta\in\Theta}$ are independent ensures that it holds for every $k \in \mathbb{N}$, $\theta_1, \theta_2, \vartheta \in \mathbb{Z}^k$ with $\theta_1 \neq \theta_2$ that $\sigma_\Omega\big((Y_{-1}^{(\theta_1,\mathfrak{z})})_{\mathfrak{z}\in\Theta}, (Y_0^{(\theta_1,\mathfrak{z})})_{\mathfrak{z}\in\Theta}, (Z^{(\theta_1,\mathfrak{z})})_{\mathfrak{z}\in\Theta}\big)$, $\sigma_\Omega\big((Y_{-1}^{(\theta_2,\mathfrak{z})})_{\mathfrak{z}\in\Theta}, (Y_0^{(\theta_2,\mathfrak{z})})_{\mathfrak{z}\in\Theta}, (Z^{(\theta_2,\mathfrak{z})})_{\mathfrak{z}\in\Theta}\big)$, and $\sigma_\Omega(Z^\vartheta)$ are independent. Combining this with (ii) proves (iii).

Fourth, note that the assumption that the family $(Y_{-1}^\theta)_{\theta\in\Theta}$ is independent, the assumption that the family $(Y_0^\theta)_{\theta\in\Theta}$ is independent, the assumption that the family $(Z^\theta)_{\theta\in\Theta}$ is independent, and the assumption that $(Y_{-1}^\theta)_{\theta\in\Theta}$, $(Y_0^\theta)_{\theta\in\Theta}$, and $(Z^\theta)_{\theta\in\Theta}$ are independent imply for every $\theta \in \Theta$ that the family

$$\mathbb{N}_0\times\mathbb{N} \ni (l,i) \mapsto \begin{cases} \sigma_\Omega\big(Y_{-1}^{(\theta,0,i)}, Y_0^{(\theta,0,i)}, Z^{(\theta,0,i)}\big) & : l = 0 \\ \sigma_\Omega\big((Y_{-1}^{(\theta,l,i,\vartheta)})_{\vartheta\in\Theta}, (Y_0^{(\theta,l,i,\vartheta)})_{\vartheta\in\Theta}, (Z^{(\theta,l,i,\vartheta)})_{\vartheta\in\Theta}, (Y_{-1}^{(\theta,-l,i,\vartheta)})_{\vartheta\in\Theta}, \\ \quad (Y_0^{(\theta,-l,i,\vartheta)})_{\vartheta\in\Theta}, (Z^{(\theta,-l,i,\vartheta)})_{\vartheta\in\Theta}, Z^{(\theta,l,i)}\big) & : l \neq 0 \end{cases}$$

is independent. This, (30), and (ii) ensure for every $\theta \in \Theta$ that the family

$$\big[\mathbb{N}_0 \times \mathbb{N} \ni (l,i) \mapsto \big(Y_l^{(\theta,l,i)}, Y_{l-1}^{(\theta,-l,i)}, Z^{(\theta,l,i)}\big)\big] = (R^{\theta,l,i})_{(l,i)\in\mathbb{N}_0\times\mathbb{N}} \tag{35}$$

is independent. This finishes the proof of (iv).

Fifth, we establish (v) by induction on $n \in \mathbb{N}$. For the base case $n = 1$ note that the assumption that Y_{-1}^θ, $\theta \in \Theta$, are identically distributed, the assumption that Y_0^θ, $\theta \in \Theta$, are identically distributed, the assumption that Z^θ, $\theta \in \Theta$, are identically distributed, the assumption that $(Y_{-1}^\theta)_{\theta\in\Theta}$, $(Y_0^\theta)_{\theta\in\Theta}$, and $(Z^\theta)_{\theta\in\Theta}$ are independent, and (30) establish for every $\theta \in \Theta$, $i \in \mathbb{N}$ that

$$R^{\theta,0,i} = \big(Y_0^{(\theta,0,i)}, Y_{-1}^{(\theta,0,i)}, Z^{(\theta,0,i)}\big) \quad \text{and} \quad (Y_0^0, Y_{-1}^1, Z^0) \tag{36}$$

are identically distributed. In particular, this shows for every $i \in \mathbb{N}$ that $R^{\theta,0,i}$, $\theta \in \Theta$, are identically distributed. Combining this with (35) proves that $\big(\Omega \ni \omega \mapsto (R^{\theta,0,i}(\omega))_{i\in\mathbb{N}} \in (\mathcal{Y}\times\mathcal{Y}\times\mathcal{Z})^{\mathbb{N}}\big)$, $\theta \in \Theta$, are identically distributed. The fact that $\forall\, \theta \in \Theta, \omega \in \Omega \colon Y_1^\theta(\omega) = \Psi_1\big((R^{\theta,0,i}(\omega))_{i\in\mathbb{N}}\big)$ hence implies that Y_1^θ, $\theta \in \Theta$, are identically distributed. This, the assumption that Y_{-1}^θ, $\theta \in \Theta$, are identically distributed, and the assumption that Y_0^θ, $\theta \in \Theta$, are identically distributed show (v) in the base case $n = 1$. For the induction step $\mathbb{N} \ni n-1 \to n \in \{2, 3, \ldots\}$ let $n \in \{2, 3, \ldots\}$ and assume for every $l \in \{-1, 0, 1, \ldots, n-1\}$ that Y_l^θ, $\theta \in \Theta$, are identically distributed. This, the assumption that Z^θ, $\theta \in \Theta$, are identically distributed, (iii), and (30) ensure for every $\theta \in \Theta$, $l \in \{1, \ldots, n-1\}$, $i \in \mathbb{N}$ that

$$R^{\theta,l,i} = \big(Y_l^{(\theta,l,i)}, Y_{l-1}^{(\theta,-l,i)}, Z^{(\theta,l,i)}\big) \quad \text{and} \quad (Y_l^0, Y_{l-1}^1, Z^0) \tag{37}$$

are identically distributed. Combining this and (36) establishes for every $l \in \{0, 1, \ldots, n-1\}$, $i \in \mathbb{N}$ that $R^{\theta,l,i}$, $\theta \in \Theta$, are identically distributed. This and (35) demonstrate that $\big(\Omega \ni \omega \mapsto (R^{\theta,l,i}(\omega))_{(l,i)\in\{0,1,\ldots,n-1\}\times\mathbb{N}} \in (\mathcal{Y}\times\mathcal{Y}\times\mathcal{Z})^{\{0,1,\ldots,n-1\}\times\mathbb{N}}\big)$, $\theta \in \Theta$, are identically distributed. Therefore, the fact that $\forall\, \theta \in \Theta, \omega \in \Omega \colon Y_n^\theta(\omega) = \Psi_n\big((R^{\theta,l,i}(\omega))_{(l,i)\in\{0,1,\ldots,n-1\}\times\mathbb{N}}\big)$ shows that Y_n^θ, $\theta \in \Theta$, are identically distributed. Induction hence proves (v).

Sixth, observe that (36) and (37) establish (vi). The proof of Proposition 2.8 is thus complete. □

2.4 Error analysis

Proposition 2.9. *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, let $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be a separable $\mathbb{R}$-Banach space, let $C, c \in (0,\infty)$, $(\mathfrak{c}_k)_{k\in\mathbb{N}_0} \subseteq (0,\infty)$, $\Theta = \cup_{n=1}^\infty \mathbb{Z}^n$, $y \in \mathcal{Y}$, for every $n \in \mathbb{N}$ let $(M_{n,l})_{l\in\{0,1,\ldots,n\}} \subseteq \mathbb{N}$ satisfy $M_{n,1} \geq M_{n,2} \geq \ldots \geq M_{n,n}$, let $(\mathcal{Z}, \mathscr{Z})$ be a measurable space, let $Z^\theta \colon \Omega \to \mathcal{Z}$, $\theta \in \Theta$, be i.i.d. $\mathscr{F}/\mathscr{Z}$-measurable functions, let $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}}, \|\cdot\|_{\mathcal{H}})$ be a separable $\mathbb{R}$-Hilbert space, let $\mathscr{S} = \sigma_{L(\mathcal{Y},\mathcal{H})}\big(\big\{\{\varphi \in L(\mathcal{Y},\mathcal{H}) \colon \varphi(x) \in \mathcal{B}\} \subseteq L(\mathcal{Y},\mathcal{H}) \colon x \in \mathcal{Y}, \mathcal{B} \in \mathscr{B}(\mathcal{H})\big\}\big)$, let $\psi_k \colon \Omega \to L(\mathcal{Y},\mathcal{H})$, $k \in \mathbb{N}_0$, be $\mathscr{F}/\mathscr{S}$-measurable functions, let $\Phi_l \colon \mathcal{Y}\times\mathcal{Y}\times\mathcal{Z} \to \mathcal{Y}$, $l \in \mathbb{N}_0$, be $(\mathscr{B}(\mathcal{Y})\otimes\mathscr{B}(\mathcal{Y})\otimes\mathscr{Z})/\mathscr{B}(\mathcal{Y})$-measurable functions, let $Y_{-1}^\theta \colon \Omega \to \mathcal{Y}$, $\theta \in \Theta$, be i.i.d. $\mathscr{F}/\mathscr{B}(\mathcal{Y})$-measurable functions, let $Y_0^\theta \colon \Omega \to \mathcal{Y}$, $\theta \in \Theta$, be i.i.d. $\mathscr{F}/\mathscr{B}(\mathcal{Y})$-measurable functions, assume that $(Y_{-1}^\theta)_{\theta\in\Theta}$, $(Y_0^\theta)_{\theta\in\Theta}$, $(Z^\theta)_{\theta\in\Theta}$, and $(\psi_k)_{k\in\mathbb{N}_0}$ are independent, let $Y_n^\theta \colon \Omega \to \mathcal{Y}$, $\theta \in \Theta$, $n \in \mathbb{N}$, satisfy for all $n \in \mathbb{N}$, $\theta \in \Theta$ that*

$$Y_n^\theta = \sum_{l=0}^{n-1} \tfrac{1}{M_{n,l}} \left[\sum_{i=1}^{M_{n,l}} \Phi_l\big(Y_l^{(\theta,l,i)}, Y_{l-1}^{(\theta,-l,i)}, Z^{(\theta,l,i)}\big) \right], \tag{38}$$

and assume for all $k \in \mathbb{N}_0$, $n \in \mathbb{N}$ that $\mathbb{E}\big[\|\Phi_k(Y_k^0, Y_{k-1}^1, Z^0)\|_{\mathcal{Y}}\big] < \infty$ and

$$\max\big\{\mathbb{E}\big[\|\psi_k(\Phi_0(Y_0^0, Y_{-1}^1, Z^0))\|_{\mathcal{H}}^2\big], \mathbb{1}_{\mathbb{N}}(k)\, \mathbb{E}\big[\|\psi_k(Y_0^0 - y)\|_{\mathcal{H}}^2\big]\big\} \leq \tfrac{C^2}{\mathfrak{c}_k}, \tag{39}$$

$$\mathbb{E}\big[\|\psi_k(\Phi_n(Y_n^0, Y_{n-1}^1, Z^0))\|_{\mathcal{H}}^2\big] \leq c\, \mathbb{E}\big[\|\psi_{k+1}(Y_n^0 - Y_{n-1}^1)\|_{\mathcal{H}}^2\big], \tag{40}$$

$$\mathbb{E}\left[\left\|\psi_k\left(y - \sum_{l=0}^{n-1} \mathbb{E}\big[\Phi_l(Y_l^0, Y_{l-1}^1, Z^0)\big]\right)\right\|_{\mathcal{H}}^2\right] \leq \tfrac{2c}{M_{n,n}}\, \mathbb{E}\big[\|\psi_{k+1}(Y_{n-1}^0 - y)\|_{\mathcal{H}}^2\big]. \tag{41}$$

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

$$\begin{aligned}&\big(\mathbb{E}\big[\|\psi_0(Y_N^0 - y)\|_{\mathcal{H}}^2\big]\big)^{1/2}\\&\leq C(1+4c)^{N/2}\Bigg[\min\Bigg(\Bigg\{\min\{M_{l_k,0}\mathfrak{c}_k, M_{l_k,1}\mathfrak{c}_{k+1}\}\prod_{j=1}^{k} M_{l_{j-1},l_j+1}\colon\\&\quad \begin{array}{l}k \in \mathbb{N}\cap[0,N-1],\,(l_i)_{i\in\{0,1,\ldots,k\}}\subseteq\\ \{1,2,\ldots,N\},\,N=l_0>l_1>\ldots>l_k\end{array}\Bigg\}\cup\big\{\min\{M_{N,0}\mathfrak{c}_0, M_{N,1}\mathfrak{c}_1\}\big\}\Bigg)\Bigg]^{-1/2}<\infty.\end{aligned} \tag{42}$$

Proof of Proposition 2.9. First of all, note that the assumption that $\forall\, l \in \mathbb{N}_0\colon \mathbb{E}\big[\|\Phi_l(Y_l^0, Y_{l-1}^1, Z^0)\|_{\mathcal{Y}}\big] < \infty$ and (vi) in Proposition 2.8 establish for all $l \in \mathbb{N}_0$, $i \in \mathbb{N}$ that

$$\mathbb{E}\Big[\big\|\Phi_l\big(Y_l^{(0,l,i)}, Y_{l-1}^{(0,-l,i)}, Z^{(0,l,i)}\big)\big\|_{\mathcal{Y}}\Big] = \mathbb{E}\big[\|\Phi_l(Y_l^0, Y_{l-1}^1, Z^0)\|_{\mathcal{Y}}\big] < \infty. \tag{43}$$

This, (i) in Proposition 2.8, and (38) ensure for all $n \in \mathbb{N}$ that $Y_n^0\colon \Omega \to \mathcal{Y}$ is an $\mathscr{F}/\mathscr{B}(\mathcal{Y})$-measurable function and

$$\begin{aligned}\mathbb{E}\big[\|Y_n^0\|_{\mathcal{Y}}\big] &= \mathbb{E}\left[\left\|\sum_{l=0}^{n-1}\tfrac{1}{M_{n,l}}\left[\sum_{i=1}^{M_{n,l}}\Phi_l\big(Y_l^{(0,l,i)}, Y_{l-1}^{(0,-l,i)}, Z^{(0,l,i)}\big)\right]\right\|_{\mathcal{Y}}\right]\\&\leq \sum_{l=0}^{n-1}\tfrac{1}{M_{n,l}}\left[\sum_{i=1}^{M_{n,l}}\mathbb{E}\Big[\big\|\Phi_l\big(Y_l^{(0,l,i)}, Y_{l-1}^{(0,-l,i)}, Z^{(0,l,i)}\big)\big\|_{\mathcal{Y}}\Big]\right]\\&= \sum_{l=0}^{n-1}\tfrac{1}{M_{n,l}}\left[\sum_{i=1}^{M_{n,l}}\mathbb{E}\big[\|\Phi_l(Y_l^0, Y_{l-1}^1, Z^0)\|_{\mathcal{Y}}\big]\right]\\&= \sum_{l=0}^{n-1}\mathbb{E}\big[\|\Phi_l(Y_l^0, Y_{l-1}^1, Z^0)\|_{\mathcal{Y}}\big] < \infty.\end{aligned} \tag{44}$$

In addition, (ii) in Proposition 2.8 yields for all $n \in \mathbb{N}$, $\theta \in \Theta$ that

$$\begin{aligned}\sigma_\Omega(Y_n^\theta) &\subseteq \sigma_\Omega\big((Y_{-1}^{(\theta,\vartheta)})_{\vartheta\in\Theta}, (Y_0^{(\theta,\vartheta)})_{\vartheta\in\Theta}, (Z^{(\theta,\vartheta)})_{\vartheta\in\Theta}\big)\\&\subseteq \sigma_\Omega\big((Y_{-1}^{\vartheta})_{\vartheta\in\Theta}, (Y_0^{\vartheta})_{\vartheta\in\Theta}, (Z^{\vartheta})_{\vartheta\in\Theta}\big).\end{aligned} \tag{45}$$

Note that this implies that $\sigma_\Omega\big((Y_n^\theta)_{(n,\theta)\in(\mathbb{N}_0\cup\{-1\})\times\Theta}, (Z^\theta)_{\theta\in\Theta}\big) \subseteq \sigma_\Omega\big((Y_{-1}^\theta)_{\theta\in\Theta}, (Y_0^\theta)_{\theta\in\Theta}, (Z^\theta)_{\theta\in\Theta}\big)$. This and the assumption that $(Y_{-1}^\theta)_{\theta\in\Theta}$, $(Y_0^\theta)_{\theta\in\Theta}$, $(Z^\theta)_{\theta\in\Theta}$, and $(\psi_k)_{k\in\mathbb{N}_0}$ are independent demonstrate for every $k \in \mathbb{N}_0$ that

$$\sigma_\Omega\big((Y_n^\theta)_{(n,\theta)\in(\mathbb{N}_0\cup\{-1\})\times\Theta}, (Z^\theta)_{\theta\in\Theta}\big) \quad \text{and} \quad \psi_k \tag{46}$$

are independent. Corollary 2.5 and (44) hence show for all $k \in \mathbb{N}_0$, $n \in \mathbb{N}$ that

$$\mathbb{E}\big[\|\psi_k(Y_n^0 - y)\|_{\mathcal{H}}^2\big] = \mathbb{E}\big[\|\psi_k(Y_n^0 - \mathbb{E}[Y_n^0])\|_{\mathcal{H}}^2\big] + \mathbb{E}\big[\|\psi_k(\mathbb{E}[Y_n^0] - y)\|_{\mathcal{H}}^2\big]. \tag{47}$$

Next observe that (38), (43), (iv) in Proposition 2.8, (46), and Corollary 2.7 (with $n = \sum_{l=0}^{\mathfrak{n}-1} M_{\mathfrak{n},l}$, $\psi = \psi_k$ for $\mathfrak{n} \in \mathbb{N}$, $k \in \mathbb{N}_0$ in the notation of Corollary 2.7) prove for all $k \in \mathbb{N}_0$,

$n \in \mathbb{N}$ that

$$
\begin{aligned}
&\mathbb{E}\big[\|\psi_k(Y_n^0 - \mathbb{E}[Y_n^0])\|_{\mathcal{H}}^2\big] \\
&= \mathbb{E}\Bigg[\bigg\|\psi_k\bigg(\sum_{l=0}^{n-1} \tfrac{1}{M_{n,l}} \sum_{i=1}^{M_{n,l}} \big(\Phi_l\big(Y_l^{(0,l,i)}, Y_{l-1}^{(0,-l,i)}, Z^{(0,l,i)}\big) - \mathbb{E}\big[\Phi_l\big(Y_l^{(0,l,i)}, Y_{l-1}^{(0,-l,i)}, Z^{(0,l,i)}\big)\big]\big)\bigg)\bigg\|_{\mathcal{H}}^2\Bigg] \\
&= \mathbb{E}\Bigg[\bigg\|\sum_{l=0}^{n-1} \sum_{i=1}^{M_{n,l}} \psi_k\Big(\tfrac{1}{M_{n,l}}\big(\Phi_l\big(Y_l^{(0,l,i)}, Y_{l-1}^{(0,-l,i)}, Z^{(0,l,i)}\big) - \mathbb{E}\big[\Phi_l\big(Y_l^{(0,l,i)}, Y_{l-1}^{(0,-l,i)}, Z^{(0,l,i)}\big)\big]\big)\Big)\bigg\|_{\mathcal{H}}^2\Bigg] \\
&= \sum_{l=0}^{n-1} \sum_{i=1}^{M_{n,l}} \mathbb{E}\Big[\big\|\tfrac{1}{M_{n,l}}\psi_k\big(\Phi_l\big(Y_l^{(0,l,i)}, Y_{l-1}^{(0,-l,i)}, Z^{(0,l,i)}\big) - \mathbb{E}\big[\Phi_l\big(Y_l^{(0,l,i)}, Y_{l-1}^{(0,-l,i)}, Z^{(0,l,i)}\big)\big]\big)\big\|_{\mathcal{H}}^2\Big] \\
&= \sum_{l=0}^{n-1} \tfrac{1}{(M_{n,l})^2}\Bigg[\sum_{i=1}^{M_{n,l}} \mathbb{E}\Big[\big\|\psi_k\big(\Phi_l\big(Y_l^{(0,l,i)}, Y_{l-1}^{(0,-l,i)}, Z^{(0,l,i)}\big) \\
&\qquad\qquad - \mathbb{E}\big[\Phi_l\big(Y_l^{(0,l,i)}, Y_{l-1}^{(0,-l,i)}, Z^{(0,l,i)}\big)\big]\big)\big\|_{\mathcal{H}}^2\Big]\Bigg].
\end{aligned} \tag{48}
$$

Moreover, the fact that it holds for all $k \in \mathbb{N}_0$, $x \in \mathcal{Y}$ that $\Omega \ni \omega \mapsto [\psi_k(\omega)](x) \in \mathcal{H}$ is an $\mathscr{F}/\mathscr{B}(\mathcal{H})$-measurable function (cf. Lemma 2.1) and the fact that it holds for all $k \in \mathbb{N}_0$, $\omega \in \Omega$ that $\mathcal{Y} \ni x \mapsto [\psi_k(\omega)](x) \in \mathcal{H}$ is a continuous function demonstrate that

$$
\mathcal{Y} \times \Omega \ni (x, \omega) \mapsto [\psi_k(\omega)](x) \in \mathcal{H} \tag{49}
$$

is a continuous random field. This, (46), (vi) in Proposition 2.8, and Hutzenthaler, Jentzen, & von Wurstemberger [69, Lemma 3.5] (with $S = \mathcal{Y}$, $E = \mathcal{H}$, $U = V = (\mathcal{Y} \times \Omega \ni (x, \omega) \mapsto [\psi_k(\omega)](x) \in \mathcal{H})$, $X = \Phi_l(Y_l^{(0,l,i)}, Y_{l-1}^{(0,-l,i)}, Z^{(0,l,i)})$, $Y = \Phi_l(Y_l^0, Y_{l-1}^1, Z^0)$ for $i \in \mathbb{N}$, $l, k \in \mathbb{N}_0$ in the notation of [69, Lemma 3.5]) ensure for all $k, l \in \mathbb{N}_0$, $i \in \mathbb{N}$ that

$$
\begin{aligned}
&\Omega \ni \omega \mapsto [\psi_k(\omega)]\big(\Phi_l\big(Y_l^{(0,l,i)}(\omega), Y_{l-1}^{(0,-l,i)}(\omega), Z^{(0,l,i)}(\omega)\big)\big) \in \mathcal{H} \quad \text{and} \\
&\Omega \ni \omega \mapsto [\psi_k(\omega)]\big(\Phi_l(Y_l^0(\omega), Y_{l-1}^1(\omega), Z^0(\omega))\big) \in \mathcal{H}
\end{aligned} \tag{50}
$$

are identically distributed. This, (48), (43), (46), and Corollary 2.5 (with $y = 0$, $Y = \Phi_l(Y_l^{(0,l,i)}, Y_{l-1}^{(0,-l,i)}, Z^{(0,l,i)})$, $\psi = \psi_k$ for $i \in \{1, 2, \ldots, M_{n,l}\}$, $l \in \{0, 1, \ldots, n-1\}$, $n \in \mathbb{N}$, $k \in \mathbb{N}_0$ in the notation of Corollary 2.5) imply for all $k \in \mathbb{N}_0$, $n \in \mathbb{N}$ that

$$
\begin{aligned}
&\mathbb{E}\big[\|\psi_k(Y_n^0 - \mathbb{E}[Y_n^0])\|_{\mathcal{H}}^2\big] \le \sum_{l=0}^{n-1} \tfrac{1}{(M_{n,l})^2}\Bigg[\sum_{i=1}^{M_{n,l}} \mathbb{E}\Big[\big\|\psi_k\big(\Phi_l\big(Y_l^{(0,l,i)}, Y_{l-1}^{(0,-l,i)}, Z^{(0,l,i)}\big)\big)\big\|_{\mathcal{H}}^2\Big]\Bigg] \\
&= \sum_{l=0}^{n-1} \tfrac{1}{(M_{n,l})^2}\Bigg[\sum_{i=1}^{M_{n,l}} \mathbb{E}\big[\|\psi_k(\Phi_l(Y_l^0, Y_{l-1}^1, Z^0))\|_{\mathcal{H}}^2\big]\Bigg] \\
&= \sum_{l=0}^{n-1} \Big(\tfrac{1}{M_{n,l}}\, \mathbb{E}\big[\|\psi_k(\Phi_l(Y_l^0, Y_{l-1}^1, Z^0))\|_{\mathcal{H}}^2\big]\Big) \\
&= \tfrac{1}{M_{n,0}}\, \mathbb{E}\big[\|\psi_k(\Phi_0(Y_0^0, Y_{-1}^1, Z^0))\|_{\mathcal{H}}^2\big] + \sum_{l=1}^{n-1} \Big(\tfrac{1}{M_{n,l}}\, \mathbb{E}\big[\|\psi_k(\Phi_l(Y_l^0, Y_{l-1}^1, Z^0))\|_{\mathcal{H}}^2\big]\Big).
\end{aligned} \tag{51}
$$

Assumptions (39)–(40), the fact that $\forall\, a, b \in \mathbb{R}\colon (a+b)^2 \le 2(a^2 + b^2)$, (49), (46), (v) in Proposition 2.8, [69, Lemma 3.5] (with $S = \mathcal{Y}$, $E = \mathcal{H}$, $U = V = (\mathcal{Y} \times \Omega \ni (x, \omega) \mapsto [\psi_k(\omega)](x) \in \mathcal{H})$, $X = Y_l^1 - y$, $Y = Y_l^0 - y$ for $l, k \in \mathbb{N}_0$ in the notation of [69, Lemma 3.5]), and the assumption that $\forall\, n \in \mathbb{N}\colon M_{n,1} \ge M_{n,2} \ge \ldots \ge M_{n,n}$ hence prove for all $k \in \mathbb{N}_0$,

$n \in \mathbb{N}$ that

$$\begin{aligned}
&\mathbb{E}\big[\|\psi_k(Y_n^0 - \mathbb{E}[Y_n^0])\|_{\mathcal{H}}^2\big] \le \tfrac{C^2}{M_{n,0}\mathfrak{c}_k} + \left[\sum_{l=1}^{n-1} \tfrac{c}{M_{n,l}} \, \mathbb{E}\big[\|\psi_{k+1}(Y_l^0 - Y_{l-1}^1)\|_{\mathcal{H}}^2\big]\right] \\
&= \tfrac{C^2}{M_{n,0}\mathfrak{c}_k} + \left[\sum_{l=1}^{n-1} \tfrac{c}{M_{n,l}} \, \mathbb{E}\big[\|\psi_{k+1}(Y_l^0 - y) + \psi_{k+1}(y - Y_{l-1}^1)\|_{\mathcal{H}}^2\big]\right] \\
&\le \tfrac{C^2}{M_{n,0}\mathfrak{c}_k} + \left[\sum_{l=1}^{n-1} \tfrac{c}{M_{n,l}} \, \mathbb{E}\big[(\|\psi_{k+1}(Y_l^0 - y)\|_{\mathcal{H}} + \|\psi_{k+1}(Y_{l-1}^1 - y)\|_{\mathcal{H}})^2\big]\right] \\
&\le \tfrac{C^2}{M_{n,0}\mathfrak{c}_k} + \left[\sum_{l=1}^{n-1} \tfrac{2c}{M_{n,l}} \, \mathbb{E}\big[\|\psi_{k+1}(Y_l^0 - y)\|_{\mathcal{H}}^2\big]\right] + \left[\sum_{l=1}^{n-1} \tfrac{2c}{M_{n,l}} \, \mathbb{E}\big[\|\psi_{k+1}(Y_{l-1}^1 - y)\|_{\mathcal{H}}^2\big]\right] \\
&= \tfrac{C^2}{M_{n,0}\mathfrak{c}_k} + \left[\sum_{l=1}^{n-1} \tfrac{2c}{M_{n,l}} \, \mathbb{E}\big[\|\psi_{k+1}(Y_l^0 - y)\|_{\mathcal{H}}^2\big]\right] + \left[\sum_{l=0}^{n-2} \tfrac{2c}{M_{n,l+1}} \, \mathbb{E}\big[\|\psi_{k+1}(Y_l^0 - y)\|_{\mathcal{H}}^2\big]\right] \\
&\le \tfrac{C^2}{M_{n,0}\mathfrak{c}_k} + \left[\sum_{l=1}^{n-1} \tfrac{2c}{M_{n,l+1}} \, \mathbb{E}\big[\|\psi_{k+1}(Y_l^0 - y)\|_{\mathcal{H}}^2\big]\right] + \left[\sum_{l=0}^{n-2} \tfrac{2c}{M_{n,l+1}} \, \mathbb{E}\big[\|\psi_{k+1}(Y_l^0 - y)\|_{\mathcal{H}}^2\big]\right] \\
&= \tfrac{C^2}{M_{n,0}\mathfrak{c}_k} + \left[\sum_{l=0}^{n-1} \tfrac{2(2-\mathbb{1}_{\{0\}}(l)-\mathbb{1}_{\{n-1\}}(l))c}{M_{n,l+1}} \, \mathbb{E}\big[\|\psi_{k+1}(Y_l^0 - y)\|_{\mathcal{H}}^2\big]\right].
\end{aligned} \tag{52}$$

Furthermore, note that (44), (38), (43), and (vi) in Proposition 2.8 show for all $n \in \mathbb{N}$ that

$$\begin{aligned}
\mathbb{E}[Y_n^0] &= \mathbb{E}\left[\sum_{l=0}^{n-1} \tfrac{1}{M_{n,l}} \left[\sum_{i=1}^{M_{n,l}} \Phi_l\big(Y_l^{(0,l,i)}, Y_{l-1}^{(0,-l,i)}, Z^{(0,l,i)}\big)\right]\right] \\
&= \sum_{l=0}^{n-1} \tfrac{1}{M_{n,l}} \left[\sum_{i=1}^{M_{n,l}} \mathbb{E}\big[\Phi_l\big(Y_l^{(0,l,i)}, Y_{l-1}^{(0,-l,i)}, Z^{(0,l,i)}\big)\big]\right] \\
&= \sum_{l=0}^{n-1} \tfrac{1}{M_{n,l}} \left[\sum_{i=1}^{M_{n,l}} \mathbb{E}\big[\Phi_l(Y_l^0, Y_{l-1}^1, Z^0)\big]\right] = \sum_{l=0}^{n-1} \mathbb{E}\big[\Phi_l(Y_l^0, Y_{l-1}^1, Z^0)\big].
\end{aligned} \tag{53}$$

This and assumption (41) establish for all $k \in \mathbb{N}_0$, $n \in \mathbb{N}$ that

$$\begin{aligned}
\mathbb{E}\big[\|\psi_k(\mathbb{E}[Y_n^0] - y)\|_{\mathcal{H}}^2\big] &= \mathbb{E}\left[\left\|\psi_k\left(\left[\sum_{l=0}^{n-1} \mathbb{E}\big[\Phi_l(Y_l^0, Y_{l-1}^1, Z^0)\big]\right] - y\right)\right\|_{\mathcal{H}}^2\right] \\
&\le \tfrac{2c}{M_{n,n}} \, \mathbb{E}\big[\|\psi_{k+1}(Y_{n-1}^0 - y)\|_{\mathcal{H}}^2\big].
\end{aligned} \tag{54}$$

Combining (47) with (52) and assumption (39) hence proves for all $k \in \mathbb{N}_0$, $n \in \mathbb{N}$ that

$$\begin{aligned}
&\mathbb{E}\big[\|\psi_k(Y_n^0 - y)\|_{\mathcal{H}}^2\big] \le \tfrac{C^2}{M_{n,0}\mathfrak{c}_k} + \left[\sum_{l=0}^{n-1} \tfrac{2(2-\mathbb{1}_{\{0\}}(l))c}{M_{n,l+1}} \, \mathbb{E}\big[\|\psi_{k+1}(Y_l^0 - y)\|_{\mathcal{H}}^2\big]\right] \\
&= \tfrac{C^2}{M_{n,0}\mathfrak{c}_k} + \tfrac{2c}{M_{n,1}} \, \mathbb{E}\big[\|\psi_{k+1}(Y_0^0 - y)\|_{\mathcal{H}}^2\big] + \left[\sum_{l=1}^{n-1} \tfrac{4c}{M_{n,l+1}} \, \mathbb{E}\big[\|\psi_{k+1}(Y_l^0 - y)\|_{\mathcal{H}}^2\big]\right] \\
&\le \tfrac{C^2}{M_{n,0}\mathfrak{c}_k} + \tfrac{2C^2c}{M_{n,1}\mathfrak{c}_{k+1}} + \left[\sum_{l=1}^{n-1} \tfrac{4c}{M_{n,l+1}} \, \mathbb{E}\big[\|\psi_{k+1}(Y_l^0 - y)\|_{\mathcal{H}}^2\big]\right].
\end{aligned} \tag{55}$$

Next we introduce some additional notation. For the remainder of this proof let $N \in \mathbb{N}$, let $\varepsilon_n \in [0, \infty]$, $n \in \{1, 2, \dots, N\}$, satisfy for all $n \in \{1, 2, \dots, N\}$ that

$$\begin{aligned}
\varepsilon_n = \max\Bigg(\Bigg\{ &\left(M_{l_{k-1},n+1} \prod_{j=1}^{k-1} M_{l_{j-1},l_j+1}\right)^{-1} \mathbb{E}\big[\|\psi_k(Y_n^0 - y)\|_{\mathcal{H}}^2\big] : \\
&\begin{array}{l} k \in \mathbb{N} \cap [0, N-n],\ (l_i)_{i \in \{0,1,\dots,k-1\}} \subseteq \{n+1, \\ \quad n+2, \dots, N\},\ N = l_0 > l_1 > \dots > l_{k-1} \end{array} \Bigg\} \cup \big\{ \mathbb{1}_{\{N\}}(n) \, \mathbb{E}\big[\|\psi_0(Y_N^0 - y)\|_{\mathcal{H}}^2\big] \big\}\Bigg),
\end{aligned} \tag{56}$$

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

$$a = C^2(1+2c)\Bigg[\min\Bigg(\Bigg\{\min\{M_{l_k,0}\mathfrak{c}_k, M_{l_k,1}\mathfrak{c}_{k+1}\}\prod_{j=1}^{k} M_{l_{j-1},l_j+1} \colon \begin{array}{c} k \in \mathbb{N} \cap [0, N-1], \, (l_i)_{i \in \{0,1,\ldots,k\}} \subseteq \\ \{1, 2, \ldots, N\}, \, N = l_0 > l_1 > \ldots > l_k \end{array}\Bigg\} \cup \big\{\min\{M_{N,0}\mathfrak{c}_0, M_{N,1}\mathfrak{c}_1\}\big\}\Bigg)\Bigg]^{-1}. \tag{57}$$

Observe that (55)–(57) establish for all $n \in \mathbb{N} \cap [0, N-1]$, $k \in \{1, 2, \ldots, N-n\}$, $(l_i)_{i \in \{0,1,\ldots,k-1\}} \subseteq \{n+1, n+2, \ldots, N\}$ with $l_0 = N$ and $\forall\, i \in \mathbb{N} \cap [0, k-1] \colon l_{i-1} > l_i$ that

$$\begin{aligned}
&\Bigg(M_{l_{k-1},n+1}\prod_{j=1}^{k-1} M_{l_{j-1},l_j+1}\Bigg)^{-1} \mathbb{E}\big[\|\psi_k(Y_n^0 - y)\|_{\mathcal{H}}^2\big] \\
&\le \Big(\tfrac{C^2}{M_{n,0}\mathfrak{c}_k} + \tfrac{2C^2 c}{M_{n,1}\mathfrak{c}_{k+1}}\Big)\Bigg(M_{l_{k-1},n+1}\prod_{j=1}^{k-1} M_{l_{j-1},l_j+1}\Bigg)^{-1} \\
&\quad + 4c\sum_{\ell=1}^{n-1}\Bigg[\Bigg(M_{n,\ell+1}M_{l_{k-1},n+1}\prod_{j=1}^{k-1} M_{l_{j-1},l_j+1}\Bigg)^{-1} \mathbb{E}\big[\|\psi_{k+1}(Y_\ell^0 - y)\|_{\mathcal{H}}^2\big]\Bigg] \\
&\le C^2(1+2c)\Bigg[\max_{l_k \in \{1,2,\ldots,l_{k-1}-1\}}\Bigg(\min\{M_{l_k,0}\mathfrak{c}_k, M_{l_k,1}\mathfrak{c}_{k+1}\}M_{l_{k-1},l_k+1}\prod_{j=1}^{k-1} M_{l_{j-1},l_j+1}\Bigg)^{-1}\Bigg] \\
&\quad + 4c\sum_{\ell=1}^{n-1}\max_{l_k \in \{\ell+1,\ell+2,\ldots,l_{k-1}-1\}}\Bigg[\Bigg(M_{l_k,\ell+1}M_{l_{k-1},l_k+1}\prod_{j=1}^{k-1} M_{l_{j-1},l_j+1}\Bigg)^{-1} \mathbb{E}\big[\|\psi_{k+1}(Y_\ell^0 - y)\|_{\mathcal{H}}^2\big]\Bigg] \\
&= C^2(1+2c)\Bigg[\min_{l_k \in \{1,2,\ldots,l_{k-1}-1\}}\Bigg(\min\{M_{l_k,0}\mathfrak{c}_k, M_{l_k,1}\mathfrak{c}_{k+1}\}\prod_{j=1}^{k} M_{l_{j-1},l_j+1}\Bigg)\Bigg]^{-1} \\
&\quad + 4c\sum_{\ell=1}^{n-1}\max_{l_k \in \{\ell+1,\ell+2,\ldots,l_{k-1}-1\}}\Bigg[\Bigg(M_{l_k,\ell+1}\prod_{j=1}^{k} M_{l_{j-1},l_j+1}\Bigg)^{-1} \mathbb{E}\big[\|\psi_{k+1}(Y_\ell^0 - y)\|_{\mathcal{H}}^2\big]\Bigg] \\
&\le a + 4c\sum_{\ell=1}^{n-1}\varepsilon_\ell.
\end{aligned} \tag{58}$$

In addition, (55)–(57) ensure that

$$\begin{aligned}
\varepsilon_N &= \mathbb{E}\big[\|\psi_0(Y_N^0 - y)\|_{\mathcal{H}}^2\big] \le \tfrac{C^2}{M_{N,0}\mathfrak{c}_0} + \tfrac{2C^2 c}{M_{N,1}\mathfrak{c}_1} + \Bigg[\sum_{\ell=1}^{N-1}\tfrac{4c}{M_{N,\ell+1}}\,\mathbb{E}\big[\|\psi_1(Y_\ell^0 - y)\|_{\mathcal{H}}^2\big]\Bigg] \\
&\le C^2(1+2c)(\min\{M_{N,0}\mathfrak{c}_0, M_{N,1}\mathfrak{c}_1\})^{-1} + 4c\Bigg[\sum_{\ell=1}^{N-1}(M_{N,\ell+1})^{-1}\,\mathbb{E}\big[\|\psi_1(Y_\ell^0 - y)\|_{\mathcal{H}}^2\big]\Bigg] \\
&\le a + 4c\sum_{\ell=1}^{N-1}\varepsilon_\ell.
\end{aligned} \tag{59}$$

This, (56), and (58) show for all $n \in \{1, 2, \ldots, N\}$ that

$$\varepsilon_n \le a + 4c\sum_{\ell=1}^{n-1}\varepsilon_\ell. \tag{60}$$

The fact that $a + c < \infty$ and the discrete Gronwall-type inequality in Agarwal [2, Corollary 4.1.2] hence establish for all $n \in \{1, 2, \ldots, N\}$ that

$$\varepsilon_n \le a(1+4c)^{n-1} < \infty. \tag{61}$$

This and (56)–(57) imply that

$$\begin{aligned}
&\mathbb{E}\big[\|\psi_0(Y_N^0 - y)\|_{\mathcal{H}}^2\big] = \varepsilon_N \leq a(1+4c)^{N-1} \\
&\leq C^2(1+4c)^N \Bigg[\min\Bigg(\Bigg\{\min\{M_{l_k,0}\mathfrak{c}_k, M_{l_k,1}\mathfrak{c}_{k+1}\} \prod_{j=1}^{k} M_{l_{j-1},l_j+1} \colon \\
&\quad \begin{matrix} k \in \mathbb{N} \cap [0, N-1],\, (l_i)_{i\in\{0,1,\ldots,k\}} \subseteq \\ \{1,2,\ldots,N\},\, N = l_0 > l_1 > \ldots > l_k \end{matrix}\Bigg\} \cup \big\{\min\{M_{N,0}\mathfrak{c}_0, M_{N,1}\mathfrak{c}_1\}\big\}\Bigg)\Bigg]^{-1} < \infty.
\end{aligned} \tag{62}$$

The proof of Proposition 2.9 is thus complete. □

Corollary 2.10. *Let* $(\Omega, \mathscr{F}, \mathbb{P})$ *be a probability space, let* $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ *be a separable* $\mathbb{R}$*-Banach space, let* $C, c \in (0,\infty)$*,* $(\mathfrak{c}_k)_{k\in\mathbb{N}_0} \subseteq (0,\infty)$*,* $\Theta = \cup_{n=1}^{\infty} \mathbb{Z}^n$*,* $M \in \mathbb{N}$*,* $y \in \mathcal{Y}$*, let* $(\mathcal{Z}, \mathscr{Z})$ *be a measurable space, let* $Z^\theta \colon \Omega \to \mathcal{Z}$*,* $\theta \in \Theta$*, be i.i.d.* $\mathscr{F}/\mathscr{Z}$*-measurable functions, let* $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}}, \|\cdot\|_{\mathcal{H}})$ *be a separable* $\mathbb{R}$*-Hilbert space, let* $\mathscr{S} = \sigma_{L(\mathcal{Y},\mathcal{H})}\big(\big\{\{\varphi \in L(\mathcal{Y},\mathcal{H}) \colon \varphi(x) \in \mathcal{B}\} \subseteq L(\mathcal{Y},\mathcal{H}) \colon x \in \mathcal{Y}, \mathcal{B} \in \mathscr{B}(\mathcal{H})\big\}\big)$*, let* $\psi_k \colon \Omega \to L(\mathcal{Y},\mathcal{H})$*,* $k \in \mathbb{N}_0$*, be* $\mathscr{F}/\mathscr{S}$*-measurable functions, let* $\Phi_l \colon \mathcal{Y}\times\mathcal{Y}\times\mathcal{Z} \to \mathcal{Y}$*,* $l \in \mathbb{N}_0$*, be* $(\mathscr{B}(\mathcal{Y})\otimes\mathscr{B}(\mathcal{Y})\otimes\mathscr{Z})/\mathscr{B}(\mathcal{Y})$*-measurable functions, let* $Y_{-1}^\theta \colon \Omega \to \mathcal{Y}$*,* $\theta \in \Theta$*, be i.i.d.* $\mathscr{F}/\mathscr{B}(\mathcal{Y})$*-measurable functions, let* $Y_0^\theta \colon \Omega \to \mathcal{Y}$*,* $\theta \in \Theta$*, be i.i.d.* $\mathscr{F}/\mathscr{B}(\mathcal{Y})$*-measurable functions, assume that* $(Y_{-1}^\theta)_{\theta\in\Theta}$*,* $(Y_0^\theta)_{\theta\in\Theta}$*,* $(Z^\theta)_{\theta\in\Theta}$*, and* $(\psi_k)_{k\in\mathbb{N}_0}$ *are independent, let* $Y_n^\theta \colon \Omega \to \mathcal{Y}$*,* $\theta \in \Theta$*,* $n \in \mathbb{N}$*, satisfy for all* $n \in \mathbb{N}$*,* $\theta \in \Theta$ *that*

$$Y_n^\theta = \sum_{l=0}^{n-1} \tfrac{1}{M^{n-l}} \left[\sum_{i=1}^{M^{n-l}} \Phi_l\big(Y_l^{(\theta,l,i)}, Y_{l-1}^{(\theta,-l,i)}, Z^{(\theta,l,i)}\big)\right], \tag{63}$$

and assume for all $k \in \mathbb{N}_0$*,* $n \in \mathbb{N}$ *that* $\mathbb{E}\big[\|\Phi_k(Y_k^0, Y_{k-1}^1, Z^0)\|_{\mathcal{Y}}\big] < \infty$ *and*

$$\max\big\{\mathbb{E}\big[\|\psi_k(\Phi_0(Y_0^0, Y_{-1}^1, Z^0))\|_{\mathcal{H}}^2\big], \mathbb{1}_{\mathbb{N}}(k)\,\mathbb{E}\big[\|\psi_k(Y_0^0 - y)\|_{\mathcal{H}}^2\big]\big\} \leq \tfrac{C^2}{\mathfrak{c}_k}, \tag{64}$$

$$\mathbb{E}\big[\|\psi_k(\Phi_n(Y_n^0, Y_{n-1}^1, Z^0))\|_{\mathcal{H}}^2\big] \leq c\,\mathbb{E}\big[\|\psi_{k+1}(Y_n^0 - Y_{n-1}^1)\|_{\mathcal{H}}^2\big], \tag{65}$$

$$\mathbb{E}\left[\left\|\psi_k\bigg(y - \sum_{l=0}^{n-1} \mathbb{E}\big[\Phi_l(Y_l^0, Y_{l-1}^1, Z^0)\big]\bigg)\right\|_{\mathcal{H}}^2\right] \leq 2c\,\mathbb{E}\big[\|\psi_{k+1}(Y_{n-1}^0 - y)\|_{\mathcal{H}}^2\big]. \tag{66}$$

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

$$\big(\mathbb{E}\big[\|\psi_0(Y_N^0 - y)\|_{\mathcal{H}}^2\big]\big)^{1/2} \leq C\big[\tfrac{1+4c}{M}\big]^{N/2} \max_{k\in\{0,1,\ldots,N\}} \sqrt{\tfrac{M^k}{\mathfrak{c}_k}} < \infty. \tag{67}$$

Proof of Corollary 2.10. Throughout this proof let $\mathfrak{M}_{n,l} \in \mathbb{N}$, $l \in \{0,1,\ldots,n\}$, $n \in \mathbb{N}$, be the natural numbers which satisfy for all $n \in \mathbb{N}$, $l \in \{0,1,\ldots,n\}$ that $\mathfrak{M}_{n,l} = M^{n-l}$. Note that it holds for all $n \in \mathbb{N}$ that $\mathfrak{M}_{n,1} \geq \mathfrak{M}_{n,2} \geq \ldots \geq \mathfrak{M}_{n,n}$. The fact that $\forall\, n \in \mathbb{N} \colon \mathfrak{M}_{n,n} = 1$ and Proposition 2.9 hence ensure for all $N \in \mathbb{N}$ that

$$\begin{aligned}
&\big(\mathbb{E}\big[\|\psi_0(Y_N^0 - y)\|_{\mathcal{H}}^2\big]\big)^{1/2} \\
&\leq C(1+4c)^{N/2} \Bigg[\min\Bigg(\Bigg\{\min\{\mathfrak{M}_{l_k,0}\mathfrak{c}_k, \mathfrak{M}_{l_k,1}\mathfrak{c}_{k+1}\} \prod_{j=1}^{k} \mathfrak{M}_{l_{j-1},l_j+1} \colon \\
&\quad \begin{matrix} k \in \mathbb{N} \cap [0, N-1],\, (l_i)_{i\in\{0,1,\ldots,k\}} \subseteq \\ \{1,2,\ldots,N\},\, N = l_0 > l_1 > \ldots > l_k \end{matrix}\Bigg\} \cup \big\{\min\{\mathfrak{M}_{N,0}\mathfrak{c}_0, \mathfrak{M}_{N,1}\mathfrak{c}_1\}\big\}\Bigg)\Bigg]^{-1/2}.
\end{aligned} \tag{68}$$

Next observe that it holds for all $N \in \mathbb{N}$, $k \in \mathbb{N} \cap [0, N-1]$, $(l_i)_{i\in\{0,1,\dots,k\}} \subseteq \{1, 2, \dots, N\}$ with $l_0 = N$ and $\forall\, i \in \{1, 2, \dots, k\}\colon l_{i-1} > l_i$ that

$$\begin{aligned}\min\{\mathfrak{M}_{l_k,0}\mathfrak{c}_k, \mathfrak{M}_{l_k,1}\mathfrak{c}_{k+1}\} \prod_{j=1}^{k} \mathfrak{M}_{l_{j-1},l_j+1} &= \min\{M^{l_k}\mathfrak{c}_k, M^{l_k-1}\mathfrak{c}_{k+1}\} \prod_{j=1}^{k} M^{l_{j-1}-l_j-1} \\ = \min\{M^{l_k}\mathfrak{c}_k, M^{l_k-1}\mathfrak{c}_{k+1}\} M^{l_0-l_k-k} &= \min\{M^{N-k}\mathfrak{c}_k, M^{N-(k+1)}\mathfrak{c}_{k+1}\}.\end{aligned} \tag{69}$$

This and (68) establish for all $N \in \mathbb{N}$ that

$$\begin{aligned}&\big(\mathbb{E}\big[\|\psi_0(Y_N^0 - y)\|_{\mathcal{H}}^2\big]\big)^{1/2} \\ &\leq C(1+4c)^{N/2}\Big[\min\Big(\big\{\min\{M^{N-k}\mathfrak{c}_k, M^{N-(k+1)}\mathfrak{c}_{k+1}\}\colon k \in \mathbb{N} \cap [0, N-1]\big\} \\ &\qquad \cup \big\{\min\{M^N\mathfrak{c}_0, M^{N-1}\mathfrak{c}_1\}\big\}\Big)\Big]^{-1/2} \\ &= C(1+4c)^{N/2}\left[\min_{k\in\{0,1,\dots,N-1\}} \min\{M^{N-k}\mathfrak{c}_k, M^{N-(k+1)}\mathfrak{c}_{k+1}\}\right]^{-1/2} \\ &= C(1+4c)^{N/2}\left[\min_{k\in\{0,1,\dots,N\}} (M^{N-k}\mathfrak{c}_k)\right]^{-1/2} \\ &= C\big[\tfrac{1+4c}{M}\big]^{N/2}\left[\min_{k\in\{0,1,\dots,N\}} \tfrac{\mathfrak{c}_k}{M^k}\right]^{-1/2} \\ &= C\big[\tfrac{1+4c}{M}\big]^{N/2} \max_{k\in\{0,1,\dots,N\}} \sqrt{\tfrac{M^k}{\mathfrak{c}_k}} < \infty.\end{aligned} \tag{70}$$

The proof of Corollary 2.10 is thus complete. □

2.5 Cost analysis

Proposition 2.11. *Let $M \in (0, \infty)$, $(\alpha_l)_{l\in\mathbb{N}_0}, (\beta_l)_{l\in\mathbb{N}_0}, (\gamma_l)_{l\in\mathbb{N}_0}, (\mathrm{Cost}_n)_{n\in\mathbb{N}_0\cup\{-1\}} \subseteq [0, \infty)$ satisfy for all $n \in \mathbb{N}$ that*

$$\mathrm{Cost}_n \leq \sum_{l=0}^{n-1} \big[M^{n-l}(\alpha_l \mathrm{Cost}_l + \beta_l \mathrm{Cost}_{l-1} + \gamma_l)\big]. \tag{71}$$

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

$$\mathrm{Cost}_n \leq M^n\left(\beta_0 \mathrm{Cost}_{-1} + \big(\alpha_0 + \tfrac{\beta_1}{M}\big)\mathrm{Cost}_0 + \sum_{l=0}^{n-1}[M^{-l}\gamma_l]\right) \prod_{l=1}^{n-1}\big(1 + \alpha_l + \tfrac{\beta_{l+1}}{M}\big). \tag{72}$$

Proof of Proposition 2.11. Observe that it holds for all $n \in \mathbb{N}$ that

$$\begin{aligned}
\mathrm{Cost}_n &\leq M^n \sum_{l=0}^{n-1} \big[M^{-l} (\alpha_l \mathrm{Cost}_l + \beta_l \mathrm{Cost}_{l-1} + \gamma_l) \big] \\
&= M^n \Bigg(\beta_0 \mathrm{Cost}_{-1} + \Bigg[\sum_{l=0}^{n-1} M^{-l} (\alpha_l \mathrm{Cost}_l + \gamma_l) \Bigg] + \Bigg[\sum_{l=1}^{n-1} M^{-l} \beta_l \mathrm{Cost}_{l-1} \Bigg] \Bigg) \\
&= M^n \Bigg(\beta_0 \mathrm{Cost}_{-1} + \Bigg[\sum_{l=0}^{n-1} M^{-l} (\alpha_l \mathrm{Cost}_l + \gamma_l) \Bigg] + \tfrac{1}{M} \Bigg[\sum_{l=0}^{n-2} M^{-l} \beta_{l+1} \mathrm{Cost}_l \Bigg] \Bigg) \\
&\leq M^n \Bigg(\beta_0 \mathrm{Cost}_{-1} + \Bigg[\sum_{l=0}^{n-1} M^{-l} \gamma_l \Bigg] + \Bigg[\sum_{l=0}^{n-1} M^{-l} \alpha_l \mathrm{Cost}_l \Bigg] + \tfrac{1}{M} \Bigg[\sum_{l=0}^{n-1} M^{-l} \beta_{l+1} \mathrm{Cost}_l \Bigg] \Bigg) \\
&= M^n \Bigg(\beta_0 \mathrm{Cost}_{-1} + \Bigg[\sum_{l=0}^{n-1} M^{-l} \gamma_l \Bigg] + \Bigg[\sum_{l=0}^{n-1} M^{-l} \big(\alpha_l + \tfrac{\beta_{l+1}}{M} \big) \mathrm{Cost}_l \Bigg] \Bigg) \\
&= M^n \Bigg(\beta_0 \mathrm{Cost}_{-1} + \big(\alpha_0 + \tfrac{\beta_1}{M} \big) \mathrm{Cost}_0 + \sum_{l=0}^{n-1} [M^{-l} \gamma_l] \Bigg) + M^n \Bigg[\sum_{l=1}^{n-1} M^{-l} \big(\alpha_l + \tfrac{\beta_{l+1}}{M} \big) \mathrm{Cost}_l \Bigg].
\end{aligned} \tag{73}$$

Theorem 4.1.1 in Agarwal [2] (with $a = 1$, $u(k) = \mathrm{Cost}_k$, $p(k) = M^k(\beta_0 \mathrm{Cost}_{-1} + (\alpha_0 + \frac{\beta_1}{M}) \mathrm{Cost}_0 + \sum_{l=0}^{k-1} [M^{-l} \gamma_l])$, $q(k) = M^k$, $f(k) = M^{-k}(\alpha_k + \frac{\beta_{k+1}}{M})$ for $k \in \mathbb{N}$ in the notation of [2, Theorem 4.1.1]) hence establishes for all $n \in \mathbb{N}$ that

$$\begin{aligned}
\mathrm{Cost}_n &\leq M^n \Bigg(\beta_0 \mathrm{Cost}_{-1} + \big(\alpha_0 + \tfrac{\beta_1}{M} \big) \mathrm{Cost}_0 + \sum_{l=0}^{n-1} [M^{-l} \gamma_l] \Bigg) \\
&\quad + M^n \sum_{l=1}^{n-1} \Bigg[M^l \Bigg(\beta_0 \mathrm{Cost}_{-1} + \big(\alpha_0 + \tfrac{\beta_1}{M} \big) \mathrm{Cost}_0 + \sum_{i=0}^{l-1} [M^{-i} \gamma_i] \Bigg) M^{-l} \big(\alpha_l + \tfrac{\beta_{l+1}}{M} \big) \\
&\qquad \cdot \prod_{i=l+1}^{n-1} \big[1 + M^i M^{-i} \big(\alpha_i + \tfrac{\beta_{i+1}}{M} \big) \big] \Bigg] \\
&\leq M^n \Bigg(\beta_0 \mathrm{Cost}_{-1} + \big(\alpha_0 + \tfrac{\beta_1}{M} \big) \mathrm{Cost}_0 + \sum_{l=0}^{n-1} [M^{-l} \gamma_l] \Bigg) \\
&\quad \cdot \Bigg(1 + \sum_{l=1}^{n-1} \Bigg[\big(\alpha_l + \tfrac{\beta_{l+1}}{M} \big) \prod_{i=l+1}^{n-1} \big(1 + \alpha_i + \tfrac{\beta_{i+1}}{M} \big) \Bigg] \Bigg).
\end{aligned} \tag{74}$$

This and [2, Problem 1.9.10] show for all $n \in \mathbb{N}$ that

$$\mathrm{Cost}_n \leq M^n \Bigg(\beta_0 \mathrm{Cost}_{-1} + \big(\alpha_0 + \tfrac{\beta_1}{M} \big) \mathrm{Cost}_0 + \sum_{l=0}^{n-1} [M^{-l} \gamma_l] \Bigg) \prod_{l=1}^{n-1} \big(1 + \alpha_l + \tfrac{\beta_{l+1}}{M} \big). \tag{75}$$

The proof of Proposition 2.11 is thus complete. □

Lemma 2.12. *Let $a, b \in [0, \infty)$. Then it holds for all $n \in \mathbb{N}$ that*

$$(an + b) b^{n-1} \leq (a + b)^n. \tag{76}$$

Proof of Lemma 2.12. We prove (76) by induction on $n \in \mathbb{N}$. Note that the base case $n = 1$ is clear. For the induction step $\mathbb{N} \ni n - 1 \to n \in \{2, 3, \dots\}$ let $n \in \{2, 3, \dots\}$ and assume that $(a(n-1)+b)b^{n-2} \leq (a+b)^{n-1}$. This ensures that

$$\begin{aligned}(an+b)b^{n-1} &= ab^{n-1} + (a(n-1)+b)b^{n-1} \leq ab^{n-1} + b(a+b)^{n-1} \\ &\leq a(a+b)^{n-1} + b(a+b)^{n-1} = (a+b)^n.\end{aligned} \tag{77}$$

Induction hence completes the proof of Lemma 2.12. □

Corollary 2.13. *Let $M \in [1,\infty)$, $\mathfrak{z}, \alpha, \beta, \gamma \in [0,\infty)$, $(\mathrm{Cost}_n)_{n\in\mathbb{N}_0\cup\{-1\}} \subseteq [0,\infty)$ satisfy for all $n \in \mathbb{N}$ that $\mathrm{Cost}_{-1} = \mathrm{Cost}_0 = 0$ and*

$$\mathrm{Cost}_n \leq M^n \mathfrak{z} + \sum_{l=0}^{n-1} \big[M^{n-l} (\alpha \mathrm{Cost}_l + \beta \mathrm{Cost}_{l-1} + \gamma \mathfrak{z}) \big]. \tag{78}$$

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

$$\mathrm{Cost}_n \leq (1+\alpha+\beta+\gamma)^n M^n \mathfrak{z}. \tag{79}$$

Proof of Corollary 2.13. Note that Proposition 2.11 demonstrates for all $n \in \mathbb{N}$ that

$$\begin{aligned}\mathrm{Cost}_n &\leq M^n \left(\mathfrak{z} + \gamma \mathfrak{z} \sum_{l=0}^{n-1} M^{-l} \right) \prod_{l=1}^{n-1} \big(1 + \alpha + \tfrac{\beta}{M}\big) \\ &\leq \left(1 + \gamma \sum_{l=0}^{n-1} M^{-l} \right) (1+\alpha+\beta)^{n-1} M^n \mathfrak{z}.\end{aligned} \tag{80}$$

In addition, observe that it holds for all $n \in \mathbb{N}$ that

$$\sum_{l=0}^{n-1} M^{-l} \leq \sum_{l=0}^{n-1} 1 = n. \tag{81}$$

Furthermore, Lemma 2.12 implies for all $n \in \mathbb{N}$ that

$$(1+\gamma n)(1+\alpha+\beta)^{n-1} \leq (\gamma n + 1 + \alpha + \beta)(1+\alpha+\beta)^{n-1} \leq (1+\alpha+\beta+\gamma)^n. \tag{82}$$

This, (80), and (81) prove for all $n \in \mathbb{N}$ that

$$\mathrm{Cost}_n \leq (1+\gamma n)(1+\alpha+\beta)^{n-1} M^n \mathfrak{z} \leq (1+\alpha+\beta+\gamma)^n M^n \mathfrak{z}. \tag{83}$$

The proof of Corollary 2.13 is thus complete. □

2.6 Complexity analysis

Theorem 2.14. *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, let $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be a separable $\mathbb{R}$-Banach space, let $\mathfrak{z}, \gamma \in [0,\infty)$, $\mathfrak{B}, \mathfrak{b}, C \in [1,\infty)$, $c \in (0,\infty)$, $(\mathfrak{c}_k)_{k\in\mathbb{N}_0} \subseteq (0,\infty)$, $\Theta = \cup_{n=1}^\infty \mathbb{Z}^n$, $(M_j)_{j\in\mathbb{N}} \subseteq \mathbb{N}$, $y, \mathfrak{y}_{-1}, \mathfrak{y}_0 \in \mathcal{Y}$ satisfy $\liminf_{j\to\infty} M_j = \infty$, $\sup_{j\in\mathbb{N}} {}^{M_{j+1}}/_{M_j} \leq \mathfrak{B}$, and $\forall\, n \in \mathbb{N}\colon \max_{k\in\{0,1,\dots,n\}} {}^{(M_n)^k}/_{\mathfrak{c}_k} \leq \mathfrak{b}^n$, let $(\mathcal{Z}, \mathscr{Z})$ be a measurable space, let $Z^\theta\colon \Omega \to \mathcal{Z}$, $\theta \in \Theta$, be i.i.d. $\mathscr{F}/\mathscr{Z}$-measurable functions, let $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}}, \|\cdot\|_{\mathcal{H}})$ be a separable $\mathbb{R}$-Hilbert space, let $\mathscr{S} = \sigma_{L(\mathcal{Y},\mathcal{H})}\big(\{\{\varphi \in L(\mathcal{Y},\mathcal{H})\colon \varphi(x) \in \mathcal{B}\} \subseteq L(\mathcal{Y},\mathcal{H})\colon x \in \mathcal{Y}, \mathcal{B} \in \mathscr{B}(\mathcal{H})\}\big)$, let $\psi_k\colon \Omega \to L(\mathcal{Y},\mathcal{H})$, $k \in \mathbb{N}_0$, be $\mathscr{F}/\mathscr{S}$-measurable functions, assume that $(Z^\theta)_{\theta\in\Theta}$ and $(\psi_k)_{k\in\mathbb{N}_0}$ are independent, let $\Phi_l\colon \mathcal{Y}\times\mathcal{Y}\times\mathcal{Z} \to \mathcal{Y}$, $l \in \mathbb{N}_0$, be $(\mathscr{B}(\mathcal{Y})\otimes\mathscr{B}(\mathcal{Y})\otimes\mathscr{Z})/\mathscr{B}(\mathcal{Y})$-measurable functions, let $Y^\theta_{n,j}\colon \Omega \to \mathcal{Y}$, $\theta \in \Theta$, $j \in \mathbb{N}$, $n \in (\mathbb{N}_0 \cup \{-1\})$, satisfy for all $n, j \in \mathbb{N}$, $\theta \in \Theta$ that $Y^\theta_{-1,j} = \mathfrak{y}_{-1}$, $Y^\theta_{0,j} = \mathfrak{y}_0$, and*

$$Y^\theta_{n,j} = \sum_{l=0}^{n-1} \tfrac{1}{(M_j)^{n-l}} \left[\sum_{i=1}^{(M_j)^{n-l}} \Phi_l\big(Y^{(\theta,l,i)}_{l,j}, Y^{(\theta,-l,i)}_{l-1,j}, Z^{(\theta,l,i)}\big) \right], \tag{84}$$

let $(\mathrm{Cost}_{n,j})_{(n,j)\in(\mathbb{N}_0\cup\{-1\})\times\mathbb{N}}\subseteq[0,\infty)$ *satisfy for all* $n,j\in\mathbb{N}$ *that* $\mathrm{Cost}_{-1,j}=\mathrm{Cost}_{0,j}=0$ *and*

$$\mathrm{Cost}_{n,j}\leq(M_j)^n\mathfrak{z}+\sum_{l=0}^{n-1}\big[(M_j)^{n-l}(\mathrm{Cost}_{l,j}+\mathrm{Cost}_{l-1,j}+\gamma\mathfrak{z})\big], \tag{85}$$

and assume for all $k\in\mathbb{N}_0$, $n,j\in\mathbb{N}$ *that* $\mathbb{E}\big[\|\Phi_k(Y^0_{k,j},Y^1_{k-1,j},Z^0)\|_{\mathcal{Y}}\big]<\infty$ *and*

$$\max\big\{\mathbb{E}\big[\|\psi_k(\Phi_0(\mathfrak{y}_0,\mathfrak{y}_{-1},Z^0))\|^2_{\mathcal{H}}\big],\mathbb{1}_{\mathbb{N}}(k)\,\mathbb{E}\big[\|\psi_k(\mathfrak{y}_0-y)\|^2_{\mathcal{H}}\big]\big\}\leq\tfrac{C^2}{\mathfrak{c}_k}, \tag{86}$$

$$\mathbb{E}\Big[\big\|\psi_k\big(\Phi_n\big(Y^0_{n,j},Y^1_{n-1,j},Z^0\big)\big)\big\|^2_{\mathcal{H}}\Big]\leq c\,\mathbb{E}\Big[\big\|\psi_{k+1}\big(Y^0_{n,j}-Y^1_{n-1,j}\big)\big\|^2_{\mathcal{H}}\Big], \tag{87}$$

$$\mathbb{E}\left[\left\|\psi_k\left(y-\sum_{l=0}^{n-1}\mathbb{E}\big[\Phi_l\big(Y^0_{l,j},Y^1_{l-1,j},Z^0\big)\big]\right)\right\|^2_{\mathcal{H}}\right]\leq 2c\,\mathbb{E}\Big[\big\|\psi_{k+1}\big(Y^0_{n-1,j}-y\big)\big\|^2_{\mathcal{H}}\Big]. \tag{88}$$

Then

(i) it holds for all $n\in\mathbb{N}$ *that*

$$\big(\mathbb{E}\big[\|\psi_0(Y^0_{n,n}-y)\|^2_{\mathcal{H}}\big]\big)^{1/2}\leq C\left[\frac{\mathfrak{b}(1+4c)}{M_n}\right]^{n/2}<\infty, \tag{89}$$

(ii) it holds for all $n\in\mathbb{N}$ *that* $\mathrm{Cost}_{n,n}\leq(3+\gamma)^n(M_n)^n\mathfrak{z}$, *and*

(iii) there exists $(N_\varepsilon)_{\varepsilon\in(0,1]}\subseteq\mathbb{N}$ *such that it holds for all* $\varepsilon\in(0,1]$, $\delta\in(0,\infty)$ *that* $\sup_{n\in\{N_\varepsilon,N_\varepsilon+1,\dots\}}\big(\mathbb{E}\big[\|\psi_0(Y^0_{n,n}-y)\|^2_{\mathcal{H}}\big]\big)^{1/2}\leq\varepsilon$ *and*

$$\mathrm{Cost}_{N_\varepsilon,N_\varepsilon}\leq\mathfrak{z}\mathfrak{b}\mathfrak{c}_1(3+\gamma)C^{2(1+\delta)}\left(1+\sup_{n\in\mathbb{N}}\left[\frac{[\mathfrak{B}\mathfrak{b}^2(3+\gamma)(1+4c)]^{(1+\delta)}}{(M_n)^\delta}\right]^n\right)\varepsilon^{-2(1+\delta)}<\infty. \tag{90}$$

Proof of Theorem 2.14. Throughout this proof let $(N_\varepsilon)_{\varepsilon\in(0,1]}\subseteq\mathbb{N}$ be the family of natural numbers which satisfies for all $\varepsilon\in(0,1]$ that

$$N_\varepsilon=\min\left\{\mathfrak{N}\in\mathbb{N}\colon\sup_{n\in\{\mathfrak{N},\mathfrak{N}+1,\dots\}}C\left[\frac{\mathfrak{b}(1+4c)}{M_n}\right]^{n/2}\leq\varepsilon\right\}. \tag{91}$$

Observe that Corollary 2.10 and the assumption that $\forall\,n\in\mathbb{N}\colon\max_{k\in\{0,1,\dots,n\}}{}^{(M_n)^k}/_{\mathfrak{c}_k}\leq\mathfrak{b}^n$ establish for all $n\in\mathbb{N}$ that

$$\big(\mathbb{E}\big[\|\psi_0(Y^0_{n,n}-y)\|^2_{\mathcal{H}}\big]\big)^{1/2}\leq C\left[\frac{1+4c}{M_n}\right]^{n/2}\max_{k\in\{0,1,\dots,n\}}\sqrt{\tfrac{(M_n)^k}{\mathfrak{c}_k}}\leq C\left[\frac{\mathfrak{b}(1+4c)}{M_n}\right]^{n/2}<\infty. \tag{92}$$

This proves (i). In addition, (85) and Corollary 2.13 demonstrate for all $n\in\mathbb{N}$ that

$$\mathrm{Cost}_{n,n}\leq(3+\gamma)^n(M_n)^n\mathfrak{z}. \tag{93}$$

This finishes the proof of (ii). It thus remains to show (iii). Observe that (92) and (91) ensure for all $\varepsilon\in(0,1]$ that

$$\sup_{n\in\{N_\varepsilon,N_\varepsilon+1,\dots\}}\big(\mathbb{E}\big[\|\psi_0(Y^0_{n,n}-y)\|^2_{\mathcal{H}}\big]\big)^{1/2}\leq\sup_{n\in\{N_\varepsilon,N_\varepsilon+1,\dots\}}C\left[\frac{\mathfrak{b}(1+4c)}{M_n}\right]^{n/2}\leq\varepsilon. \tag{94}$$

Furthermore, note that (91) implies for all $\varepsilon \in (0,1]$ with $N_\varepsilon \geq 2$ that

$$C\left[\frac{\mathfrak{b}(1+4c)}{M_{N_\varepsilon - 1}}\right]^{(N_\varepsilon-1)/2} > \varepsilon. \tag{95}$$

This, (93), the assumption that $\sup_{j\in\mathbb{N}} {}^{M_{j+1}}/{}_{M_j} \leq \mathfrak{B}$, and the fact that $\forall\, n \in \mathbb{N}\colon {}^{M_n}/{}_{\mathfrak{c}_1} \leq \mathfrak{b}^n$ show for all $\varepsilon \in (0,1]$, $\delta \in (0,\infty)$ with $N_\varepsilon \geq 2$ that

$$\begin{aligned}
\mathrm{Cost}_{N_\varepsilon,N_\varepsilon} &\leq (3+\gamma)^{N_\varepsilon}(M_{N_\varepsilon})^{N_\varepsilon}\mathfrak{z} \\
&\leq (3+\gamma)^{N_\varepsilon}(M_{N_\varepsilon})^{N_\varepsilon}\mathfrak{z}\left[C\left[\frac{\mathfrak{b}(1+4c)}{M_{N_\varepsilon-1}}\right]^{(N_\varepsilon-1)/2}\varepsilon^{-1}\right]^{2(1+\delta)} \\
&= \mathfrak{z}C^{2(1+\delta)}\varepsilon^{-2(1+\delta)}\left[\frac{(3+\gamma)^{N_\varepsilon}(M_{N_\varepsilon})^{N_\varepsilon}[\mathfrak{b}(1+4c)]^{(N_\varepsilon-1)(1+\delta)}}{(M_{N_\varepsilon-1})^{(N_\varepsilon-1)(1+\delta)}}\right] \\
&\leq \mathfrak{z}C^{2(1+\delta)}\varepsilon^{-2(1+\delta)}\sup_{n\in\mathbb{N}}\left[\frac{(3+\gamma)^{n+1}(M_{n+1})^{n+1}[\mathfrak{b}(1+4c)]^{n(1+\delta)}}{(M_n)^{n(1+\delta)}}\right] \\
&\leq \mathfrak{z}(3+\gamma)C^{2(1+\delta)}\varepsilon^{-2(1+\delta)}\sup_{n\in\mathbb{N}}\left[\frac{M_{n+1}(M_{n+1})^n[\mathfrak{b}(3+\gamma)(1+4c)]^{n(1+\delta)}}{(M_n)^n(M_n)^{n\delta}}\right] \\
&\leq \mathfrak{z}\mathfrak{c}_1(3+\gamma)C^{2(1+\delta)}\varepsilon^{-2(1+\delta)}\sup_{n\in\mathbb{N}}\left[\frac{\mathfrak{b}^{n+1}\mathfrak{B}^n[\mathfrak{b}(3+\gamma)(1+4c)]^{n(1+\delta)}}{(M_n)^{n\delta}}\right] \\
&\leq \mathfrak{z}\mathfrak{b}\mathfrak{c}_1(3+\gamma)C^{2(1+\delta)}\varepsilon^{-2(1+\delta)}\sup_{n\in\mathbb{N}}\left[\frac{[\mathfrak{B}\mathfrak{b}^2(3+\gamma)(1+4c)]^{(1+\delta)}}{(M_n)^{\delta}}\right]^n \\
&\leq \mathfrak{z}\mathfrak{b}\mathfrak{c}_1(3+\gamma)C^{2(1+\delta)}\left(1+\sup_{n\in\mathbb{N}}\left[\frac{[\mathfrak{B}\mathfrak{b}^2(3+\gamma)(1+4c)]^{(1+\delta)}}{(M_n)^{\delta}}\right]^n\right)\varepsilon^{-2(1+\delta)}.
\end{aligned} \tag{96}$$

Moreover, (85), the fact that ${}^{M_1}/{}_{\mathfrak{c}_1} \leq \mathfrak{b}$, and the fact that $C \geq 1$ ensure for all $\varepsilon \in (0,1]$, $\delta \in (0,\infty)$ that

$$\begin{aligned}
\mathrm{Cost}_{1,1} &\leq \mathfrak{z}(1+\gamma)M_1 \leq \mathfrak{z}\mathfrak{b}\mathfrak{c}_1(3+\gamma) \\
&\leq \mathfrak{z}\mathfrak{b}\mathfrak{c}_1(3+\gamma)C^{2(1+\delta)}\left(1+\sup_{n\in\mathbb{N}}\left[\frac{[\mathfrak{B}\mathfrak{b}^2(3+\gamma)(1+4c)]^{(1+\delta)}}{(M_n)^{\delta}}\right]^n\right)\varepsilon^{-2(1+\delta)}.
\end{aligned} \tag{97}$$

Combining this with (96) establishes for all $\varepsilon \in (0,1]$, $\delta \in (0,\infty)$ that

$$\mathrm{Cost}_{N_\varepsilon,N_\varepsilon} \leq \mathfrak{z}\mathfrak{b}\mathfrak{c}_1(3+\gamma)C^{2(1+\delta)}\left(1+\sup_{n\in\mathbb{N}}\left[\frac{[\mathfrak{B}\mathfrak{b}^2(3+\gamma)(1+4c)]^{(1+\delta)}}{(M_n)^{\delta}}\right]^n\right)\varepsilon^{-2(1+\delta)} < \infty. \tag{98}$$

The proof of Theorem 2.14 is thus complete. □

Corollary 2.15. *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, let $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be a separable $\mathbb{R}$-Banach space, let $\mathfrak{z}, \gamma \in [0,\infty)$, $\mathfrak{B}, \kappa, C \in [1,\infty)$, $c \in (0,\infty)$, $\Theta = \cup_{n=1}^\infty \mathbb{Z}^n$, $(M_j)_{j\in\mathbb{N}} \subseteq \mathbb{N}$, $y, \mathfrak{y}_{-1}, \mathfrak{y}_0 \in \mathcal{Y}$ satisfy $\liminf_{j\to\infty} M_j = \infty$, $\sup_{j\in\mathbb{N}} {}^{M_{j+1}}/{}_{M_j} \leq \mathfrak{B}$, and $\sup_{j\in\mathbb{N}} {}^{M_j}/{}_j \leq \kappa$, let $(\mathcal{Z}, \mathscr{Z})$ be a measurable space, let $Z^\theta\colon \Omega \to \mathcal{Z}$, $\theta \in \Theta$, be i.i.d. $\mathscr{F}/\mathscr{Z}$-measurable functions, let $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}}, \|\cdot\|_{\mathcal{H}})$ be a separable $\mathbb{R}$-Hilbert space, let $\mathscr{S} = \sigma_{L(\mathcal{Y},\mathcal{H})}\big(\{\{\varphi \in L(\mathcal{Y},\mathcal{H})\colon \varphi(x) \in \mathcal{B}\} \subseteq L(\mathcal{Y},\mathcal{H})\colon x \in \mathcal{Y}, \mathcal{B} \in \mathscr{B}(\mathcal{H})\}\big)$, let $\psi_k\colon \Omega \to L(\mathcal{Y},\mathcal{H})$, $k \in \mathbb{N}_0$, be $\mathscr{F}/\mathscr{S}$-measurable functions, assume that $(Z^\theta)_{\theta\in\Theta}$ and $(\psi_k)_{k\in\mathbb{N}_0}$ are independent, let $\Phi_l\colon \mathcal{Y}\times\mathcal{Y}\times\mathcal{Z} \to \mathcal{Y}$, $l \in \mathbb{N}_0$, be $(\mathscr{B}(\mathcal{Y})\otimes\mathscr{B}(\mathcal{Y})\otimes\mathscr{Z})/\mathscr{B}(\mathcal{Y})$-measurable functions, let $Y^\theta_{n,j}\colon \Omega \to \mathcal{Y}$, $\theta \in \Theta$, $j \in \mathbb{N}$, $n \in (\mathbb{N}_0 \cup \{-1\})$, satisfy for all $n, j \in \mathbb{N}$, $\theta \in \Theta$ that $Y^\theta_{-1,j} = \mathfrak{y}_{-1}$, $Y^\theta_{0,j} = \mathfrak{y}_0$, and*

$$Y^\theta_{n,j} = \sum_{l=0}^{n-1} \tfrac{1}{(M_j)^{n-l}}\left[\sum_{i=1}^{(M_j)^{n-l}} \Phi_l\big(Y^{(\theta,l,i)}_{l,j}, Y^{(\theta,-l,i)}_{l-1,j}, Z^{(\theta,l,i)}\big)\right], \tag{99}$$

let $(\mathrm{Cost}_{n,j})_{(n,j)\in(\mathbb{N}_0\cup\{-1\})\times\mathbb{N}} \subseteq [0,\infty)$ *satisfy for all* $n,j\in\mathbb{N}$ *that* $\mathrm{Cost}_{-1,j} = \mathrm{Cost}_{0,j} = 0$ *and*

$$\mathrm{Cost}_{n,j} \leq (M_j)^n \mathfrak{z} + \sum_{l=0}^{n-1} \big[(M_j)^{n-l} (\mathrm{Cost}_{l,j} + \mathrm{Cost}_{l-1,j} + \gamma\mathfrak{z}) \big], \tag{100}$$

and assume for all $k\in\mathbb{N}_0$, $n,j\in\mathbb{N}$, $u,v\in\mathcal{Y}$ *that* $\mathbb{E}\big[\|\Phi_k(Y^0_{k,j}, Y^1_{k-1,j}, Z^0)\|_{\mathcal{Y}}\big] < \infty$ *and*

$$\max\big\{\mathbb{E}\big[\|\psi_k(\Phi_0(\mathfrak{y}_0,\mathfrak{y}_{-1},Z^0))\|^2_{\mathcal{H}}\big], \mathbb{1}_{\mathbb{N}}(k)\,\mathbb{E}\big[\|\psi_k(\mathfrak{y}_0 - y)\|^2_{\mathcal{H}}\big]\big\} \leq \tfrac{C^2}{k!}, \tag{101}$$

$$\mathbb{E}\big[\|\psi_k(\Phi_n(u,v,Z^0))\|^2_{\mathcal{H}}\big] \leq c\,\mathbb{E}\big[\|\psi_{k+1}(u-v)\|^2_{\mathcal{H}}\big], \tag{102}$$

$$\mathbb{E}\left[\left\|\psi_k\left(y - \sum_{l=0}^{n-1} \mathbb{E}\big[\Phi_l\big(Y^0_{l,j}, Y^1_{l-1,j}, Z^0\big)\big]\right)\right\|^2_{\mathcal{H}}\right] \leq 2c\,\mathbb{E}\Big[\big\|\psi_{k+1}\big(Y^0_{n-1,j} - y\big)\big\|^2_{\mathcal{H}}\Big]. \tag{103}$$

Then

(i) it holds for all $n\in\mathbb{N}$ *that*

$$\big(\mathbb{E}\big[\|\psi_0(Y^0_{n,n} - y)\|^2_{\mathcal{H}}\big]\big)^{1/2} \leq C\left[\frac{e^{\kappa}(1+4c)}{M_n}\right]^{n/2} < \infty, \tag{104}$$

(ii) it holds for all $n\in\mathbb{N}$ *that* $\mathrm{Cost}_{n,n} \leq (3+\gamma)^n (M_n)^n \mathfrak{z}$*, and*

(iii) there exists $(N_\varepsilon)_{\varepsilon\in(0,1]} \subseteq \mathbb{N}$ *such that it holds for all* $\varepsilon\in(0,1]$, $\delta\in(0,\infty)$ *that* $\sup_{n\in\{N_\varepsilon, N_\varepsilon+1,\ldots\}} \big(\mathbb{E}\big[\|\psi_0(Y^0_{n,n} - y)\|^2_{\mathcal{H}}\big]\big)^{1/2} \leq \varepsilon$ *and*

$$\mathrm{Cost}_{N_\varepsilon,N_\varepsilon} \leq \mathfrak{z}(3+\gamma)e^{\kappa}C^{2(1+\delta)}\left(1 + \sup_{n\in\mathbb{N}}\left[\frac{[\mathfrak{B}e^{2\kappa}(3+\gamma)(1+4c)]^{(1+\delta)}}{(M_n)^{\delta}}\right]^n\right)\varepsilon^{-2(1+\delta)} < \infty. \tag{105}$$

Proof of Corollary 2.15. Throughout this proof let $(\mathfrak{c}_k)_{k\in\mathbb{N}_0} \subseteq (0,\infty)$ be the family of real numbers which satisfies for all $k\in\mathbb{N}_0$ that $\mathfrak{c}_k = k!$. Note that the assumption that $\sup_{j\in\mathbb{N}} {}^{M_j}/_j \leq \kappa$ ensures for all $n\in\mathbb{N}$ that

$$\max_{k\in\{0,1,\ldots,n\}} \tfrac{(M_n)^k}{\mathfrak{c}_k} = \max_{k\in\{0,1,\ldots,n\}} \tfrac{(M_n)^k}{k!} \leq \sum_{k=0}^{\infty} \tfrac{(M_n)^k}{k!} = e^{M_n} \leq e^{\kappa n} = (e^{\kappa})^n. \tag{106}$$

Next observe that (ii) in Proposition 2.8 implies for all $n,j\in\mathbb{N}$, $\theta\in\Theta$ that $\sigma_\Omega(Y^\theta_{n,j}) \subseteq \sigma_\Omega\big((Z^{(\theta,\vartheta)})_{\vartheta\in\Theta}\big)$. This demonstrates for all $n,j\in\mathbb{N}$ that $\sigma_\Omega\big(Y^0_{n,j}, Y^1_{n-1,j}\big) \subseteq \sigma_\Omega\big((Z^{(0,\theta)})_{\theta\in\Theta}, (Z^{(1,\theta)})_{\theta\in\Theta}\big)$. The fact that it holds for every $k\in\mathbb{N}_0$ that $\sigma_\Omega\big((Z^{(0,\theta)})_{\theta\in\Theta}, (Z^{(1,\theta)})_{\theta\in\Theta}\big)$, Z^0, and ψ_k are independent hence shows for every $k\in\mathbb{N}_0$, $n,j\in\mathbb{N}$ that

$$\sigma_\Omega\big(Y^0_{n,j}, Y^1_{n-1,j}\big) \quad \text{and} \quad \sigma_\Omega(\psi_k, Z^0) \tag{107}$$

are independent. Furthermore, the fact that it holds for all $k\in\mathbb{N}_0$, $x\in\mathcal{Y}$ that $\Omega\ni\omega\mapsto[\psi_k(\omega)](x)\in\mathcal{H}$ is an $\mathscr{F}/\mathscr{B}(\mathcal{H})$-measurable function (cf. Lemma 2.1) and the fact that it holds for all $k\in\mathbb{N}_0$, $\omega\in\Omega$ that $\mathcal{Y}\ni x\mapsto[\psi_k(\omega)](x)\in\mathcal{H}$ is a continuous function yield that

$$\mathcal{Y}\times\mathcal{Y}\times\Omega \ni (u,v,\omega) \mapsto [\psi_k(\omega)](u-v) \in \mathcal{H} \tag{108}$$

is a continuous random field. Moreover, note that the fact that $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ is separable ensures that $\mathscr{B}(\mathcal{Y})\otimes\mathscr{B}(\mathcal{Y}) = \mathscr{B}(\mathcal{Y}\times\mathcal{Y})$. This, (i) in Corollary 2.3, (107), [68, Lemma 2.2] (with $\mathcal{G} = \sigma_\Omega(\psi_k, Z^0)$, $(S,\mathcal{S}) = (\mathcal{Y}\times\mathcal{Y}, \mathscr{B}(\mathcal{Y})\otimes\mathscr{B}(\mathcal{Y}))$, $U = (\mathcal{Y}\times\mathcal{Y}\times\Omega \ni (u,v,\omega) \mapsto \|[\psi_k(\omega)](\Phi_n(u,v,Z^0(\omega)))\|_{\mathcal{H}}^2 \in [0,\infty))$, $Y = (Y^0_{n,j}, Y^1_{n-1,j})$ for $j,n\in\mathbb{N}$, $k\in\mathbb{N}_0$ in the notation of [68, Lemma 2.2]), (102), (108), and [68, Lemma 2.3] (with $S = \mathcal{Y}\times\mathcal{Y}$, $U = (\mathcal{Y}\times\mathcal{Y}\times\Omega \ni (u,v,\omega) \mapsto \|[\psi_k(\omega)](u-v)\|_{\mathcal{H}}^2 \in [0,\infty))$, $Y = (Y^0_{n,j}, Y^1_{n-1,j})$ for $j,n\in\mathbb{N}$, $k\in\mathbb{N}_0$ in the notation of [68, Lemma 2.3]) establish for all $k\in\mathbb{N}_0$, $n,j\in\mathbb{N}$ that

$$\begin{aligned}
&\mathbb{E}\Big[\big\|\psi_k\big(\Phi_n\big(Y^0_{n,j}, Y^1_{n-1,j}, Z^0\big)\big)\big\|_{\mathcal{H}}^2\Big] \\
&= \int_{\mathcal{Y}\times\mathcal{Y}} \mathbb{E}\big[\|\psi_k(\Phi_n(u,v,Z^0))\|_{\mathcal{H}}^2\big]\,\big(\big(Y^0_{n,j}, Y^1_{n-1,j}\big)(\mathbb{P})_{\mathscr{B}(\mathcal{Y})\otimes\mathscr{B}(\mathcal{Y})}\big)(\mathrm{d}u,\mathrm{d}v) \\
&= \int_{\mathcal{Y}\times\mathcal{Y}} \mathbb{E}\big[\|\psi_k(\Phi_n(u,v,Z^0))\|_{\mathcal{H}}^2\big]\,\big(\big(Y^0_{n,j}, Y^1_{n-1,j}\big)(\mathbb{P})_{\mathscr{B}(\mathcal{Y}\times\mathcal{Y})}\big)(\mathrm{d}u,\mathrm{d}v) \\
&\le c\int_{\mathcal{Y}\times\mathcal{Y}} \mathbb{E}\big[\|\psi_{k+1}(u-v)\|_{\mathcal{H}}^2\big]\,\big(\big(Y^0_{n,j}, Y^1_{n-1,j}\big)(\mathbb{P})_{\mathscr{B}(\mathcal{Y}\times\mathcal{Y})}\big)(\mathrm{d}u,\mathrm{d}v) \\
&= c\,\mathbb{E}\Big[\big\|\psi_{k+1}\big(Y^0_{n,j} - Y^1_{n-1,j}\big)\big\|_{\mathcal{H}}^2\Big].
\end{aligned} \tag{109}$$

Combining (106) and (109) with Theorem 2.14 shows (i)–(iii). The proof of Corollary 2.15 is thus complete. □

Lemma 2.16. *Let $\kappa\in[1,\infty)$, $(M_j)_{j\in\mathbb{N}}\subseteq\mathbb{N}$ satisfy for all $j\in\mathbb{N}$ that $M_j < M_{j+1}$ and $M_j \le \kappa j$. Then*

(i) it holds for all $j\in\mathbb{N}$ that $j \le M_j \le \kappa j$,

(ii) it holds that $\liminf_{j\to\infty} M_j = \infty$, and

(iii) it holds that $\sup_{j\in\mathbb{N}} {}^{M_{j+1}}/_{M_j} \le 2\kappa$.

Proof of Lemma 2.16. Note that the assumption that $\forall\, j\in\mathbb{N}\colon M_j < M_{j+1}$ and induction show (i). Next observe that (i) implies (ii). Furthermore, the assumption that $\forall\, j\in\mathbb{N}\colon M_j \le \kappa j$ and (i) ensure for all $j\in\mathbb{N}$ that

$$\frac{M_{j+1}}{M_j} \le \frac{\kappa(j+1)}{j} = \kappa + \frac{\kappa}{j} \le 2\kappa. \tag{110}$$

The proof of Lemma 2.16 is thus complete. □

3 MLP for semi-linear heat equations

In this section we employ the abstract framework for generalised MLP approximations developed in Section 2 to prove that appropriate MLP approximations, which essentially are generalised versions of the MLP approximations proposed in Hutzenthaler et al. [68], are able to overcome the curse of dimensionality in the numerical approximation of semi-linear heat equations (see Theorem 3.17 and Corollary 3.18 in Subsection 3.3.2 below).

In the context of applying the abstract complexity result about generalised MLP approximations in Corollary 2.15 above to numerical approximations for semi-linear heat equations, the separable $\mathbb{R}$-Banach space $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ in Corollary 2.15 is chosen to be a subspace of the vector space of real-valued at most polynomially growing continuous functions defined on $[0,T]\times\mathbb{R}^d$ equipped with a suitable polynomial growth norm, where $T\in(0,\infty)$, $d\in\mathbb{N}$ (see (138)–(139) below). In Subsection 3.1 we derive several elementary and well-known properties of these and related function spaces and their elements. In particular, Subsection 3.1.1 deals with completeness and separability of such function spaces. Lemma 3.1 recalls that the vector space of real-valued at most polynomially growing continuous functions defined on a non-empty subset of $\mathbb{R}^d$ equipped with an appropriate polynomial growth norm is complete. Thereafter, we state in Proposition 3.2 the well-known fact that the vector space of real-valued continuous functions defined on a non-empty compact subset of $\mathbb{R}^d$ equipped with the uniform norm is a separable $\mathbb{R}$-Banach space, which follows directly from Lemma 3.1 and, e.g., Conway [30, Theorem 6.6 in Chapter V]. Using Proposition 3.2 we deduce the elementary fact that also the vector space of real-valued continuous functions with compact support defined on a non-empty closed subset of $\mathbb{R}^d$ equipped with a suitable polynomial growth norm is separable (see Lemma 3.3). Subsection 3.1.1 is concluded by the well-known result in Proposition 3.4, which establishes a characterisation of the above mentioned choice for the vector space $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ (see (138)–(139) below) and shows that it is indeed a separable $\mathbb{R}$-Banach space. Subsequently, we provide in Lemmas 3.5–3.6 and Corollary 3.7 in Subsection 3.1.2 three elementary results about sufficient conditions under which suitable functions and suitable compositions of functions grow strictly slower than a given polynomial order. These results are used to ensure well-definedness of certain functions introduced in Subsection 3.2.1 (see (142) below). Furthermore, Lemma 3.8 in Subsection 3.1.3 offers an elementary polynomial growth estimate for suitable compositions of functions.

In Subsection 3.2 we specify a number of the objects appearing in Corollary 2.15 above for the example of MLP approximations for semi-linear heat equations and verify that the main assumptions of Corollary 2.15 are fulfilled in this context. In particular, we first present in Setting 3.1 in Subsection 3.2.1 the framework which we refer to throughout Subsection 3.2. In the subsection that follows, Subsection 3.2.2, we establish measurability properties of several of the involved functions (see Lemmas 3.9–3.10). Subsequently, Lemma 3.11 in Subsection 3.2.3 shows that the MLP approximations introduced in (137) in Setting 3.1 fit into the abstract framework for generalised MLP approximations developed in Section 2 (see (99) above). Moreover, Subsection 3.2.4 is devoted to proving certain integrability properties of the MLP approximations introduced in (137) in Setting 3.1 (see Lemma 3.12), while in Subsection 3.2.5 we verify that the estimates assumed in (101)–(103) in Corollary 2.15 hold true for the functions introduced in Setting 3.1 (see Lemmas 3.13–3.15).

Finally, in Subsection 3.3 we combine the results from Subsection 3.2 with Corollary 2.15 to obtain a complexity analysis for MLP approximations for semi-linear heat equations. In Proposition 3.16 in Subsection 3.3.1 this is done for semi-linear heat equations of fixed space dimension $d \in \mathbb{N}$ (cf. [68, Theorem 3.8]). Thereafter, Proposition 3.16 is used to establish Theorem 3.17 in Subsection 3.3.2, which reveals that the MLP approximations in (229) overcome the curse of dimensionality in the numerical approximation of semi-linear heat equations and which essentially is a slight generalisation of [68, Theorem 1.1]. The last result in this section, Corollary 3.18, is a direct consequence of Theorem 3.17 and describes the special case of Theorem 3.17 in which the non-linearity in the semi-linear heat equations is the same for every dimension (see (i) in Corollary 3.18) and in which the constants in the complexity estimate are not given explicitly (see (ii) in Corollary 3.18).

3.1 Properties of spaces of at most polynomially growing continuous functions

3.1.1 Completeness and separability

Lemma 3.1. *Let $d \in \mathbb{N}$, $p \in [0,\infty)$, let $\mathcal{A} \subseteq \mathbb{R}^d$ be a non-empty set, let $\mathcal{V} = \{v \in C(\mathcal{A},\mathbb{R})\colon \sup_{x\in\mathcal{A}} {|v(x)|}/{\max\{1,\|x\|_{\mathbb{R}^d}^p\}} < \infty\}$, and let $\|\cdot\|_{\mathcal{V}}\colon \mathcal{V} \to [0,\infty)$ satisfy for all $v \in \mathcal{V}$ that $\|v\|_{\mathcal{V}} = \sup_{x\in\mathcal{A}} {|v(x)|}/{\max\{1,\|x\|_{\mathbb{R}^d}^p\}}$. Then it holds that $(\mathcal{V},\|\cdot\|_{\mathcal{V}})$ is an $\mathbb{R}$-Banach space.*

Proof of Lemma 3.1. Observe that it holds that $(\mathcal{V},\|\cdot\|_{\mathcal{V}})$ is a normed $\mathbb{R}$-vector space. It thus remains to prove that $(\mathcal{V},\|\cdot\|_{\mathcal{V}})$ is complete. For this let $\mathcal{W} \subseteq C(\mathcal{A},\mathbb{R})$ be the set given by

$$\mathcal{W} = \{w \in C(\mathcal{A},\mathbb{R})\colon \sup_{x\in\mathcal{A}} |w(x)| < \infty\}, \tag{111}$$

let $\|\cdot\|_{\mathcal{W}}\colon \mathcal{W} \to [0,\infty)$ satisfy for all $w \in \mathcal{W}$ that $\|w\|_{\mathcal{W}} = \sup_{x\in\mathcal{A}} |w(x)|$, and let $I\colon \mathcal{V} \to \mathcal{W}$ and $J\colon \mathcal{W} \to \mathcal{V}$ satisfy for all $v \in \mathcal{V}$, $w \in \mathcal{W}$, $x \in \mathcal{A}$ that $[I(v)](x) = {v(x)}/{\max\{1,\|x\|_{\mathbb{R}^d}^p\}}$ and $[J(w)](x) = w(x)\max\{1,\|x\|_{\mathbb{R}^d}^p\}$. Note that $(\mathcal{W},\|\cdot\|_{\mathcal{W}})$ is a normed $\mathbb{R}$-vector space. Furthermore, Jentzen, Mazzonetto, & Salimova [72, Corollary 2.3] shows that

$$(\mathcal{W},\|\cdot\|_{\mathcal{W}}) \tag{112}$$

is complete. Next observe that it holds for all $v \in \mathcal{V}$ that

$$\|I(v)\|_{\mathcal{W}} = \sup_{x\in\mathcal{A}} |[I(v)](x)| = \sup_{x\in\mathcal{A}}\left[\frac{|v(x)|}{\max\{1,\|x\|_{\mathbb{R}^d}^p\}}\right] = \|v\|_{\mathcal{V}}. \tag{113}$$

In addition, note that it holds for all $w \in \mathcal{W}$, $x \in \mathcal{A}$ that

$$[I(J(w))](x) = \frac{[J(w)](x)}{\max\{1,\|x\|_{\mathbb{R}^d}^p\}} = \frac{w(x)\max\{1,\|x\|_{\mathbb{R}^d}^p\}}{\max\{1,\|x\|_{\mathbb{R}^d}^p\}} = w(x). \tag{114}$$

Combining this with (113) ensures that $I\colon \mathcal{V} \to \mathcal{W}$ is a bijective linear isometry and $I^{-1} = J$. This and (112) establish that $(\mathcal{V},\|\cdot\|_{\mathcal{V}}) = (I^{-1}(\mathcal{W}),\|\cdot\|_{\mathcal{V}})$ is complete and thus finish the proof of Lemma 3.1. ☐

Proposition 3.2. *Let $d \in \mathbb{N}$, let $\mathcal{A} \subseteq \mathbb{R}^d$ be a non-empty compact set, and let $\|\cdot\|_{C(\mathcal{A},\mathbb{R})}\colon C(\mathcal{A},\mathbb{R}) \to [0,\infty)$ satisfy for all $f \in C(\mathcal{A},\mathbb{R})$ that $\|f\|_{C(\mathcal{A},\mathbb{R})} = \sup_{x\in\mathcal{A}} |f(x)|$. Then it holds that $(C(\mathcal{A},\mathbb{R}),\|\cdot\|_{C(\mathcal{A},\mathbb{R})})$ is a separable $\mathbb{R}$-Banach space.*

Lemma 3.3. *Let $d \in \mathbb{N}$, $p \in [0,\infty)$, let $\mathcal{A} \subseteq \mathbb{R}^d$ be a non-empty closed set, and let $|||\cdot|||\colon C_c(\mathcal{A},\mathbb{R}) \to [0,\infty)$ satisfy for all $f \in C_c(\mathcal{A},\mathbb{R})$ that $|||f||| = \sup_{x\in\mathcal{A}} {|f(x)|}/{\max\{1,\|x\|_{\mathbb{R}^d}^p\}}$. Then it holds that $(C_c(\mathcal{A},\mathbb{R}),|||\cdot|||)$ is a separable normed $\mathbb{R}$-vector space.*

Proof of Lemma 3.3. Throughout this proof let $\mathbf{y} \in \mathcal{A}$, let $N = \min([\|\mathbf{y}\|_{\mathbb{R}^d},\infty) \cap \mathbb{N})$, let $\mathcal{S}_n \subseteq C_c(\mathcal{A},\mathbb{R})$, $n \in \{N, N+1,\ldots\}$, be the sets which satisfy for all $n \in \{N,N+1,\ldots\}$ that

$$\mathcal{S}_n = \{f \in C(\mathcal{A},\mathbb{R})\colon \{x \in \mathcal{A}\colon f(x) \neq 0\} \subseteq [-n,n]^d\}, \tag{115}$$

let $[\![\cdot]\!]_n\colon C(\mathcal{A}\cap[-n,n]^d,\mathbb{R}) \to [0,\infty)$, $n \in \{N,N+1,\ldots\}$, satisfy for all $n \in \{N,N+1,\ldots\}$, $f \in C(\mathcal{A}\cap[-n,n]^d,\mathbb{R})$ that

$$[\![f]\!]_n = \sup_{x\in\mathcal{A}\cap[-n,n]^d}\left[\frac{|f(x)|}{\max\{1,\|x\|_{\mathbb{R}^d}^p\}}\right], \tag{116}$$

and let $I_n\colon \mathcal{S}_n \to C(\mathcal{A}\cap[-n,n]^d,\mathbb{R})$, $n \in \{N, N+1, \dots\}$, satisfy for all $n \in \{N, N+1, \dots\}$, $f \in \mathcal{S}_n$ that $I_n(f) = f|_{\mathcal{A}\cap[-n,n]^d}$. Note that (115) proves for all $n \in \{N, N+1, \dots\}$, $f \in \mathcal{S}_n$ that

$$[\![I_n(f)]\!]_n = \sup_{x\in\mathcal{A}\cap[-n,n]^d} \left[\frac{|f(x)|}{\max\{1, \|x\|_{\mathbb{R}^d}^p\}}\right] = \sup_{x\in\mathcal{A}} \left[\frac{|f(x)|}{\max\{1, \|x\|_{\mathbb{R}^d}^p\}}\right] = |\!|\!| f |\!|\!|. \tag{117}$$

This and the fact that it holds for all $n \in \{N, N+1, \dots\}$ that $(\mathcal{S}_n, |\!|\!|\cdot|\!|\!||_{\mathcal{S}_n})$ and $(C(\mathcal{A}\cap[-n,n]^d,\mathbb{R}), [\![\cdot]\!]_n)$ are normed $\mathbb{R}$-vector spaces ensure for all $n \in \{N, N+1, \dots\}$ that

$$I_n\colon \mathcal{S}_n \to C(\mathcal{A}\cap[-n,n]^d,\mathbb{R}) \tag{118}$$

is a linear isometry. Next observe that it holds for all $n \in \{N, N+1, \dots\}$, $f \in C(\mathcal{A}\cap[-n,n]^d,\mathbb{R})$ that

$$[\![f]\!]_n \le \sup_{x\in\mathcal{A}\cap[-n,n]^d} |f(x)| \le \sup_{x\in\mathcal{A}\cap[-n,n]^d} \left[\frac{|f(x)|\big(n\sqrt{d}\big)^p}{\max\{1, \|x\|_{\mathbb{R}^d}^p\}}\right] = [\![f]\!]_n \big(n\sqrt{d}\big)^p. \tag{119}$$

In addition, the assumption that $\mathcal{A} \subseteq \mathbb{R}^d$ is a closed set and the fact that $\mathbf{y} \in \mathcal{A}$ ensure for all $n \in \{N, N+1, \dots\}$ that $\mathcal{A}\cap[-n,n]^d$ is a non-empty compact set. Proposition 3.2 and (119) hence show for all $n \in \{N, N+1, \dots\}$ that $(C(\mathcal{A}\cap[-n,n]^d,\mathbb{R}), [\![\cdot]\!]_n)$ is a separable $\mathbb{R}$-Banach space. This implies for all $n \in \{N, N+1, \dots\}$ that $(I_n(\mathcal{S}_n), [\![\cdot]\!]_n|_{I_n(\mathcal{S}_n)})$ is a separable normed $\mathbb{R}$-vector space. Combining this with (118) hence establishes for all $n \in \{N, N+1, \dots\}$ that

$$(\mathcal{S}_n, |\!|\!|\cdot|\!|\!||_{\mathcal{S}_n}) \tag{120}$$

is a separable normed $\mathbb{R}$-vector space. Furthermore, the assumption that $\mathcal{A} \subseteq \mathbb{R}^d$ is a closed set and (115) demonstrate that

$$\begin{aligned} C_{\mathrm{c}}(\mathcal{A},\mathbb{R}) &= \Big\{ f \in C(\mathcal{A},\mathbb{R})\colon \Big(\exists\, n \in \mathbb{N}\colon \overline{\{x \in \mathcal{A}\colon f(x) \neq 0\}}^{\mathbb{R}^d} \subseteq \mathcal{A}\cap[-n,n]^d \Big) \Big\} \\ &= \big\{ f \in C(\mathcal{A},\mathbb{R})\colon \big(\exists\, n \in \mathbb{N}\colon \{x \in \mathcal{A}\colon f(x) \neq 0\} \subseteq [-n,n]^d \big) \big\} \\ &= \big\{ f \in C(\mathcal{A},\mathbb{R})\colon \big(\exists\, n \in \{N, N+1, \dots\}\colon \{x \in \mathcal{A}\colon f(x) \neq 0\} \subseteq [-n,n]^d \big) \big\} \\ &= \bigcup_{n=N}^{\infty} \big\{ f \in C(\mathcal{A},\mathbb{R})\colon \{x \in \mathcal{A}\colon f(x) \neq 0\} \subseteq [-n,n]^d \big\} = \bigcup_{n=N}^{\infty} \mathcal{S}_n. \end{aligned} \tag{121}$$

This and (120) establish that $(C_{\mathrm{c}}(\mathcal{A},\mathbb{R}), |\!|\!|\cdot|\!|\!|)$ is a separable normed $\mathbb{R}$-vector space. The proof of Lemma 3.3 is thus complete. □

Proposition 3.4. *Let $d \in \mathbb{N}$, $T \in (0,\infty)$, $p \in [0,\infty)$, let $\mathcal{Y} = \big\{ y \in C([0,T]\times\mathbb{R}^d,\mathbb{R})\colon \limsup_{\mathbb{N}\ni n\to\infty} \sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\ge n} {}^{|y(t,x)|}\!/\!_{\|x\|_{\mathbb{R}^d}^p} = 0 \big\}$, and let $\|\cdot\|_{\mathcal{Y}}\colon \mathcal{Y} \to [0,\infty)$ satisfy for all $y \in \mathcal{Y}$ that $\|y\|_{\mathcal{Y}} = \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} {}^{|y(t,x)|}\!/\!_{\max\{1,\|x\|_{\mathbb{R}^d}^p\}}$. Then*

(i) it holds that $\mathcal{Y} = \overline{C_{\mathrm{c}}([0,T]\times\mathbb{R}^d,\mathbb{R})}^{\mathcal{Y}}$ and

(ii) it holds that $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ is a separable $\mathbb{R}$-Banach space.

Proof of Proposition 3.4. Throughout this proof let $\tau_n \in C(\mathbb{R}^d,[0,1])$, $n \in \mathbb{N}$, satisfy for all $n \in \mathbb{N}$, $x \in \mathbb{R}^d$ that

$$\tau_n(x) = \max\{\min\{n+1-\|x\|_{\mathbb{R}^d}, 1\}, 0\} = \begin{cases} 1 & \colon \|x\|_{\mathbb{R}^d} \le n \\ n+1-\|x\|_{\mathbb{R}^d} & \colon n \le \|x\|_{\mathbb{R}^d} \le n+1 \\ 0 & \colon n+1 \le \|x\|_{\mathbb{R}^d} \end{cases}, \tag{122}$$

let $\mathbf{y} \in \mathcal{Y}$, and let $y_n \in C([0,T] \times \mathbb{R}^d, \mathbb{R})$, $n \in \mathbb{N}$, satisfy for all $n \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that $y_n(t,x) = \tau_n(x)\,\mathbf{y}(t,x)$. Note that it holds that $(y_n)_{n\in\mathbb{N}} \subseteq C_c([0,T] \times \mathbb{R}^d, \mathbb{R}) \subseteq \mathcal{Y}$ and

$$\begin{aligned} \limsup_{\mathbb{N}\ni n\to\infty} \|\mathbf{y} - y_n\|_{\mathcal{Y}} &= \limsup_{\mathbb{N}\ni n\to\infty} \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \left[\frac{|\mathbf{y}(t,x) - \tau_n(x)\,\mathbf{y}(t,x)|}{\max\{1, \|x\|_{\mathbb{R}^d}^p\}}\right] \\ &= \limsup_{\mathbb{N}\ni n\to\infty} \sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq n} \left[\frac{(1-\tau_n(x))|\mathbf{y}(t,x)|}{\|x\|_{\mathbb{R}^d}^p}\right] \\ &\leq \limsup_{\mathbb{N}\ni n\to\infty} \sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq n} \frac{|\mathbf{y}(t,x)|}{\|x\|_{\mathbb{R}^d}^p} = 0. \end{aligned} \tag{123}$$

This proves that $\mathcal{Y} \subseteq \overline{C_c([0,T] \times \mathbb{R}^d, \mathbb{R})}^{\mathcal{Y}}$. In addition, observe that the fact that $\mathcal{Y} \supseteq C_c([0,T] \times \mathbb{R}^d, \mathbb{R})$ ensures that $\mathcal{Y} \supseteq \overline{C_c([0,T] \times \mathbb{R}^d, \mathbb{R})}^{\mathcal{Y}}$. This finishes the proof of (i). It thus remains to show (ii). For this let $\mathcal{V} \subseteq C([0,T] \times \mathbb{R}^d, \mathbb{R})$ be the set given by

$$\mathcal{V} = \left\{ v \in C([0,T] \times \mathbb{R}^d, \mathbb{R}) \colon \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \left[\tfrac{|v(t,x)|}{\max\{1, \|x\|_{\mathbb{R}^d}^p\}}\right] < \infty \right\}, \tag{124}$$

let $\|\cdot\|_{\mathcal{V}} \colon \mathcal{V} \to [0,\infty)$ satisfy for all $v \in \mathcal{V}$ that

$$\|v\|_{\mathcal{V}} = \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \left[\frac{|v(t,x)|}{\max\{1, \|x\|_{\mathbb{R}^d}^p\}}\right], \tag{125}$$

let $\mathbf{v} \in \mathcal{V}$, and let $(v_n)_{n\in\mathbb{N}} \subseteq \mathcal{Y} \subseteq \mathcal{V}$ be a sequence which satisfies $\limsup_{\mathbb{N}\ni n\to\infty} \|\mathbf{v} - v_n\|_{\mathcal{V}} = 0$. Note that this implies that

$$\begin{aligned} &\limsup_{\mathbb{N}\ni n\to\infty} \sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq n} \frac{|\mathbf{v}(t,x)|}{\|x\|_{\mathbb{R}^d}^p} \\ &\leq \limsup_{\mathbb{N}\ni m\to\infty} \limsup_{\mathbb{N}\ni n\to\infty} \sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq n} \frac{|\mathbf{v}(t,x) - v_m(t,x)|}{\|x\|_{\mathbb{R}^d}^p} \\ &\quad + \limsup_{\mathbb{N}\ni m\to\infty} \limsup_{\mathbb{N}\ni n\to\infty} \sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq n} \frac{|v_m(t,x)|}{\|x\|_{\mathbb{R}^d}^p} \\ &\leq \limsup_{\mathbb{N}\ni m\to\infty} \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \left[\frac{|\mathbf{v}(t,x) - v_m(t,x)|}{\max\{1, \|x\|_{\mathbb{R}^d}^p\}}\right] = \limsup_{\mathbb{N}\ni m\to\infty} \|\mathbf{v} - v_m\|_{\mathcal{V}} = 0. \end{aligned} \tag{126}$$

This establishes that $\mathbf{v} \in \mathcal{Y}$. Therefore, it holds that $\mathcal{Y} \subseteq \mathcal{V}$ is a closed set. The fact that $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$ is an $\mathbb{R}$-Banach space (cf. Lemma 3.1) hence demonstrates that $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}}) = (\mathcal{Y}, \|\cdot\|_{\mathcal{V}}|_{\mathcal{Y}})$ is an $\mathbb{R}$-Banach space. Moreover, note that the fact that $(C_c([0,T] \times \mathbb{R}^d, \mathbb{R}), \|\cdot\|_{\mathcal{Y}}|_{C_c([0,T]\times\mathbb{R}^d,\mathbb{R})})$ is a separable normed $\mathbb{R}$-vector space (cf. Lemma 3.3) and (i) assure that

$$(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}}) = \left(\overline{C_c([0,T] \times \mathbb{R}^d, \mathbb{R})}^{\mathcal{Y}}, \|\cdot\|_{\mathcal{Y}}|_{\overline{C_c([0,T]\times\mathbb{R}^d,\mathbb{R})}^{\mathcal{Y}}}\right) \tag{127}$$

is separable. This establishes (ii). The proof of Proposition 3.4 is thus complete. □

3.1.2 Sufficient conditions for strictly slower growth

Lemma 3.5. *Let $d \in \mathbb{N}$, $T \in (0,\infty)$, $p \in [0,\infty)$, $q \in (p,\infty)$ and let $y \in C([0,T] \times \mathbb{R}^d, \mathbb{R})$ satisfy $\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} {}^{|y(t,x)|}\!/\!_{\max\{1,\|x\|_{\mathbb{R}^d}^p\}} < \infty$. Then*

$$\limsup_{\mathbb{N}\ni n\to\infty} \sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq n} \frac{|y(t,x)|}{\|x\|_{\mathbb{R}^d}^q} = 0. \tag{128}$$

Proof of Lemma 3.5. Throughout this proof let $C \in [0,\infty)$ be the real number which satisfies $C = \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} {}^{|y(t,x)|}\!/\!_{\max\{1,\|x\|^p_{\mathbb{R}^d}\}}$. Observe that it holds that

$$\begin{aligned}
&\limsup_{\mathbb{N}\ni n\to\infty} \sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq n} \frac{|y(t,x)|}{\|x\|^q_{\mathbb{R}^d}} \\
&= \limsup_{\mathbb{N}\ni n\to\infty} \sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq n} \left[\frac{|y(t,x)|}{\max\{1,\|x\|^p_{\mathbb{R}^d}\}} \frac{\max\{1,\|x\|^p_{\mathbb{R}^d}\}}{\|x\|^q_{\mathbb{R}^d}}\right] \\
&\leq C \limsup_{\mathbb{N}\ni n\to\infty} \sup_{x\in\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq n} \left[\frac{\max\{1,\|x\|^p_{\mathbb{R}^d}\}}{\|x\|^q_{\mathbb{R}^d}}\right] \\
&= C \limsup_{\mathbb{N}\ni n\to\infty} \sup_{x\in\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq n} \frac{1}{\|x\|^{q-p}_{\mathbb{R}^d}} = C \limsup_{\mathbb{N}\ni n\to\infty} \frac{1}{n^{q-p}} = 0.
\end{aligned} \tag{129}$$

The proof of Lemma 3.5 is thus complete. □

Lemma 3.6. *Let $d\in\mathbb{N}$, $T,q\in(0,\infty)$, let $\mathcal{Y} = \{y\in C([0,T]\times\mathbb{R}^d,\mathbb{R})\colon \limsup_{\mathbb{N}\ni n\to\infty} \sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\,\|x\|_{\mathbb{R}^d}\geq n} {}^{|y(t,x)|}\!/\!_{\|x\|^q_{\mathbb{R}^d}} = 0\}$, let $\varrho = (\varrho_1,\varrho_2)\in C([0,T]\times\mathbb{R}^d,[0,T]\times\mathbb{R}^d)$ satisfy $\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} {}^{\|\varrho_2(t,x)\|_{\mathbb{R}^d}}\!/\!_{\max\{1,\|x\|_{\mathbb{R}^d}\}} < \infty$, and let $y\in\mathcal{Y}$. Then it holds that $y\circ\varrho\in\mathcal{Y}$.*

Proof of Lemma 3.6. Throughout this proof let $\varepsilon\in(0,\infty)$ and let $L,\mathfrak{N},N\in\mathbb{N}$ satisfy $\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} {}^{\|\varrho_2(t,x)\|_{\mathbb{R}^d}}\!/\!_{\max\{1,\|x\|_{\mathbb{R}^d}\}} \leq L$, $\sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\,\|x\|_{\mathbb{R}^d}\geq\mathfrak{N}} {}^{|y(t,x)|}\!/\!_{\|x\|^q_{\mathbb{R}^d}} \leq \frac{\varepsilon}{L^q}$, and $\varepsilon^{-1/q}\sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\,\|x\|_{\mathbb{R}^d}\leq\mathfrak{N}} |y(t,x)|^{1/q} \leq N$. Observe that it holds for all $t\in[0,T]$, $x\in\mathbb{R}^d$ with $\|x\|_{\mathbb{R}^d}\geq N$ and $\|\varrho_2(t,x)\|_{\mathbb{R}^d}\leq\mathfrak{N}$ that

$$|y(\varrho(t,x))| = |y(\varrho_1(t,x),\varrho_2(t,x))| \leq \sup_{(s,\mathbf{x})\in[0,T]\times\mathbb{R}^d,\,\|\mathbf{x}\|_{\mathbb{R}^d}\leq\mathfrak{N}} |y(s,\mathbf{x})| \leq \varepsilon N^q \leq \varepsilon\|x\|^q_{\mathbb{R}^d}. \tag{130}$$

In addition, note that it holds for all $t\in[0,T]$, $x\in\mathbb{R}^d$ with $\|x\|_{\mathbb{R}^d}\geq 1$ and $\|\varrho_2(t,x)\|_{\mathbb{R}^d}\geq\mathfrak{N}$ that

$$\begin{aligned}
|y(\varrho(t,x))| &\leq \left[\sup_{(s,\mathbf{x})\in[0,T]\times\mathbb{R}^d,\,\|\mathbf{x}\|_{\mathbb{R}^d}\geq\mathfrak{N}} \frac{|y(s,\mathbf{x})|}{\|\mathbf{x}\|^q_{\mathbb{R}^d}}\right] \|\varrho_2(t,x)\|^q_{\mathbb{R}^d} \\
&\leq \frac{\varepsilon}{L^q}\cdot L^q\max\{1,\|x\|^q_{\mathbb{R}^d}\} = \varepsilon\|x\|^q_{\mathbb{R}^d}.
\end{aligned} \tag{131}$$

This and (130) establish for all $n\in\{N,N+1,\ldots\}$ that

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\,\|x\|_{\mathbb{R}^d}\geq n} \frac{|y(\varrho(t,x))|}{\|x\|^q_{\mathbb{R}^d}} \leq \sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\,\|x\|_{\mathbb{R}^d}\geq N} \frac{|y(\varrho(t,x))|}{\|x\|^q_{\mathbb{R}^d}} \leq \varepsilon. \tag{132}$$

The fact that $y\circ\varrho\in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ thus completes the proof of Lemma 3.6. □

Corollary 3.7. *Let $d\in\mathbb{N}$, $T\in(0,\infty)$, $L,p\in[0,\infty)$, $q\in(p,\infty)$, let $\mathcal{Y} = \{y\in C([0,T]\times\mathbb{R}^d,\mathbb{R})\colon \limsup_{\mathbb{N}\ni n\to\infty}\sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\,\|x\|_{\mathbb{R}^d}\geq n} {}^{|y(t,x)|}\!/\!_{\|x\|^q_{\mathbb{R}^d}} = 0\}$, and let $\varrho = (\varrho_1,\varrho_2)\in C([0,T],[0,T]\times\mathbb{R}^d)$, $f\in C([0,T]\times\mathbb{R}^d\times\mathbb{R},\mathbb{R})$, $y\in\mathcal{Y}$ satisfy for all $t\in[0,T]$, $x\in\mathbb{R}^d$, $v,w\in\mathbb{R}$ that $|f(t,x,0)|\leq L\max\{1,\|x\|^p_{\mathbb{R}^d}\}$ and $|f(t,x,v)-f(t,x,w)|\leq L|v-w|$. Then*

(i) it holds that $([0,T]\times\mathbb{R}^d\ni(t,x)\mapsto y(\varrho_1(t),x+\varrho_2(t))\in\mathbb{R})\in\mathcal{Y}$ and

(ii) it holds that $([0,T]\times\mathbb{R}^d\ni(t,x)\mapsto f(\varrho_1(t),x+\varrho_2(t),y(t,x))\in\mathbb{R})\in\mathcal{Y}$.

Proof of Corollary 3.7. Note that it holds that

$$\begin{aligned}\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\left[\frac{\|x+\varrho_2(t)\|_{\mathbb{R}^d}}{\max\{1,\|x\|_{\mathbb{R}^d}\}}\right] &\le \sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\left[\frac{\|x\|_{\mathbb{R}^d}}{\max\{1,\|x\|_{\mathbb{R}^d}\}}+\frac{\|\varrho_2(t)\|_{\mathbb{R}^d}}{\max\{1,\|x\|_{\mathbb{R}^d}\}}\right]\\ &\le 1+\sup_{t\in[0,T]}\|\varrho_2(t)\|_{\mathbb{R}^d}<\infty.\end{aligned} \tag{133}$$

Lemma 3.6 (with $d = d$, $T = T$, $q = q$, $\mathcal{Y} = \mathcal{Y}$, $\varrho = ([0,T]\times\mathbb{R}^d \ni (t,x) \mapsto (\varrho_1(t), x+\varrho_2(t)) \in [0,T]\times\mathbb{R}^d)$, $y = y$ in the notation of Lemma 3.6) hence shows that $([0,T]\times\mathbb{R}^d \ni (t,x) \mapsto y(\varrho_1(t), x+\varrho_2(t)) \in \mathbb{R}) \in \mathcal{Y}$. This proves (i). Next observe that Lemma 3.5 (with $d = d$, $T = T$, $p = p$, $q = q$, $y = ([0,T]\times\mathbb{R}^d \ni (t,x) \mapsto f(t,x,0) \in \mathbb{R})$ in the notation of Lemma 3.5) ensures that $([0,T]\times\mathbb{R}^d \ni (t,x) \mapsto f(t,x,0) \in \mathbb{R}) \in \mathcal{Y}$. Combining this with (i) implies that $([0,T]\times\mathbb{R}^d \ni (t,x) \mapsto f(\varrho_1(t), x+\varrho_2(t), 0) \in \mathbb{R}) \in \mathcal{Y}$. Therefore, we obtain that

$$\begin{aligned}&\limsup_{\mathbb{N}\ni n\to\infty}\sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\,\|x\|_{\mathbb{R}^d}\ge n}\left[\frac{|f(\varrho_1(t),x+\varrho_2(t),y(t,x))|}{\|x\|_{\mathbb{R}^d}^q}\right]\\ &\le \limsup_{\mathbb{N}\ni n\to\infty}\sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\,\|x\|_{\mathbb{R}^d}\ge n}\left[\frac{|f(\varrho_1(t),x+\varrho_2(t),0)|}{\|x\|_{\mathbb{R}^d}^q}\right]\\ &\quad+\limsup_{\mathbb{N}\ni n\to\infty}\sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\,\|x\|_{\mathbb{R}^d}\ge n}\left[\frac{|f(\varrho_1(t),x+\varrho_2(t),y(t,x))-f(\varrho_1(t),x+\varrho_2(t),0)|}{\|x\|_{\mathbb{R}^d}^q}\right]\\ &\le L\limsup_{\mathbb{N}\ni n\to\infty}\sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\,\|x\|_{\mathbb{R}^d}\ge n}\frac{|y(t,x)|}{\|x\|_{\mathbb{R}^d}^q}=0.\end{aligned} \tag{134}$$

This and the fact that $([0,T]\times\mathbb{R}^d \ni (t,x) \mapsto f(\varrho_1(t), x+\varrho_2(t), y(t,x)) \in \mathbb{R}) \in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ establish (ii). The proof of Corollary 3.7 is thus complete. □

3.1.3 Growth estimate for compositions

Lemma 3.8. *Let $d \in \mathbb{N}$, $T \in (0,\infty)$, $p \in [0,\infty)$, $L \in [1,\infty)$, let $[\![\cdot]\!]\colon C([0,T]\times\mathbb{R}^d,\mathbb{R}) \to [0,\infty]$ satisfy for all $v \in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ that $[\![v]\!] = \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} {}^{|v(t,x)|}/_{\max\{1,\|x\|_{\mathbb{R}^d}^p\}}$, let $\varrho = (\varrho_1,\varrho_2) \in C([0,T]\times\mathbb{R}^d, [0,T]\times\mathbb{R}^d)$ satisfy for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that $\|\varrho_2(t,x)\|_{\mathbb{R}^d} \le L\max\{1,\|x\|_{\mathbb{R}^d}\}$, and let $v \in C([0,T]\times\mathbb{R}^d,\mathbb{R})$. Then it holds that $v\circ\varrho \in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ and $[\![v\circ\varrho]\!] \le L^p[\![v]\!]$.*

Proof of Lemma 3.8. Observe that it holds that $v\circ\varrho \in C([0,T]\times\mathbb{R}^d,\mathbb{R})$. In addition, note that it holds that

$$\begin{aligned}[\![v\circ\varrho]\!] &= \sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\left[\frac{|v(\varrho(t,x))|}{\max\{1,\|x\|_{\mathbb{R}^d}^p\}}\right] \le [\![v]\!]\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\left[\frac{\max\{1,\|\varrho_2(t,x)\|_{\mathbb{R}^d}^p\}}{\max\{1,\|x\|_{\mathbb{R}^d}^p\}}\right]\\ &\le [\![v]\!]\sup_{x\in\mathbb{R}^d}\left[\frac{\max\bigl\{1,L^p\max\{1,\|x\|_{\mathbb{R}^d}^p\}\bigr\}}{\max\{1,\|x\|_{\mathbb{R}^d}^p\}}\right] = L^p[\![v]\!].\end{aligned} \tag{135}$$

The proof of Lemma 3.8 is thus complete. □

3.2 Verification of the assumed properties

3.2.1 Setting

Setting 3.1. *Let $d \in \mathbb{N}$, $\xi \in \mathbb{R}^d$, $T \in (0,\infty)$, $L, p \in [0,\infty)$, $q \in (p,\infty)$, $\Theta = \cup_{n=1}^\infty \mathbb{Z}^n$, $(M_j)_{j\in\mathbb{N}} \subseteq \mathbb{N}$, let $(\Omega,\mathscr{F},\mathbb{P})$ be a probability space, let $\mathbf{U}\colon \Omega \to [0,1]$ and $U^\theta\colon \Omega \to [0,1]$,*

$\theta \in \Theta$, be on $[0,1]$ uniformly distributed random variables, let $\mathbf{W}\colon [0,T]\times\Omega \to \mathbb{R}^d$ and $W^\theta\colon [0,T]\times\Omega\to\mathbb{R}^d$, $\theta\in\Theta$, be standard Brownian motions with continuous sample paths, assume that (U^θ, W^θ), $\theta \in \Theta$, are independent, assume that $\mathbf{U}$, $\mathbf{W}$, $(U^\theta)_{\theta\in\Theta}$, and $(W^\theta)_{\theta\in\Theta}$ are independent, let $f \in C([0,T]\times\mathbb{R}^d\times\mathbb{R},\mathbb{R})$, $g\in C(\mathbb{R}^d,\mathbb{R})$, $y \in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ satisfy for all $t\in[0,T]$, $x\in\mathbb{R}^d$, $v,w\in\mathbb{R}$ that $\max\{|f(t,x,0)|,|g(x)|\} \leq L\max\{1,\|x\|_{\mathbb{R}^d}^p\}$, $|f(t,x,v)-f(t,x,w)| \leq L|v-w|$, $\sup_{(s,\mathbf{x})\in[0,T]\times\mathbb{R}^d} {}^{|y(s,\mathbf{x})|}\!/\!{}_{\max\{1,\|\mathbf{x}\|_{\mathbb{R}^d}^p\}} < \infty$, and

$$y(t,x) = \mathbb{E}\bigg[g(x+\mathbf{W}_{T-t}) + \int_t^T f\big(s, x+\mathbf{W}_{s-t}, y(s,x+\mathbf{W}_{s-t})\big)\,\mathrm{d}s\bigg], \tag{136}$$

let $Y^\theta_{n,j}\colon [0,T]\times\mathbb{R}^d\times\Omega\to\mathbb{R}$, $\theta\in\Theta$, $j\in\mathbb{N}$, $n\in(\mathbb{N}_0\cup\{-1\})$, satisfy for all $n,j\in\mathbb{N}$, $\theta\in\Theta$, $t\in[0,T]$, $x\in\mathbb{R}^d$ that $Y^\theta_{-1,j}(t,x) = Y^\theta_{0,j}(t,x) = 0$ and

$$\begin{aligned} Y^\theta_{n,j}(T-t,x) &= \tfrac{1}{(M_j)^n}\Bigg[\sum_{i=1}^{(M_j)^n} g\big(x+W_t^{(\theta,0,i)}\big)\Bigg] + \sum_{l=0}^{n-1}\tfrac{t}{(M_j)^{n-l}}\Bigg[\sum_{i=1}^{(M_j)^{n-l}} \\ &\Big[f\Big(T-t+U^{(\theta,l,i)}t, x+W^{(\theta,l,i)}_{U^{(\theta,l,i)}t}, Y^{(\theta,l,i)}_{l,j}\big(T-t+U^{(\theta,l,i)}t, x+W^{(\theta,l,i)}_{U^{(\theta,l,i)}t}\big)\Big) \\ &- \mathbb{1}_{\mathbb{N}}(l) f\Big(T-t+U^{(\theta,l,i)}t, x+W^{(\theta,l,i)}_{U^{(\theta,l,i)}t}, Y^{(\theta,-l,i)}_{l-1,j}\big(T-t+U^{(\theta,l,i)}t, x+W^{(\theta,l,i)}_{U^{(\theta,l,i)}t}\big)\Big)\Big]\Bigg], \end{aligned} \tag{137}$$

let $\mathcal{Y}\subseteq C([0,T]\times\mathbb{R}^d,\mathbb{R})$ be the set given by

$$\mathcal{Y} = \left\{ v \in C([0,T]\times\mathbb{R}^d,\mathbb{R})\colon \limsup_{\mathbb{N}\ni n\to\infty}\ \sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq n} \frac{|v(t,x)|}{\|x\|^q_{\mathbb{R}^d}} = 0\right\}, \tag{138}$$

let $\|\cdot\|_{\mathcal{Y}}\colon \mathcal{Y}\to[0,\infty)$ satisfy for all $v\in\mathcal{Y}$ that

$$\|v\|_{\mathcal{Y}} = \sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\left[\frac{|v(t,x)|}{\max\{1,\|x\|^q_{\mathbb{R}^d}\}}\right], \tag{139}$$

let $\mathcal{Z} = [0,1]\times C([0,T],\mathbb{R}^d)$, let $\mathbf{d}_{\mathcal{Z}}\colon \mathcal{Z}\times\mathcal{Z}\to[0,\infty)$ satisfy for all $\mathfrak{z} = (\mathfrak{u},\mathfrak{w}), \mathfrak{Z} = (\mathfrak{U},\mathfrak{W}) \in \mathcal{Z}$ that

$$\mathbf{d}_{\mathcal{Z}}(\mathfrak{z},\mathfrak{Z}) = |\mathfrak{u}-\mathfrak{U}| + \|\mathfrak{w}-\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)} = |\mathfrak{u}-\mathfrak{U}| + \sup_{t\in[0,T]}\|\mathfrak{w}(t)-\mathfrak{W}(t)\|_{\mathbb{R}^d}, \tag{140}$$

let $Z^\theta\colon\Omega\to\mathcal{Z}$, $\theta\in\Theta$, satisfy for all $\theta\in\Theta$ that $Z^\theta = (U^\theta, W^\theta)$, let $\psi_k\colon\Omega\to\mathcal{Y}^*$, $k\in\mathbb{N}_0$, satisfy for all $k\in\mathbb{N}_0$, $\omega\in\Omega$, $v\in\mathcal{Y}$ that

$$[\psi_k(\omega)](v) = \begin{cases} v(0,\xi) & \colon k=0 \\ \sqrt{\frac{(\mathbf{U}(\omega))^{k-1}}{(k-1)!}}\, v\big(\mathbf{U}(\omega)T, \xi+\mathbf{W}_{\mathbf{U}(\omega)T}(\omega)\big) & \colon k\in\mathbb{N} \end{cases}, \tag{141}$$

and let $\Phi_l\colon \mathcal{Y}\times\mathcal{Y}\times\mathcal{Z}\to\mathcal{Y}$, $l\in\mathbb{N}_0$, satisfy for all $l\in\mathbb{N}_0$, $v,w\in\mathcal{Y}$, $\mathfrak{z}=(\mathfrak{u},\mathfrak{w})\in\mathcal{Z}$, $t\in[0,T]$, $x\in\mathbb{R}^d$ that

$$\begin{aligned} &[\Phi_l(v,w,\mathfrak{z})](T-t,x) \\ &= \begin{cases} g(x+\mathfrak{w}_t) + tf\big(T-t+\mathfrak{u}t, x+\mathfrak{w}_{\mathfrak{u}t}, v(T-t+\mathfrak{u}t, x+\mathfrak{w}_{\mathfrak{u}t})\big) & \colon l=0 \\ t\big[f\big(T-t+\mathfrak{u}t, x+\mathfrak{w}_{\mathfrak{u}t}, v(T-t+\mathfrak{u}t,x+\mathfrak{w}_{\mathfrak{u}t})\big) & \\ \quad - f\big(T-t+\mathfrak{u}t, x+\mathfrak{w}_{\mathfrak{u}t}, w(T-t+\mathfrak{u}t,x+\mathfrak{w}_{\mathfrak{u}t})\big)\big] & \colon l\in\mathbb{N} \end{cases} \end{aligned} \tag{142}$$

(cf. Lemma 3.5 and Corollary 3.7).

3.2.2 Measurability

Lemma 3.9. *Assume Setting 3.1 and let* $\mathscr{S} = \sigma_{\mathcal{Y}^*}\big(\{\{\varphi \in \mathcal{Y}^* \colon \varphi(v) \in \mathcal{B}\} \subseteq \mathcal{Y}^* \colon v \in \mathcal{Y},\, \mathcal{B} \in \mathscr{B}(\mathbb{R})\}\big)$. *Then it holds for all* $k \in \mathbb{N}_0$ *that* $\psi_k \colon \Omega \to \mathcal{Y}^*$ *is an* $\mathscr{F}/\mathscr{S}$*-measurable function.*

Proof of Lemma 3.9. Note that it holds for all $k \in \mathbb{N}_0$, $v \in \mathcal{Y}$ that $\Omega \ni \omega \mapsto [\psi_k(\omega)](v) \in \mathbb{R}$ is an $\mathscr{F}/\mathscr{B}(\mathbb{R})$-measurable function. Lemma 2.1 (with $\mathcal{E} = \mathcal{Y}$, $(\mathcal{F}, \mathscr{F}) = (\mathbb{R}, \mathscr{B}(\mathbb{R}))$, $(\mathcal{G}, \mathscr{G}) = (\Omega, \mathscr{F})$, $\mathcal{S} = \mathcal{Y}^*$, $\mathscr{S} = \mathscr{S}$, $\psi = \psi_k$ for $k \in \mathbb{N}_0$ in the notation of Lemma 2.1) hence proves for all $k \in \mathbb{N}_0$ that $\psi_k \colon \Omega \to \mathcal{Y}^*$ is an $\mathscr{F}/\mathscr{S}$-measurable function. The proof of Lemma 3.9 is thus complete. ☐

Lemma 3.10. *Assume Setting 3.1. Then*

(i) it holds for all $l \in \mathbb{N}_0$ *that* $\Phi_l \colon \mathcal{Y} \times \mathcal{Y} \times \mathcal{Z} \to \mathcal{Y}$ *is a continuous function and*

(ii) it holds for all $l \in \mathbb{N}_0$ *that* $\Phi_l \colon \mathcal{Y} \times \mathcal{Y} \times \mathcal{Z} \to \mathcal{Y}$ *is a* $(\mathscr{B}(\mathcal{Y}) \otimes \mathscr{B}(\mathcal{Y}) \otimes \mathscr{Z})/\mathscr{B}(\mathcal{Y})$*-measurable function.*

Proof of Lemma 3.10. Throughout this proof let $\varphi_1 \colon \mathcal{Z} \to C([0,T], \mathbb{R}^d)$, $\varphi_2, \varphi_3, F \colon \mathcal{Y} \times \mathcal{Z} \to \mathcal{Y}$, $\varphi_4 \colon \mathcal{Y} \to \mathcal{Y}$, $\Psi_1, \Psi_2 \colon \mathcal{Y} \times \mathcal{Z} \to \mathcal{Y} \times \mathcal{Z}$, $\mathfrak{g} \in \mathcal{Y}$, $G \colon \mathcal{Z} \to \mathcal{Y}$ satisfy for all $v, w \in \mathcal{Y}$, $\mathfrak{z} = (\mathfrak{u}, \mathfrak{w}) \in \mathcal{Z}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$[\varphi_1(\mathfrak{z})](t) = \mathfrak{w}_{\mathfrak{u}t}, \qquad \Psi_1(v, \mathfrak{z}) = (v, \mathfrak{u}, \varphi_1(\mathfrak{z})), \tag{143}$$

$$[\varphi_2(v, \mathfrak{z})](t,x) = v(T - t + \mathfrak{u}t, x + \mathfrak{w}_t), \qquad \Psi_2(v, \mathfrak{z}) = (\varphi_2(v, \mathfrak{z}), \mathfrak{z}), \tag{144}$$

$$[\varphi_3(v, \mathfrak{z})](t,x) = tf\big(T - t + \mathfrak{u}t, x + \mathfrak{w}_t, v(t,x)\big), \qquad \mathfrak{g}(t,x) = g(x), \tag{145}$$

$$[\varphi_4(v)](t,x) = v(T-t, x), \qquad G(\mathfrak{z}) = \varphi_4(\varphi_2(\mathfrak{g}, \mathfrak{z})), \tag{146}$$

and $F = \varphi_4 \circ \varphi_3 \circ \Psi_2 \circ \Psi_1$ (cf. Corollary 3.7). Note that it holds for all $v \in \mathcal{Y}$, $\mathfrak{z} = (\mathfrak{u}, \mathfrak{w}) \in \mathcal{Z}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$\begin{aligned} [G(\mathfrak{z})](T-t, x) &= [\varphi_4(\varphi_2(\mathfrak{g}, \mathfrak{z}))](T-t, x) = [\varphi_2(\mathfrak{g}, \mathfrak{z})](t,x) \\ &= \mathfrak{g}(T - t + \mathfrak{u}t, x + \mathfrak{w}_t) = g(x + \mathfrak{w}_t) \end{aligned} \tag{147}$$

and

$$\begin{aligned} [F(v, \mathfrak{z})](T-t, x) &= \big[\varphi_4\big((\varphi_3 \circ \Psi_2 \circ \Psi_1)(v, \mathfrak{z})\big)\big](T-t, x) \\ &= [(\varphi_3 \circ \Psi_2 \circ \Psi_1)(v, \mathfrak{z})](t,x) = \big[(\varphi_3 \circ \Psi_2)\big(v, \mathfrak{u}, \varphi_1(\mathfrak{z})\big)\big](t,x) \\ &= \big[\varphi_3\big(\varphi_2\big(v, \mathfrak{u}, \varphi_1(\mathfrak{z})\big), \mathfrak{u}, \varphi_1(\mathfrak{z})\big)\big](t,x) \\ &= tf\big(T - t + \mathfrak{u}t, x + [\varphi_1(\mathfrak{z})](t), \big[\varphi_2\big(v, \mathfrak{u}, \varphi_1(\mathfrak{z})\big)\big](t,x)\big) \\ &= tf\big(T - t + \mathfrak{u}t, x + [\varphi_1(\mathfrak{z})](t), v\big(T - t + \mathfrak{u}t, x + [\varphi_1(\mathfrak{z})](t)\big)\big) \\ &= tf\big(T - t + \mathfrak{u}t, x + \mathfrak{w}_{\mathfrak{u}t}, v(T - t + \mathfrak{u}t, x + \mathfrak{w}_{\mathfrak{u}t})\big). \end{aligned} \tag{148}$$

Combining (147)–(148) with (142) ensures for all $l \in \mathbb{N}$, $v, w \in \mathcal{Y}$, $\mathfrak{z} \in \mathcal{Z}$ that

$$\Phi_0(v, w, \mathfrak{z}) = G(\mathfrak{z}) + F(v, \mathfrak{z}) \qquad \text{and} \qquad \Phi_l(v, w, \mathfrak{z}) = F(v, \mathfrak{z}) - F(w, \mathfrak{z}). \tag{149}$$

In the following we establish that $G \colon \mathcal{Z} \to \mathcal{Y}$ and $F \colon \mathcal{Y} \times \mathcal{Z} \to \mathcal{Y}$ are continuous functions.

First, we show that $\varphi_1 \colon \mathcal{Z} \to C([0,T], \mathbb{R}^d)$ is a continuous function. Throughout this paragraph let $\varepsilon \in (0, \infty)$, $\mathfrak{Z} = (\mathfrak{U}, \mathfrak{W}) \in \mathcal{Z}$ and let $\Delta, \delta \in (0, \infty)$ be real numbers which satisfy $\sup_{s,t\in[0,T],\,|s-t|\le\Delta}\|\mathfrak{W}_s-\mathfrak{W}_t\|_{\mathbb{R}^d}\le\frac{\varepsilon}{2}$ and $\delta=\min\{\frac{\Delta}{T},\frac{\varepsilon}{2}\}$. Observe that it holds for all $\mathfrak{z}=(\mathfrak{u},\mathfrak{w})\in\mathcal{Z}$ with $\mathbf{d}_{\mathcal{Z}}(\mathfrak{z},\mathfrak{Z})=|\mathfrak{u}-\mathfrak{U}|+\|\mathfrak{w}-\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)}\le\delta$ that

$$\begin{aligned}
&\|\varphi_1(\mathfrak{z})-\varphi_1(\mathfrak{Z})\|_{C([0,T],\mathbb{R}^d)}=\sup_{t\in[0,T]}\|\mathfrak{w}_{\mathfrak{u}t}-\mathfrak{W}_{\mathfrak{U}t}\|_{\mathbb{R}^d}\\
&\le\left[\sup_{t\in[0,T]}\|\mathfrak{w}_{\mathfrak{u}t}-\mathfrak{W}_{\mathfrak{u}t}\|_{\mathbb{R}^d}\right]+\left[\sup_{t\in[0,T]}\|\mathfrak{W}_{\mathfrak{u}t}-\mathfrak{W}_{\mathfrak{U}t}\|_{\mathbb{R}^d}\right]\\
&\le\left[\sup_{t\in[0,T]}\|\mathfrak{w}_{t}-\mathfrak{W}_{t}\|_{\mathbb{R}^d}\right]+\left[\sup_{s,t\in[0,T],\,|s-t|\le T\delta}\|\mathfrak{W}_{s}-\mathfrak{W}_{t}\|_{\mathbb{R}^d}\right]\\
&\le\|\mathfrak{w}-\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)}+\left[\sup_{s,t\in[0,T],\,|s-t|\le\Delta}\|\mathfrak{W}_{s}-\mathfrak{W}_{t}\|_{\mathbb{R}^d}\right]\\
&\le\delta+\tfrac{\varepsilon}{2}\le\tfrac{\varepsilon}{2}+\tfrac{\varepsilon}{2}=\varepsilon.
\end{aligned}\tag{150}$$

It thus holds that $\varphi_1\colon\mathcal{Z}\to C([0,T],\mathbb{R}^d)$ is a continuous function. Note that this ensures that $\Psi_1\colon\mathcal{Y}\times\mathcal{Z}\to\mathcal{Y}\times\mathcal{Z}$ is a continuous function.

Second, we claim that $\varphi_2\colon\mathcal{Y}\times\mathcal{Z}\to\mathcal{Y}$ is a continuous function. Throughout this paragraph let $\varepsilon\in(0,\infty)$, $\mathbf{v}\in\mathcal{Y}$, $\mathfrak{Z}=(\mathfrak{U},\mathfrak{W})\in\mathcal{Z}$ and let $N\in\mathbb{N}$, $R,\Delta,\delta\in(0,\infty)$ be real numbers which satisfy

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\,\|x\|_{\mathbb{R}^d}\ge N}\frac{|\mathbf{v}(t,x)|}{\|x\|_{\mathbb{R}^d}^q}\le\frac{\varepsilon}{12+6\,\|\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)}},\tag{151}$$

$$R=1+\|\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)},\qquad\sup_{\substack{(s,\mathbf{x}),(t,x)\in[0,T]\times\{w\in\mathbb{R}^d:\,\|w\|_{\mathbb{R}^d}\le N+2R\},\\|s-t|+\|\mathbf{x}-x\|_{\mathbb{R}^d}\le\Delta}}|\mathbf{v}(s,\mathbf{x})-\mathbf{v}(t,x)|\le\tfrac{\varepsilon}{3},\tag{152}$$

and

$$\delta=\min\left\{1,\frac{\Delta}{\max\{1,T\}},\frac{\varepsilon}{3\,(2+\|\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)})^q}\right\}.\tag{153}$$

Note that it holds for all $\mathfrak{w}\in C([0,T],\mathbb{R}^d)$, $t\in[0,T]$, $x\in\mathbb{R}^d$ with $\|\mathfrak{w}-\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)}\le1$ that

$$\begin{aligned}
\|x+\mathfrak{w}_t\|_{\mathbb{R}^d}&\le\|x\|_{\mathbb{R}^d}+\|\mathfrak{w}_t\|_{\mathbb{R}^d}\le\big(1+\|\mathfrak{w}\|_{C([0,T],\mathbb{R}^d)}\big)\max\{1,\|x\|_{\mathbb{R}^d}\}\\
&\le\big(1+\|\mathfrak{w}-\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)}+\|\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)}\big)\max\{1,\|x\|_{\mathbb{R}^d}\}\\
&\le\big(2+\|\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)}\big)\max\{1,\|x\|_{\mathbb{R}^d}\}.
\end{aligned}\tag{154}$$

This and Lemma 3.8 (with $d=d$, $T=T$, $p=q$, $L=2+\|\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)}$, $\varrho=([0,T]\times\mathbb{R}^d\ni(t,x)\mapsto(T-t+\mathfrak{u}t,x+\mathfrak{w}_t)\in[0,T]\times\mathbb{R}^d)$, $v=v-\mathbf{v}$ for $(\mathfrak{u},\mathfrak{w})\in\mathcal{Z}$, $v\in\mathcal{Y}$ with $\|\mathfrak{w}-\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)}\le1$ in the notation of Lemma 3.8) imply for all $v\in\mathcal{Y}$, $(\mathfrak{u},\mathfrak{w})\in\mathcal{Z}$ with $\|\mathfrak{w}-\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)}\le1$ that

$$\begin{aligned}
&\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\left[\frac{|v(T-t+\mathfrak{u}t,x+\mathfrak{w}_t)-\mathbf{v}(T-t+\mathfrak{u}t,x+\mathfrak{w}_t)|}{\max\{1,\|x\|_{\mathbb{R}^d}^q\}}\right]\\
&\le\big(2+\|\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)}\big)^q\|v-\mathbf{v}\|_{\mathcal{Y}}.
\end{aligned}\tag{155}$$

In addition, observe that it holds for all $\mathfrak{w}\in C([0,T],\mathbb{R}^d)$, $t\in[0,T]$, $x\in\mathbb{R}^d$ with $\|\mathfrak{w}-\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)}\le1$ and $\|x\|_{\mathbb{R}^d}\ge N+R$ that

$$\|x+\mathfrak{w}_t\|_{\mathbb{R}^d}\ge\|x\|_{\mathbb{R}^d}-\|\mathfrak{w}_t\|_{\mathbb{R}^d}\ge N+1+\|\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)}-\|\mathfrak{w}\|_{C([0,T],\mathbb{R}^d)}\ge N.\tag{156}$$

This, (154), and (151) establish for all $v \in \mathcal{Y}$, $(\mathfrak{u}, \mathfrak{w}) \in \mathcal{Z}$ with $\|\mathfrak{w} - \mathfrak{W}\|_{C([0,T],\mathbb{R}^d)} \leq 1$ that

$$\begin{aligned}
&\sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq N+R} \left[\frac{|\mathbf{v}(T-t+\mathfrak{u}t, x+\mathfrak{w}_t)|}{\|x\|_{\mathbb{R}^d}^q}\right] \\
&= \sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq N+R} \left[\frac{\|x+\mathfrak{w}_t\|_{\mathbb{R}^d}^q}{\max\{1, \|x\|_{\mathbb{R}^d}^q\}} \frac{|\mathbf{v}(T-t+\mathfrak{u}t, x+\mathfrak{w}_t)|}{\|x+\mathfrak{w}_t\|_{\mathbb{R}^d}^q}\right] \\
&\leq \left[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \frac{\|x+\mathfrak{w}_t\|_{\mathbb{R}^d}^q}{\max\{1, \|x\|_{\mathbb{R}^d}^q\}}\right]\left[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq N} \frac{|\mathbf{v}(t,x)|}{\|x\|_{\mathbb{R}^d}^q}\right] \\
&\leq \big(2 + \|\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)}\big) \frac{\varepsilon}{12 + 6\,\|\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)}} = \tfrac{\varepsilon}{6}.
\end{aligned} \tag{157}$$

Furthermore, note that it holds for all $\mathfrak{z} = (\mathfrak{u}, \mathfrak{w}) \in \mathcal{Z}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ with $\mathbf{d}_{\mathcal{Z}}(\mathfrak{z}, \mathfrak{Z}) = |\mathfrak{u} - \mathfrak{U}| + \|\mathfrak{w} - \mathfrak{W}\|_{C([0,T],\mathbb{R}^d)} \leq \delta$ and $\|x\|_{\mathbb{R}^d} \leq N + R$ that

$$\begin{aligned}
\|x + \mathfrak{w}_t\|_{\mathbb{R}^d} &\leq \|x\|_{\mathbb{R}^d} + \|\mathfrak{w}_t\|_{\mathbb{R}^d} \leq N + R + \|\mathfrak{w}\|_{C([0,T],\mathbb{R}^d)} \\
&\leq N + R + \|\mathfrak{w} - \mathfrak{W}\|_{C([0,T],\mathbb{R}^d)} + \|\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)} \\
&\leq N + R + 1 + \|\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)} = N + 2R
\end{aligned} \tag{158}$$

and

$$\begin{aligned}
&|T - t + \mathfrak{u}t - (T - t + \mathfrak{U}t)| + \|x + \mathfrak{w}_t - (x + \mathfrak{W}_t)\|_{\mathbb{R}^d} = |\mathfrak{u} - \mathfrak{U}|t + \|\mathfrak{w}_t - \mathfrak{W}_t\|_{\mathbb{R}^d} \\
&\leq |\mathfrak{u} - \mathfrak{U}|T + \|\mathfrak{w} - \mathfrak{W}\|_{C([0,T],\mathbb{R}^d)} \leq \max\{1, T\}\delta.
\end{aligned} \tag{159}$$

Combining (158)–(159) with (155), (157), (153), and (152) ensures for all $v \in \mathcal{Y}$, $\mathfrak{z} = (\mathfrak{u}, \mathfrak{w}) \in \mathcal{Z}$ with $\|v - \mathbf{v}\|_{\mathcal{Y}} + \mathbf{d}_{\mathcal{Z}}(\mathfrak{z}, \mathfrak{Z}) \leq \delta$ that

$$\begin{aligned}
&\|\varphi_2(v, \mathfrak{z}) - \varphi_2(\mathbf{v}, \mathfrak{Z})\|_{\mathcal{Y}} \\
&= \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \left[\frac{|v(T-t+\mathfrak{u}t, x+\mathfrak{w}_t) - \mathbf{v}(T-t+\mathfrak{U}t, x+\mathfrak{W}_t)|}{\max\{1, \|x\|_{\mathbb{R}^d}^q\}}\right] \\
&\leq \left[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \frac{|v(T-t+\mathfrak{u}t, x+\mathfrak{w}_t) - \mathbf{v}(T-t+\mathfrak{u}t, x+\mathfrak{w}_t)|}{\max\{1, \|x\|_{\mathbb{R}^d}^q\}}\right] \\
&\quad + \left[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\leq N+R} \frac{|\mathbf{v}(T-t+\mathfrak{u}t, x+\mathfrak{w}_t) - \mathbf{v}(T-t+\mathfrak{U}t, x+\mathfrak{W}_t)|}{\max\{1, \|x\|_{\mathbb{R}^d}^q\}}\right] \\
&\quad + \left[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq N+R} \frac{|\mathbf{v}(T-t+\mathfrak{u}t, x+\mathfrak{w}_t) - \mathbf{v}(T-t+\mathfrak{U}t, x+\mathfrak{W}_t)|}{\max\{1, \|x\|_{\mathbb{R}^d}^q\}}\right] \\
&\leq \big(2 + \|\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)}\big)^q \|v - \mathbf{v}\|_{\mathcal{Y}} + \sup_{\substack{(s,\mathbf{x}),(t,x)\in[0,T]\times\{w\in\mathbb{R}^d\colon \|w\|_{\mathbb{R}^d}\leq N+2R\},\\ |s-t|+\|\mathbf{x}-x\|_{\mathbb{R}^d}\leq\max\{1,T\}\delta}} |\mathbf{v}(s,\mathbf{x}) - \mathbf{v}(t,x)| \\
&\quad + \sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq N+R} \left[\frac{|\mathbf{v}(T-t+\mathfrak{u}t, x+\mathfrak{w}_t)|}{\|x\|_{\mathbb{R}^d}^q} + \frac{|\mathbf{v}(T-t+\mathfrak{U}t, x+\mathfrak{W}_t)|}{\|x\|_{\mathbb{R}^d}^q}\right] \\
&\leq \big(2 + \|\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)}\big)^q \delta + \tfrac{2\varepsilon}{6} + \sup_{\substack{(s,\mathbf{x}),(t,x)\in[0,T]\times\{w\in\mathbb{R}^d\colon \|w\|_{\mathbb{R}^d}\leq N+2R\},\\ |s-t|+\|\mathbf{x}-x\|_{\mathbb{R}^d}\leq\Delta}} |\mathbf{v}(s,\mathbf{x}) - \mathbf{v}(t,x)| \\
&\leq \tfrac{\varepsilon}{3} + \tfrac{\varepsilon}{3} + \tfrac{\varepsilon}{3} = \varepsilon.
\end{aligned} \tag{160}$$

This proves that $\varphi_2\colon \mathcal{Y} \times \mathcal{Z} \to \mathcal{Y}$ is a continuous function. Observe that this implies that $\Psi_2\colon \mathcal{Y} \times \mathcal{Z} \to \mathcal{Y} \times \mathcal{Z}$ is a continuous function.

Third, we establish that $\varphi_3\colon \mathcal{Y} \times \mathcal{Z} \to \mathcal{Y}$ is a continuous function. Throughout this paragraph let $\varepsilon \in (0,\infty)$, $\mathbf{v} \in \mathcal{Y}$, $\mathfrak{Z} = (\mathfrak{U}, \mathfrak{W}) \in \mathcal{Z}$ and let $N \in \mathbb{N}$, $R, \Delta, \delta \in (0,\infty)$ be real numbers which satisfy $N \geq (6LT(2 + \|\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)})^p \varepsilon^{-1})^{1/(q-p)}$ and

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq N} \left[\frac{L|\mathbf{v}(t,x)| + |f(T-t+\mathfrak{U}t, x+\mathfrak{W}_t, \mathbf{v}(t,x))|}{\|x\|_{\mathbb{R}^d}^q}\right] \leq \tfrac{\varepsilon}{6T}, \tag{161}$$

$$R = 1 + \|\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)}, \quad \sup_{\substack{(s,\mathbf{x}),(t,x)\in[0,T]\times\{w\in\mathbb{R}^d\colon \|w\|_{\mathbb{R}^d}\leq N+R\},\\ \mathfrak{v}\in\mathbb{R},\, |\mathfrak{v}|\leq\|\mathbf{v}\|_{\mathcal{Y}}N^q,\, |s-t|+\|\mathbf{x}-x\|_{\mathbb{R}^d}\leq\Delta}} |f(s,\mathbf{x},\mathfrak{v}) - f(t,x,\mathfrak{v})| \leq \tfrac{\varepsilon}{3T}, \tag{162}$$

and

$$\delta = \min\left\{1, \frac{\Delta}{\max\{1,T\}}, \frac{\varepsilon}{3\max\{1,LT\}}\right\} \tag{163}$$

(cf. (ii) in Corollary 3.7). Note that it holds for all $v \in \mathcal{Y}$, $\mathfrak{z} = (\mathfrak{u}, \mathfrak{w}) \in \mathcal{Z}$ that

$$\begin{aligned}
&\|\varphi_3(v,\mathfrak{z}) - \varphi_3(\mathbf{v},\mathfrak{Z})\|_{\mathcal{Y}} \\
&= \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \left[\frac{t|f(T-t+\mathfrak{u}t, x+\mathfrak{w}_t, v(t,x)) - f(T-t+\mathfrak{U}t, x+\mathfrak{W}_t, \mathbf{v}(t,x))|}{\max\{1, \|x\|_{\mathbb{R}^d}^q\}}\right] \\
&\leq T \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \left[\frac{|f(T-t+\mathfrak{u}t, x+\mathfrak{w}_t, v(t,x)) - f(T-t+\mathfrak{u}t, x+\mathfrak{w}_t, \mathbf{v}(t,x))|}{\max\{1, \|x\|_{\mathbb{R}^d}^q\}}\right] \\
&\quad + T \sup_{\substack{(t,x)\in[0,T]\times\mathbb{R}^d,\\ \|x\|_{\mathbb{R}^d}\leq N}} \left[\frac{|f(T-t+\mathfrak{u}t, x+\mathfrak{w}_t, \mathbf{v}(t,x)) - f(T-t+\mathfrak{U}t, x+\mathfrak{W}_t, \mathbf{v}(t,x))|}{\max\{1, \|x\|_{\mathbb{R}^d}^q\}}\right] \\
&\quad + T \sup_{\substack{(t,x)\in[0,T]\times\mathbb{R}^d,\\ \|x\|_{\mathbb{R}^d}\geq N}} \left[\frac{|f(T-t+\mathfrak{u}t, x+\mathfrak{w}_t, \mathbf{v}(t,x)) - f(T-t+\mathfrak{U}t, x+\mathfrak{W}_t, \mathbf{v}(t,x))|}{\|x\|_{\mathbb{R}^d}^q}\right].
\end{aligned} \tag{164}$$

Next observe that it holds for all $\mathfrak{z} = (\mathfrak{u}, \mathfrak{w}) \in \mathcal{Z}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ with $\mathbf{d}_{\mathcal{Z}}(\mathfrak{z},\mathfrak{Z}) = |\mathfrak{u} - \mathfrak{U}| + \|\mathfrak{w} - \mathfrak{W}\|_{C([0,T],\mathbb{R}^d)} \leq \delta$ and $\|x\|_{\mathbb{R}^d} \leq N$ that $\|x + \mathfrak{w}_t\|_{\mathbb{R}^d} \leq N + R$, $|\mathbf{v}(t,x)| \leq \|\mathbf{v}\|_{\mathcal{Y}} \max\{1, \|x\|_{\mathbb{R}^d}^q\} \leq \|\mathbf{v}\|_{\mathcal{Y}} N^q$, and

$$|T - t + \mathfrak{u}t - (T - t + \mathfrak{U}t)| + \|x + \mathfrak{w}_t - (x + \mathfrak{W}_t)\|_{\mathbb{R}^d} \leq \max\{1,T\}\delta. \tag{165}$$

This and (162) show for all $\mathfrak{z} = (\mathfrak{u}, \mathfrak{w}) \in \mathcal{Z}$ with $\mathbf{d}_{\mathcal{Z}}(\mathfrak{z},\mathfrak{Z}) \leq \delta$ that

$$\begin{aligned}
&\sup_{\substack{(t,x)\in[0,T]\times\mathbb{R}^d,\\ \|x\|_{\mathbb{R}^d}\leq N}} \left[\frac{|f(T-t+\mathfrak{u}t, x+\mathfrak{w}_t, \mathbf{v}(t,x)) - f(T-t+\mathfrak{U}t, x+\mathfrak{W}_t, \mathbf{v}(t,x))|}{\max\{1, \|x\|_{\mathbb{R}^d}^q\}}\right] \\
&\leq \sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\leq N} |f(T-t+\mathfrak{u}t, x+\mathfrak{w}_t, \mathbf{v}(t,x)) - f(T-t+\mathfrak{U}t, x+\mathfrak{W}_t, \mathbf{v}(t,x))| \\
&\leq \sup_{\substack{(s,\mathbf{x}),(t,x)\in[0,T]\times\{w\in\mathbb{R}^d\colon \|w\|_{\mathbb{R}^d}\leq N+R\},\\ \mathfrak{v}\in\mathbb{R},\, |\mathfrak{v}|\leq\|\mathbf{v}\|_{\mathcal{Y}}N^q,\, |s-t|+\|\mathbf{x}-x\|_{\mathbb{R}^d}\leq\max\{1,T\}\delta}} |f(s,\mathbf{x},\mathfrak{v}) - f(t,x,\mathfrak{v})| \\
&\leq \sup_{\substack{(s,\mathbf{x}),(t,x)\in[0,T]\times\{w\in\mathbb{R}^d\colon \|w\|_{\mathbb{R}^d}\leq N+R\},\\ \mathfrak{v}\in\mathbb{R},\, |\mathfrak{v}|\leq\|\mathbf{v}\|_{\mathcal{Y}}N^q,\, |s-t|+\|\mathbf{x}-x\|_{\mathbb{R}^d}\leq\Delta}} |f(s,\mathbf{x},\mathfrak{v}) - f(t,x,\mathfrak{v})| \leq \tfrac{\varepsilon}{3T}.
\end{aligned} \tag{166}$$

Furthermore, (161) and (154) ensure for all $(\mathfrak{u},\mathfrak{w}) \in \mathcal{Z}$ with $\|\mathfrak{w} - \mathfrak{W}\|_{C([0,T],\mathbb{R}^d)} \leq 1$ that

$$\begin{aligned}
&\sup_{\substack{(t,x)\in[0,T]\times\mathbb{R}^d,\\ \|x\|_{\mathbb{R}^d}\geq N}} \left[\frac{|f(T-t+\mathfrak{u}t, x+\mathfrak{w}_t, \mathbf{v}(t,x)) - f(T-t+\mathfrak{U}t, x+\mathfrak{W}_t, \mathbf{v}(t,x))|}{\|x\|_{\mathbb{R}^d}^q}\right] \\
&\leq \sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq N} \left[\frac{|f(T-t+\mathfrak{u}t, x+\mathfrak{w}_t, \mathbf{v}(t,x)) - f(T-t+\mathfrak{u}t, x+\mathfrak{w}_t, 0)|}{\|x\|_{\mathbb{R}^d}^q}\right. \\
&\qquad\qquad \left. + \frac{|f(T-t+\mathfrak{u}t, x+\mathfrak{w}_t, 0) - f(T-t+\mathfrak{U}t, x+\mathfrak{W}_t, \mathbf{v}(t,x))|}{\|x\|_{\mathbb{R}^d}^q}\right] \\
&\leq \left[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq N} \frac{L|\mathbf{v}(t,x)| + |f(T-t+\mathfrak{U}t, x+\mathfrak{W}_t, \mathbf{v}(t,x))|}{\|x\|_{\mathbb{R}^d}^q}\right] \\
&\quad + \left[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq N} \frac{|f(T-t+\mathfrak{u}t, x+\mathfrak{w}_t, 0)|}{\|x\|_{\mathbb{R}^d}^q}\right] \\
&\leq \tfrac{\varepsilon}{6T} + L\left[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq N} \frac{\max\{1, \|x+\mathfrak{w}_t\|_{\mathbb{R}^d}^p\}}{\|x\|_{\mathbb{R}^d}^q}\right] \\
&\leq \tfrac{\varepsilon}{6T} + L\big(2 + \|\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)}\big)^p \left[\sup_{x\in\mathbb{R}^d,\, \|x\|_{\mathbb{R}^d}\geq N} \frac{1}{\|x\|_{\mathbb{R}^d}^{q-p}}\right] \\
&= \tfrac{\varepsilon}{6T} + L\big(2 + \|\mathfrak{W}\|_{C([0,T],\mathbb{R}^d)}\big)^p \frac{1}{N^{q-p}} \leq \tfrac{\varepsilon}{6T} + \tfrac{\varepsilon}{6T} = \tfrac{\varepsilon}{3T}.
\end{aligned} \tag{167}$$

Combining (164) with (166), (167), and (163) establishes for all $v \in \mathcal{Y}$, $\mathfrak{z} = (\mathfrak{u},\mathfrak{w}) \in \mathcal{Z}$ with $\|v - \mathbf{v}\|_{\mathcal{Y}} + \mathbf{d}_{\mathcal{Z}}(\mathfrak{z}, \mathfrak{Z}) \leq \delta$ that

$$\|\varphi_3(v,\mathfrak{z}) - \varphi_3(\mathbf{v},\mathfrak{Z})\|_{\mathcal{Y}} \leq LT\|v - \mathbf{v}\|_{\mathcal{Y}} + \tfrac{2\varepsilon}{3} \leq LT\delta + \tfrac{2\varepsilon}{3} \leq \tfrac{\varepsilon}{3} + \tfrac{2\varepsilon}{3} = \varepsilon. \tag{168}$$

From this it follows that $\varphi_3\colon \mathcal{Y}\times\mathcal{Z} \to \mathcal{Y}$ is a continuous function.

As a next step observe that the fact that φ_2, Ψ_1, Ψ_2, and φ_3 are continuous functions, the fact that $\mathcal{Z} \ni \mathfrak{z} \mapsto (\mathfrak{g},\mathfrak{z}) \in \mathcal{Y}\times\mathcal{Z}$ is a continuous function, and the fact that $\varphi_4\colon \mathcal{Y}\to\mathcal{Y}$ is a linear isometry demonstrate that $G\colon \mathcal{Z}\to\mathcal{Y}$ and $F\colon \mathcal{Y}\times\mathcal{Z}\to\mathcal{Y}$ are continuous functions. Combining this with (149) proves (i). Finally, the fact that $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ is a separable $\mathbb{R}$-Banach space (cf. (ii) in Proposition 3.4), the fact that $(\mathcal{Z}, \mathbf{d}_{\mathcal{Z}})$ is a separable metric space, and (i) establish (ii). The proof of Lemma 3.10 is thus complete. □

3.2.3 Recursive formulation

Lemma 3.11. *Assume Setting 3.1. Then*

(i) it holds for all $n \in (\mathbb{N}_0 \cup \{-1\})$, $j \in \mathbb{N}$, $\theta \in \Theta$ that $Y^\theta_{n,j}(\Omega) \subseteq \mathcal{Y}$,

(ii) it holds for all $n, j \in \mathbb{N}$, $\theta\in\Theta$ that $Y^\theta_{-1,j} = Y^\theta_{0,j} = 0$ and

$$Y^\theta_{n,j} = \sum_{l=0}^{n-1} \tfrac{1}{(M_j)^{n-l}} \left[\sum_{i=1}^{(M_j)^{n-l}} \Phi_l\big(Y^{(\theta,l,i)}_{l,j}, Y^{(\theta,-l,i)}_{l-1,j}, Z^{(\theta,l,i)}\big)\right], \tag{169}$$

and

(iii) it holds for all $n \in (\mathbb{N}_0\cup\{-1\})$, $j\in\mathbb{N}$, $\theta\in\Theta$ that $\Omega \ni \omega \mapsto Y^\theta_{n,j}(\omega) \in \mathcal{Y}$ is an $\mathscr{F}/\mathscr{B}(\mathcal{Y})$-measurable function.

Proof of Lemma 3.11. We show (i)–(ii) by induction on $n \in \mathbb{N}$. For the base case $n = 1$ note that the fact that $\forall\, j \in \mathbb{N}$, $\theta \in \Theta$: $Y^\theta_{-1,j} = Y^\theta_{0,j} = 0$ implies for all $j \in \mathbb{N}$, $\theta \in \Theta$ that

$$Y^\theta_{-1,j}, Y^\theta_{0,j} \in \mathcal{Y}. \tag{170}$$

Next observe that (137) and (142) ensure for all $j \in \mathbb{N}$, $\theta \in \Theta$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$\begin{aligned} Y^\theta_{1,j}(T-t,x) &= \tfrac{1}{M_j}\left[\sum_{i=1}^{M_j} g\big(x + W^{(\theta,0,i)}_t\big)\right] \\ &\quad + \tfrac{t}{M_j}\left[\sum_{i=1}^{M_j} f\Big(T-t+U^{(\theta,0,i)}t, x + W^{(\theta,0,i)}_{U^{(\theta,0,i)}t}, Y^{(\theta,0,i)}_{0,j}\big(T-t+U^{(\theta,0,i)}t, x+W^{(\theta,0,i)}_{U^{(\theta,0,i)}t}\big)\Big)\right] \\ &= \tfrac{1}{M_j}\left[\sum_{i=1}^{M_j}\big[\Phi_0\big(Y^{(\theta,0,i)}_{0,j}, Y^{(\theta,0,i)}_{-1,j}, Z^{(\theta,0,i)}\big)\big](T-t,x)\right]. \end{aligned} \tag{171}$$

This and (170) prove (i)–(ii) in the base case $n = 1$. For the induction step $\mathbb{N} \ni n-1 \to n \in \{2,3,\dots\}$ let $n \in \{2,3,\dots\}$ and assume for all $l \in \{-1,0,1,\dots,n-1\}$, $j \in \mathbb{N}$, $\theta \in \Theta$ that $Y^\theta_{l,j}(\Omega) \subseteq \mathcal{Y}$. Equations (137) and (142) hence demonstrate for all $j \in \mathbb{N}$, $\theta \in \Theta$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$\begin{aligned} Y^\theta_{n,j}(T-t,x) &= \tfrac{1}{(M_j)^n}\left[\sum_{i=1}^{(M_j)^n} g\big(x+W^{(\theta,0,i)}_t\big)\right] + \sum_{l=0}^{n-1}\tfrac{t}{(M_j)^{n-l}}\left[\sum_{i=1}^{(M_j)^{n-l}}\right. \\ &\quad \Big[f\Big(T-t+U^{(\theta,l,i)}t, x+W^{(\theta,l,i)}_{U^{(\theta,l,i)}t}, Y^{(\theta,l,i)}_{l,j}\big(T-t+U^{(\theta,l,i)}t, x+W^{(\theta,l,i)}_{U^{(\theta,l,i)}t}\big)\Big) \\ &\quad \left. - \mathbb{1}_{\mathbb{N}}(l) f\Big(T-t+U^{(\theta,l,i)}t, x+W^{(\theta,l,i)}_{U^{(\theta,l,i)}t}, Y^{(\theta,-l,i)}_{l-1,j}\big(T-t+U^{(\theta,l,i)}t, x+W^{(\theta,l,i)}_{U^{(\theta,l,i)}t}\big)\Big)\Big]\right] \\ &= \tfrac{1}{(M_j)^n}\left[\sum_{i=1}^{(M_j)^n} g\big(x+W^{(\theta,0,i)}_t\big)\right] \\ &\quad + \tfrac{t}{(M_j)^n}\left[\sum_{i=1}^{(M_j)^n} f\Big(T-t+U^{(\theta,0,i)}t, x+W^{(\theta,0,i)}_{U^{(\theta,0,i)}t}, Y^{(\theta,0,i)}_{0,j}\big(T-t+U^{(\theta,0,i)}t, x+W^{(\theta,0,i)}_{U^{(\theta,0,i)}t}\big)\Big)\right] \\ &\quad + \sum_{l=1}^{n-1}\tfrac{t}{(M_j)^{n-l}}\left[\sum_{i=1}^{(M_j)^{n-l}}\right. \\ &\qquad \Big[f\Big(T-t+U^{(\theta,l,i)}t, x+W^{(\theta,l,i)}_{U^{(\theta,l,i)}t}, Y^{(\theta,l,i)}_{l,j}\big(T-t+U^{(\theta,l,i)}t, x+W^{(\theta,l,i)}_{U^{(\theta,l,i)}t}\big)\Big) \\ &\qquad \left. - f\Big(T-t+U^{(\theta,l,i)}t, x+W^{(\theta,l,i)}_{U^{(\theta,l,i)}t}, Y^{(\theta,-l,i)}_{l-1,j}\big(T-t+U^{(\theta,l,i)}t, x+W^{(\theta,l,i)}_{U^{(\theta,l,i)}t}\big)\Big)\Big]\right] \\ &= \tfrac{1}{(M_j)^n}\left[\sum_{i=1}^{(M_j)^n}\big[\Phi_0\big(Y^{(\theta,0,i)}_{0,j}, Y^{(\theta,0,i)}_{-1,j}, Z^{(\theta,0,i)}\big)\big](T-t,x)\right] \\ &\quad + \sum_{l=1}^{n-1}\tfrac{1}{(M_j)^{n-l}}\left[\sum_{i=1}^{(M_j)^{n-l}}\big[\Phi_l\big(Y^{(\theta,l,i)}_{l,j}, Y^{(\theta,-l,i)}_{l-1,j}, Z^{(\theta,l,i)}\big)\big](T-t,x)\right] \\ &= \sum_{l=0}^{n-1}\tfrac{1}{(M_j)^{n-l}}\left[\sum_{i=1}^{(M_j)^{n-l}}\big[\Phi_l\big(Y^{(\theta,l,i)}_{l,j}, Y^{(\theta,-l,i)}_{l-1,j}, Z^{(\theta,l,i)}\big)\big](T-t,x)\right]. \end{aligned} \tag{172}$$

Induction hence establishes (i)–(ii).

Furthermore, combining (i)–(ii) with (ii) in Lemma 3.10 and (i) in Proposition 2.8 shows (iii). The proof of Lemma 3.11 is thus complete. □

3.2.4 Integrability

Lemma 3.12. *Assume Setting 3.1. Then it holds for all $l \in \mathbb{N}_0$, $j \in \mathbb{N}$, $r \in [0, \infty)$ that*

$$\mathbb{E}\Big[\big\|\Phi_l\big(Y^0_{l,j}, Y^1_{l-1,j}, Z^0\big)\big\|^r_{\mathcal{Y}} + \big\|Y^0_{l-1,j}\big\|^r_{\mathcal{Y}}\Big] < \infty \tag{173}$$

(cf. (iii) in Lemma 3.11).

Proof of Lemma 3.12. First of all, note that it holds for all $(\mathfrak{u}, \mathfrak{w}) \in \mathcal{Z}$, $t \in [0, T]$, $x \in \mathbb{R}^d$ that

$$\|x + \mathfrak{w}_{\mathfrak{u}t}\|_{\mathbb{R}^d} \leq \|x\|_{\mathbb{R}^d} + \|\mathfrak{w}_{\mathfrak{u}t}\|_{\mathbb{R}^d} \leq \left(1 + \sup_{s\in[0,T]} \|\mathfrak{w}_s\|_{\mathbb{R}^d}\right) \max\{1, \|x\|_{\mathbb{R}^d}\}. \tag{174}$$

This and Lemma 3.8 (with $d = d$, $T = T$, $p = p$, $L = 1 + \sup_{t\in[0,T]}\|\mathfrak{w}_t\|_{\mathbb{R}^d}$, $\varrho = ([0, T] \times \mathbb{R}^d \ni (t, x) \mapsto (t, x + \mathfrak{w}_t) \in [0, T] \times \mathbb{R}^d)$, $v = ([0, T] \times \mathbb{R}^d \ni (t, x) \mapsto g(x) \in \mathbb{R})$ for $\mathfrak{w} \in C([0, T], \mathbb{R}^d)$ in the notation of Lemma 3.8) show for all $\mathfrak{w} \in C([0, T], \mathbb{R}^d)$ that

$$\begin{aligned}
&\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\left[\frac{|g(x + \mathfrak{w}_t)|}{\max\{1, \|x\|^q_{\mathbb{R}^d}\}}\right] \leq \sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\left[\frac{|g(x + \mathfrak{w}_t)|}{\max\{1, \|x\|^p_{\mathbb{R}^d}\}}\right]\\
&\leq \left(1 + \sup_{t\in[0,T]}\|\mathfrak{w}_t\|_{\mathbb{R}^d}\right)^p \sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\left[\frac{|g(x)|}{\max\{1, \|x\|^p_{\mathbb{R}^d}\}}\right] \leq L\left(1 + \sup_{t\in[0,T]}\|\mathfrak{w}_t\|_{\mathbb{R}^d}\right)^p.
\end{aligned} \tag{175}$$

Similarly, (174) and Lemma 3.8 (with $d = d$, $T = T$, $p = p$, $L = 1 + \sup_{t\in[0,T]}\|\mathfrak{w}_t\|_{\mathbb{R}^d}$, $\varrho = ([0, T] \times \mathbb{R}^d \ni (t, x) \mapsto (T - t + \mathfrak{u}t, x + \mathfrak{w}_{\mathfrak{u}t}) \in [0, T] \times \mathbb{R}^d)$, $v = ([0, T] \times \mathbb{R}^d \ni (t, x) \mapsto tf(t, x, 0) \in \mathbb{R})$ for $(\mathfrak{u}, \mathfrak{w}) \in \mathcal{Z}$ in the notation of Lemma 3.8) ensure for all $(\mathfrak{u}, \mathfrak{w}) \in \mathcal{Z}$ that

$$\begin{aligned}
&\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\left[\frac{|tf(T - t + \mathfrak{u}t, x + \mathfrak{w}_{\mathfrak{u}t}, 0)|}{\max\{1, \|x\|^q_{\mathbb{R}^d}\}}\right] \leq \sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\left[\frac{|tf(T - t + \mathfrak{u}t, x + \mathfrak{w}_{\mathfrak{u}t}, 0)|}{\max\{1, \|x\|^p_{\mathbb{R}^d}\}}\right]\\
&\leq \left(1 + \sup_{t\in[0,T]}\|\mathfrak{w}_t\|_{\mathbb{R}^d}\right)^p \sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\left[\frac{|tf(t, x, 0)|}{\max\{1, \|x\|^p_{\mathbb{R}^d}\}}\right]\\
&\leq T\left(1 + \sup_{t\in[0,T]}\|\mathfrak{w}_t\|_{\mathbb{R}^d}\right)^p \sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\left[\frac{|f(t, x, 0)|}{\max\{1, \|x\|^p_{\mathbb{R}^d}\}}\right] \leq LT\left(1 + \sup_{t\in[0,T]}\|\mathfrak{w}_t\|_{\mathbb{R}^d}\right)^p.
\end{aligned} \tag{176}$$

Combining (142), (175), and (176) implies for all $w \in \mathcal{Y}$, $\mathfrak{z} = (\mathfrak{u}, \mathfrak{w}) \in \mathcal{Z}$ that

$$\begin{aligned}
\|\Phi_0(0, w, \mathfrak{z})\|_{\mathcal{Y}} &= \sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\left[\frac{|[\Phi_0(0, w, \mathfrak{z})](T - t, x)|}{\max\{1, \|x\|^q_{\mathbb{R}^d}\}}\right]\\
&\leq \left[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\frac{|g(x + \mathfrak{w}_t)|}{\max\{1, \|x\|^q_{\mathbb{R}^d}\}}\right] + \left[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\frac{|tf(T - t + \mathfrak{u}t, x + \mathfrak{w}_{\mathfrak{u}t}, 0)|}{\max\{1, \|x\|^q_{\mathbb{R}^d}\}}\right]\\
&\leq L(T + 1)\left(1 + \sup_{t\in[0,T]}\|\mathfrak{w}_t\|_{\mathbb{R}^d}\right)^p.
\end{aligned} \tag{177}$$

In addition, (142), (174) and Lemma 3.8 (with $d = d$, $T = T$, $p = q$, $L = 1 + \sup_{t\in[0,T]}\|\mathfrak{w}_t\|_{\mathbb{R}^d}$, $\varrho = ([0, T] \times \mathbb{R}^d \ni (t, x) \mapsto (T - t + \mathfrak{u}t, x + \mathfrak{w}_{\mathfrak{u}t}) \in [0, T] \times \mathbb{R}^d)$, $v = v - w$

for $(\mathfrak{u},\mathfrak{w})\in\mathcal{Z}$, $v,w\in\mathcal{Y}$ in the notation of Lemma 3.8) prove for all $l\in\mathbb{N}$, $v,w\in\mathcal{Y}$, $\mathfrak{z}=(\mathfrak{u},\mathfrak{w})\in\mathcal{Z}$ that

$$\begin{aligned}
&\|\Phi_l(v,w,\mathfrak{z})\|_{\mathcal{Y}}=\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\left[\frac{|[\Phi_l(v,w,\mathfrak{z})](T-t,x)|}{\max\{1,\|x\|_{\mathbb{R}^d}^q\}}\right]\\
&=\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\left[\frac{t}{\max\{1,\|x\|_{\mathbb{R}^d}^q\}}\big|f\big(T-t+\mathfrak{u}t,x+\mathfrak{w}_{\mathfrak{u}t},v(T-t+\mathfrak{u}t,x+\mathfrak{w}_{\mathfrak{u}t})\big)\right.\\
&\qquad\qquad\left.-f\big(T-t+\mathfrak{u}t,x+\mathfrak{w}_{\mathfrak{u}t},w(T-t+\mathfrak{u}t,x+\mathfrak{w}_{\mathfrak{u}t})\big)\big|\right]\\
&\leq LT\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\left[\frac{|v(T-t+\mathfrak{u}t,x+\mathfrak{w}_{\mathfrak{u}t})-w(T-t+\mathfrak{u}t,x+\mathfrak{w}_{\mathfrak{u}t})|}{\max\{1,\|x\|_{\mathbb{R}^d}^q\}}\right]\\
&\leq LT\left(1+\sup_{t\in[0,T]}\|\mathfrak{w}_t\|_{\mathbb{R}^d}\right)^q\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\left[\frac{|v(t,x)-w(t,x)|}{\max\{1,\|x\|_{\mathbb{R}^d}^q\}}\right]\\
&=LT\left(1+\sup_{t\in[0,T]}\|\mathfrak{w}_t\|_{\mathbb{R}^d}\right)^q\|v-w\|_{\mathcal{Y}}.
\end{aligned}\tag{178}$$

Next we claim that it holds for all $l\in\mathbb{N}_0$, $j\in\mathbb{N}$, $r\in[0,\infty)$ that

$$\mathbb{E}\Big[\big\|\Phi_l\big(Y^0_{l,j},Y^1_{l-1,j},Z^0\big)\big\|_{\mathcal{Y}}^r+\big\|Y^0_{l,j}\big\|_{\mathcal{Y}}^r+\big\|Y^0_{l-1,j}\big\|_{\mathcal{Y}}^r\Big]<\infty.\tag{179}$$

We establish (179) by induction on $l\in\mathbb{N}_0$. For the base case $l=0$ observe that (177) and the fact that $\forall\,a,b,r\in[0,\infty)\colon(a+b)^r\leq 2^{\max\{r-1,0\}}(a^r+b^r)$ show for all $j\in\mathbb{N}$, $r\in[0,\infty)$ that

$$\begin{aligned}
&\mathbb{E}\Big[\big\|\Phi_0\big(Y^0_{0,j},Y^1_{-1,j},Z^0\big)\big\|_{\mathcal{Y}}^r\Big]=\mathbb{E}\Big[\big\|\Phi_0\big(0,0,U^0,W^0\big)\big\|_{\mathcal{Y}}^r\Big]\\
&\leq L^r(T+1)^r\,\mathbb{E}\left[\left(1+\sup_{t\in[0,T]}\|W^0_t\|_{\mathbb{R}^d}\right)^{pr}\right]\\
&\leq 2^{\max\{pr-1,0\}}L^r(T+1)^r\left(1+\mathbb{E}\left[\sup_{t\in[0,T]}\|W^0_t\|_{\mathbb{R}^d}^{pr}\right]\right)<\infty.
\end{aligned}\tag{180}$$

This and the fact that $\forall\,j\in\mathbb{N}$, $r\in[0,\infty)\colon\mathbb{E}\big[\|Y^0_{0,j}\|_{\mathcal{Y}}^r+\|Y^0_{-1,j}\|_{\mathcal{Y}}^r\big]=0<\infty$ prove (179) in the base case $l=0$. For the induction step $\mathbb{N}_0\ni l-1\to l\in\mathbb{N}$ let $l\in\mathbb{N}$ and assume that it holds for all $k\in\{0,1,\dots,l-1\}$, $j\in\mathbb{N}$, $r\in[0,\infty)$ that

$$\mathbb{E}\Big[\big\|\Phi_k\big(Y^0_{k,j},Y^1_{k-1,j},Z^0\big)\big\|_{\mathcal{Y}}^r+\big\|Y^0_{k,j}\big\|_{\mathcal{Y}}^r+\big\|Y^0_{k-1,j}\big\|_{\mathcal{Y}}^r\Big]<\infty.\tag{181}$$

Note that this, (ii) in Lemma 3.11, and (vi) in Proposition 2.8 ensure for all $j\in\mathbb{N}$, $r\in[1,\infty)$ that

$$\begin{aligned}
&\Big(\mathbb{E}\Big[\big\|Y^0_{l,j}\big\|_{\mathcal{Y}}^r\Big]\Big)^{1/r}=\left(\mathbb{E}\left[\left\|\sum_{k=0}^{l-1}\tfrac{1}{(M_j)^{l-k}}\left[\sum_{i=1}^{(M_j)^{l-k}}\Phi_k\big(Y^{(0,k,i)}_{k,j},Y^{(0,-k,i)}_{k-1,j},Z^{(0,k,i)}\big)\right]\right\|_{\mathcal{Y}}^r\right]\right)^{1/r}\\
&\leq\sum_{k=0}^{l-1}\tfrac{1}{(M_j)^{l-k}}\left[\sum_{i=1}^{(M_j)^{l-k}}\Big(\mathbb{E}\Big[\big\|\Phi_k\big(Y^{(0,k,i)}_{k,j},Y^{(0,-k,i)}_{k-1,j},Z^{(0,k,i)}\big)\big\|_{\mathcal{Y}}^r\Big]\Big)^{1/r}\right]\\
&=\sum_{k=0}^{l-1}\Big(\mathbb{E}\Big[\big\|\Phi_k\big(Y^0_{k,j},Y^1_{k-1,j},Z^0\big)\big\|_{\mathcal{Y}}^r\Big]\Big)^{1/r}<\infty.
\end{aligned}\tag{182}$$

Hölder's inequality, (178), the fact that $\forall\, a,b,r\in[0,\infty)\colon (a+b)^r\leq 2^{\max\{r-1,0\}}(a^r+b^r)$, (v) in Proposition 2.8, and (181) hence demonstrate for all $j\in\mathbb{N}$, $r\in[1,\infty)$ that

$$
\begin{aligned}
&\Big(\mathbb{E}\Big[\big\|\Phi_l\big(Y^0_{l,j},Y^1_{l-1,j},Z^0\big)\big\|^r_{\mathcal{Y}}\Big]\Big)^{1/r}\\
&\leq LT\bigg(\mathbb{E}\bigg[\bigg(1+\sup_{t\in[0,T]}\|W^0_t\|_{\mathbb{R}^d}\bigg)^{qr}\big\|Y^0_{l,j}-Y^1_{l-1,j}\big\|^r_{\mathcal{Y}}\bigg]\bigg)^{1/r}\\
&\leq LT\bigg(\mathbb{E}\bigg[\bigg(1+\sup_{t\in[0,T]}\|W^0_t\|_{\mathbb{R}^d}\bigg)^{2qr}\bigg]\bigg)^{1/(2r)}\Big(\mathbb{E}\Big[\big\|Y^0_{l,j}-Y^1_{l-1,j}\big\|^{2r}_{\mathcal{Y}}\Big]\Big)^{1/(2r)}\\
&\leq 2^{\max\{q-1/(2r),0\}}LT\bigg(1+\mathbb{E}\bigg[\sup_{t\in[0,T]}\|W^0_t\|^{2qr}_{\mathbb{R}^d}\bigg]\bigg)^{1/(2r)}\\
&\quad\cdot\bigg[\Big(\mathbb{E}\Big[\big\|Y^0_{l,j}\big\|^{2r}_{\mathcal{Y}}\Big]\Big)^{1/(2r)}+\Big(\mathbb{E}\Big[\big\|Y^0_{l-1,j}\big\|^{2r}_{\mathcal{Y}}\Big]\Big)^{1/(2r)}\bigg]<\infty.
\end{aligned}
\tag{183}
$$

Combining this with (182) and (181) establishes for all $j\in\mathbb{N}$, $r\in[0,\infty)$ that

$$
\mathbb{E}\Big[\big\|\Phi_l\big(Y^0_{l,j},Y^1_{l-1,j},Z^0\big)\big\|^r_{\mathcal{Y}}+\big\|Y^0_{l,j}\big\|^r_{\mathcal{Y}}+\big\|Y^0_{l-1,j}\big\|^r_{\mathcal{Y}}\Big]<\infty. \tag{184}
$$

Induction hence proves (179). The proof of Lemma 3.12 is thus complete. □

3.2.5 Estimates

Lemma 3.13. *Assume Setting 3.1 and let $C\in[0,\infty)$ be given by*

$$
C=e^{LT}\Big[\big(\mathbb{E}\big[|g(\xi+W^0_T)|^2\big]\big)^{1/2}+\sqrt{T}\big(\textstyle\int_0^T\mathbb{E}\big[|f(t,\xi+W^0_t,0)|^2\big]\,\mathrm{d}t\big)^{1/2}\Big]. \tag{185}
$$

Then it holds for all $k\in\mathbb{N}_0$ that

$$
\max\big\{\mathbb{E}\big[|\psi_k(\Phi_0(0,0,Z^0))|^2\big],\mathbb{E}\big[|\psi_k(y)|^2\big]\big\}\leq\tfrac{C^2}{k!}. \tag{186}
$$

Proof of Lemma 3.13. Throughout this proof let $F\colon C([0,T]\times\mathbb{R}^d,\mathbb{R})\to C([0,T]\times\mathbb{R}^d,\mathbb{R})$ satisfy for all $v\in C([0,T]\times\mathbb{R}^d,\mathbb{R})$, $t\in[0,T]$, $x\in\mathbb{R}^d$ that

$$
[F(v)](t,x)=f(t,x,v(t,x)). \tag{187}
$$

Observe that (141), the fact that $\mathbf{U}$ and $\mathbf{W}$ are independent, and Hutzenthaler et al. [68, Lemma 2.3] (with $S=[0,1]$, $U=([0,1]\times\Omega\ni(s,\omega)\mapsto {}^{s^{k-1}}\!/_{(k-1)!}|y(sT,\xi+\mathbf{W}_{sT}(\omega))|^2\in[0,\infty))$, $Y=\mathbf{U}$ for $k\in\mathbb{N}$ in the notation of [68, Lemma 2.3]) imply for all $k\in\mathbb{N}$ that

$$
\begin{aligned}
&\mathbb{E}\big[|\psi_k(y)|^2\big]=\mathbb{E}\Big[\tfrac{\mathbf{U}^{k-1}}{(k-1)!}|y(\mathbf{U}T,\xi+\mathbf{W}_{\mathbf{U}T})|^2\Big]\\
&=\tfrac{1}{(k-1)!}\int_0^1 s^{k-1}\,\mathbb{E}\big[|y(sT,\xi+\mathbf{W}_{sT})|^2\big]\,\mathrm{d}s\leq\tfrac{1}{k!}\bigg[\sup_{t\in[0,T]}\mathbb{E}\big[|y(t,\xi+\mathbf{W}_t)|^2\big]\bigg].
\end{aligned}
\tag{188}
$$

This, the fact that $\mathbb{E}[|\psi_0(y)|^2]=|y(0,\xi)|^2=\mathbb{E}[|y(0,\xi+\mathbf{W}_0)|^2]$, and [68, Lemma 3.4] establish for all $k\in\mathbb{N}_0$ that

$$
\begin{aligned}
&\mathbb{E}\big[|\psi_k(y)|^2\big]\leq\tfrac{1}{k!}\bigg[\sup_{t\in[0,T]}\mathbb{E}\big[|y(t,\xi+\mathbf{W}_t)|^2\big]\bigg]\\
&\leq\tfrac{e^{2LT}}{k!}\Big[\big(\mathbb{E}\big[|g(\xi+\mathbf{W}_T)|^2\big]\big)^{1/2}+\sqrt{T}\big(\textstyle\int_0^T\mathbb{E}\big[|[F(0)](t,\xi+\mathbf{W}_t)|^2\big]\,\mathrm{d}t\big)^{1/2}\Big]^2\\
&=\tfrac{e^{2LT}}{k!}\Big[\big(\mathbb{E}\big[|g(\xi+W^0_T)|^2\big]\big)^{1/2}+\sqrt{T}\big(\textstyle\int_0^T\mathbb{E}\big[|f(t,\xi+W^0_t,0)|^2\big]\,\mathrm{d}t\big)^{1/2}\Big]^2=\tfrac{C^2}{k!}.
\end{aligned}
\tag{189}
$$

Next note that (142) shows for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$[\Phi_0(0,0,Z^0)](t,x) = g(x + W^0_{T-t}) + (T-t) f\big(t + (T-t)U^0, x + W^0_{(T-t)U^0}, 0\big). \tag{190}$$

This, (141), and Hölder's inequality demonstrate for all $k \in \mathbb{N}$ that

$$\begin{aligned}
&\big(\mathbb{E}\big[|\psi_k(\Phi_0(0,0,Z^0))|^2\big]\big)^{1/2} = \mathbb{E}\Big[\tfrac{\mathbf{U}^{k-1}}{(k-1)!}|[\Phi_0(0,0,Z^0)](\mathbf{U}T, \xi + \mathbf{W}_{\mathbf{U}T})|^2\Big] \\
&= \bigg(\mathbb{E}\Big[\tfrac{\mathbf{U}^{k-1}}{(k-1)!}|g(\xi + \mathbf{W}_{\mathbf{U}T} + W^0_{(1-\mathbf{U})T}) \\
&\qquad + (1-\mathbf{U})T f\big(\mathbf{U}T + (1-\mathbf{U})U^0 T, \xi + \mathbf{W}_{\mathbf{U}T} + W^0_{(1-\mathbf{U})U^0T}, 0\big)\big|^2\Big]\bigg)^{1/2} \\
&\leq \bigg(\mathbb{E}\Big[\tfrac{\mathbf{U}^{k-1}}{(k-1)!}|g(\xi + \mathbf{W}_{\mathbf{U}T} + W^0_{(1-\mathbf{U})T})|^2\Big]\bigg)^{1/2} \\
&\quad + \bigg(\mathbb{E}\Big[\tfrac{\mathbf{U}^{k-1}}{(k-1)!}\big|(1-\mathbf{U})T f\big(\mathbf{U}T + (1-\mathbf{U})U^0 T, \xi + \mathbf{W}_{\mathbf{U}T} + W^0_{(1-\mathbf{U})U^0T}, 0\big)\big|^2\Big]\bigg)^{1/2}.
\end{aligned} \tag{191}$$

The fact that $\mathbf{U}$, $\mathbf{W}$, and W^0 are independent and [68, Lemma 2.3] (with $S = [0,1]$, $U = ([0,1] \times \Omega \ni (s,\omega) \mapsto {}^{s^{k-1}}\!/\!{}_{(k-1)!}|g(\xi + \mathbf{W}_{sT}(\omega) + W^0_{(1-s)T}(\omega))|^2 \in [0,\infty))$, $Y = \mathbf{U}$ for $k \in \mathbb{N}$ in the notation of [68, Lemma 2.3]) ensure for all $k \in \mathbb{N}$ that

$$\begin{aligned}
&\mathbb{E}\Big[\tfrac{\mathbf{U}^{k-1}}{(k-1)!}|g(\xi + \mathbf{W}_{\mathbf{U}T} + W^0_{(1-\mathbf{U})T})|^2\Big] = \int_0^1 \tfrac{s^{k-1}}{(k-1)!}\,\mathbb{E}\big[|g(\xi + \mathbf{W}_{sT} + W^0_{(1-s)T})|^2\big]\,\mathrm{d}s \\
&= \tfrac{1}{(k-1)!}\bigg[\int_0^1 s^{k-1}\,\mathrm{d}s\bigg]\mathbb{E}\big[|g(\xi + W^0_T)|^2\big] = \tfrac{1}{k!}\,\mathbb{E}\big[|g(\xi + W^0_T)|^2\big].
\end{aligned} \tag{192}$$

In addition, the fact that $\mathbf{U}$, U^0, $\mathbf{W}$, and W^0 are independent, [68, Lemma 2.3] (with $S = [0,1]$, $U = ([0,1] \times \Omega \ni (s,\omega) \mapsto {}^{s^{k-1}}\!/\!{}_{(k-1)!}|(1-s)Tf(sT + (1-s)U^0(\omega)T, \xi + \mathbf{W}_{sT}(\omega) + W^0_{(1-s)U^0(\omega)T}(\omega), 0)|^2 \in [0,\infty))$, $Y = \mathbf{U}$ for $k \in \mathbb{N}$ in the notation of [68, Lemma 2.3]), and [68, Lemma 2.10] (with $k = k$, $U = ([0,T] \times \mathbb{R}^d \times \Omega \ni (t,x,\omega) \mapsto f(t,x,0) \in \mathbb{R})$, $\mathfrak{r} = U^0$, $\mathbb{W} = W^0$ for $k \in \mathbb{N}$ in the notation of [68, Lemma 2.10]) establish for all $k \in \mathbb{N}$ that

$$\begin{aligned}
&\mathbb{E}\Big[\tfrac{\mathbf{U}^{k-1}}{(k-1)!}\big|(1-\mathbf{U})T f\big(\mathbf{U}T + (1-\mathbf{U})U^0T, \xi + \mathbf{W}_{\mathbf{U}T} + W^0_{(1-\mathbf{U})U^0T}, 0\big)\big|^2\Big] \\
&= \int_0^1 \tfrac{s^{k-1}}{(k-1)!}\,\mathbb{E}\Big[\big|(1-s)Tf\big(sT + (1-s)U^0T, \xi + \mathbf{W}_{sT} + W^0_{(1-s)U^0T}, 0\big)\big|^2\Big]\,\mathrm{d}s \\
&= \tfrac{1}{T^k}\int_0^T \tfrac{t^{k-1}}{(k-1)!}\,\mathbb{E}\Big[\big|(T-t)f\big(t + (T-t)U^0, \xi + \mathbf{W}_t + W^0_{(T-t)U^0}, 0\big)\big|^2\Big]\,\mathrm{d}t \\
&= \tfrac{1}{T^k}\int_0^T \tfrac{t^{k-1}}{(k-1)!}\,\mathbb{E}\Big[\big|(T-t)f\big(t + (T-t)U^0, \xi + \mathbf{W}_t + W^0_{t+(T-t)U^0} - W^0_t, 0\big)\big|^2\Big]\,\mathrm{d}t \\
&\leq \tfrac{T^2}{T^{k+1}}\int_0^T \tfrac{t^k}{k!}\,\mathbb{E}\big[|f(t, \xi + \mathbf{W}_t, 0)|^2\big]\,\mathrm{d}t \leq \tfrac{T}{k!}\int_0^T \mathbb{E}\big[|f(t, \xi + \mathbf{W}_t, 0)|^2\big]\,\mathrm{d}t \\
&= \tfrac{T}{k!}\int_0^T \mathbb{E}\big[|f(t, \xi + W^0_t, 0)|^2\big]\,\mathrm{d}t.
\end{aligned} \tag{193}$$

Combining (191) with (192)–(193) yields for all $k \in \mathbb{N}$ that

$$\begin{aligned}
&\mathbb{E}\big[|\psi_k(\Phi_0(0,0,Z^0))|^2\big] \\
&\leq \tfrac{1}{k!}\Big[\big(\mathbb{E}\big[|g(\xi + W^0_T)|^2\big]\big)^{1/2} + \sqrt{T}\big(\textstyle\int_0^T \mathbb{E}\big[|f(t, \xi + W^0_t, 0)|^2\big]\,\mathrm{d}t\big)^{1/2}\Big]^2 \\
&\leq \tfrac{e^{2LT}}{k!}\Big[\big(\mathbb{E}\big[|g(\xi + W^0_T)|^2\big]\big)^{1/2} + \sqrt{T}\big(\textstyle\int_0^T \mathbb{E}\big[|f(t, \xi + W^0_t, 0)|^2\big]\,\mathrm{d}t\big)^{1/2}\Big]^2 = \tfrac{C^2}{k!}.
\end{aligned} \tag{194}$$

Moreover, (190), (141), Hölder's inequality, the fact that U^0 and W^0 are independent, and [68, Lemma 2.3] imply that

$$\begin{aligned}
&\mathbb{E}\big[|\psi_0(\Phi_0(0,0,Z^0))|^2\big] = \mathbb{E}\big[|g(\xi+W^0_T) + Tf(U^0T, \xi + W^0_{U^0T}, 0)|^2\big] \\
&\leq \Big[\big(\mathbb{E}\big[|g(\xi+W^0_T)|^2\big]\big)^{1/2} + \big(T^2\,\mathbb{E}\big[|f(U^0T,\xi+W^0_{U^0T},0)|^2\big]\big)^{1/2}\Big]^2 \\
&= \Big[\big(\mathbb{E}\big[|g(\xi+W^0_T)|^2\big]\big)^{1/2} + \big(T^2 \textstyle\int_0^1 \mathbb{E}\big[|f(sT,\xi+W^0_{sT},0)|^2\big]\,\mathrm{d}s\big)^{1/2}\Big]^2 \\
&= \Big[\big(\mathbb{E}\big[|g(\xi+W^0_T)|^2\big]\big)^{1/2} + \big(T \textstyle\int_0^T \mathbb{E}\big[|f(t,\xi+W^0_t,0)|^2\big]\,\mathrm{d}t\big)^{1/2}\Big]^2 \\
&\leq e^{2LT}\Big[\big(\mathbb{E}\big[|g(\xi+W^0_T)|^2\big]\big)^{1/2} + \sqrt{T}\big(\textstyle\int_0^T \mathbb{E}\big[|f(t,\xi+W^0_t,0)|^2\big]\,\mathrm{d}t\big)^{1/2}\Big]^2 = \tfrac{C^2}{0!}.
\end{aligned} \tag{195}$$

The proof of Lemma 3.13 is thus complete. □

Lemma 3.14. *Assume Setting 3.1. Then it holds for all $k \in \mathbb{N}_0$, $n \in \mathbb{N}$, $u, v \in \mathcal{Y}$ that*

$$\mathbb{E}\big[|\psi_k(\Phi_n(u,v,Z^0))|^2\big] \leq (LT)^2\,\mathbb{E}\big[|\psi_{k+1}(u-v)|^2\big]. \tag{196}$$

Proof of Lemma 3.14. Throughout this proof let $u, v \in \mathcal{Y}$. Observe that (142) shows for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$\begin{aligned}
&|[\Phi_1(u,v,Z^0)](t,x)| \\
&= (T-t)\big|f\big(t+(T-t)U^0, x+W^0_{(T-t)U^0}, u\big(t+(T-t)U^0, x+W^0_{(T-t)U^0}\big)\big) \\
&\qquad - f\big(t+(T-t)U^0, x+W^0_{(T-t)U^0}, v\big(t+(T-t)U^0, x+W^0_{(T-t)U^0}\big)\big)\big| \\
&\leq L(T-t)\big|u\big(t+(T-t)U^0, x+W^0_{(T-t)U^0}\big) - v\big(t+(T-t)U^0, x+W^0_{(T-t)U^0}\big)\big| \\
&= L\big|(T-t)\cdot[u-v]\big(t+(T-t)U^0, x+W^0_{(T-t)U^0}\big)\big|.
\end{aligned} \tag{197}$$

Equation (141), the fact that $\mathbf{U}$, U^0, $\mathbf{W}$, and W^0 are independent, [68, Lemma 2.3], and [68, Lemma 2.10] (with $k = k$, $U = ([0,T]\times\mathbb{R}^d\times\Omega \ni (t,x,\omega) \mapsto u(t,x)-v(t,x) \in \mathbb{R})$, $\mathfrak{r} = U^0$, $\mathbb{W} = W^0$ for $k \in \mathbb{N}$ in the notation of [68, Lemma 2.10]) hence prove for all $k \in \mathbb{N}$ that

$$\begin{aligned}
&\mathbb{E}\big[|\psi_k(\Phi_1(u,v,Z^0))|^2\big] = \mathbb{E}\Big[\tfrac{\mathbf{U}^{k-1}}{(k-1)!}|[\Phi_1(u,v,Z^0)](\mathbf{U}T, \xi+\mathbf{W}_{\mathbf{U}T})|^2\Big] \\
&\leq L^2\,\mathbb{E}\Big[\tfrac{\mathbf{U}^{k-1}}{(k-1)!}\big|(1-\mathbf{U})T\cdot[u-v]\big(\mathbf{U}T+(1-\mathbf{U})U^0T, \xi+\mathbf{W}_{\mathbf{U}T}+W^0_{(1-\mathbf{U})U^0T}\big)\big|^2\Big] \\
&= L^2\int_0^1 \tfrac{s^{k-1}}{(k-1)!}\,\mathbb{E}\Big[\big|(1-s)T\cdot[u-v]\big(sT+(1-s)U^0T, \xi+\mathbf{W}_{sT}+W^0_{(1-s)U^0T}\big)\big|^2\Big]\,\mathrm{d}s \\
&= \tfrac{L^2}{T^k}\int_0^T \tfrac{t^{k-1}}{(k-1)!}\,\mathbb{E}\Big[\big|(T-t)\cdot[u-v]\big(t+(T-t)U^0, \xi+\mathbf{W}_t+W^0_{(T-t)U^0}\big)\big|^2\Big]\,\mathrm{d}t \\
&= \tfrac{L^2}{T^k}\int_0^T \tfrac{t^{k-1}}{(k-1)!}\,\mathbb{E}\Big[\big|(T-t)\cdot[u-v]\big(t+(T-t)U^0, \xi+\mathbf{W}_t+W^0_{t+(T-t)U^0}-W^0_t\big)\big|^2\Big]\,\mathrm{d}t \\
&\leq \tfrac{(LT)^2}{T^{k+1}}\int_0^T \tfrac{t^k}{k!}\,\mathbb{E}\big[|[u-v](t,\xi+\mathbf{W}_t)|^2\big]\,\mathrm{d}t \\
&= (LT)^2\int_0^1 \tfrac{s^k}{k!}\,\mathbb{E}\big[|[u-v](sT,\xi+\mathbf{W}_{sT})|^2\big]\,\mathrm{d}s \\
&= (LT)^2\,\mathbb{E}\Big[\tfrac{\mathbf{U}^k}{k!}|[u-v](\mathbf{U}T,\xi+\mathbf{W}_{\mathbf{U}T})|^2\Big] = (LT)^2\,\mathbb{E}\big[|\psi_{k+1}(u-v)|^2\big].
\end{aligned} \tag{198}$$

In addition, (141), (197) and the fact that (U^0, W^0) and $(\mathbf{U}, \mathbf{W})$ are identically distributed ensure that

$$\begin{aligned}&\mathbb{E}\big[|\psi_0(\Phi_1(u,v,Z^0))|^2\big] = \mathbb{E}\big[|[\Phi_1(u,v,Z^0)](0,\xi)|^2\big]\\&\le (LT)^2\,\mathbb{E}\big[|[u-v](U^0T, \xi + W^0_{U^0T})|^2\big]\\&= (LT)^2\,\mathbb{E}\big[|[u-v](\mathbf{U}T, \xi + \mathbf{W}_{\mathbf{U}T})|^2\big] = (LT)^2\,\mathbb{E}\big[|\psi_1(u-v)|^2\big].\end{aligned}\tag{199}$$

This, (198), and the fact that $\forall\, n \in \mathbb{N}\colon \Phi_n = \Phi_1$ complete the proof of Lemma 3.14. □

Lemma 3.15. *Assume Setting 3.1. Then it holds for all $k \in \mathbb{N}_0$, $n, j \in \mathbb{N}$ that*

$$\mathbb{E}\left[\left|\psi_k\left(y - \sum_{l=0}^{n-1} \mathbb{E}\big[\Phi_l\big(Y^0_{l,j}, Y^1_{l-1,j}, Z^0\big)\big]\right)\right|^2\right] \le (LT)^2\,\mathbb{E}\Big[\big|\psi_{k+1}\big(Y^0_{n-1,j} - y\big)\big|^2\Big].\tag{200}$$

Proof of Lemma 3.15. Throughout this proof let $\Psi_{n,j}\colon [0,T] \times \mathbb{R}^d \to [0,\infty)$, $j \in \mathbb{N}$, $n \in \mathbb{N}_0$, satisfy for all $n, j \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$\Psi_{n-1,j}(t,x) = \mathbb{E}\Big[\big|(T-t)\cdot\big[Y^0_{n-1,j} - y\big]\big(t + (T-t)U^0, x + W^0_{(T-t)U^0}\big)\big|^2\Big]\tag{201}$$

(cf. Lemma 3.12). To start with, observe that (142), (i)–(ii) in Lemma 3.11, (ii) in Lemma 3.10, and (iii) and (v) in Proposition 2.8 show for all $l, j \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$\begin{aligned}&\mathbb{E}\big[\big[\Phi_l\big(Y^0_{l,j}, Y^1_{l-1,j}, Z^0\big)\big](T-t,x)\big]\\&= t\,\mathbb{E}\big[f\big(T-t+U^0t, x+W^0_{U^0t}, Y^0_{l,j}\big(T-t+U^0t, x+W^0_{U^0t}\big)\big)\\&\qquad - f\big(T-t+U^0t, x+W^0_{U^0t}, Y^1_{l-1,j}\big(T-t+U^0t, x+W^0_{U^0t}\big)\big)\big]\\&= t\,\mathbb{E}\big[f\big(T-t+U^0t, x+W^0_{U^0t}, Y^0_{l,j}\big(T-t+U^0t, x+W^0_{U^0t}\big)\big)\\&\qquad - f\big(T-t+U^0t, x+W^0_{U^0t}, Y^0_{l-1,j}\big(T-t+U^0t, x+W^0_{U^0t}\big)\big)\big].\end{aligned}\tag{202}$$

Again (142) hence ensures for all $n, j \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$\begin{aligned}&\left[\sum_{l=0}^{n-1} \mathbb{E}\big[\Phi_l\big(Y^0_{l,j}, Y^1_{l-1,j}, Z^0\big)\big]\right](T-t,x)\\&= \sum_{l=0}^{n-1} \mathbb{E}\big[\big[\Phi_l\big(Y^0_{l,j}, Y^1_{l-1,j}, Z^0\big)\big](T-t,x)\big]\\&= \mathbb{E}[g(x+W^0_t)] + t\,\mathbb{E}\big[f\big(T-t+U^0t, x+W^0_{U^0t}, Y^0_{0,j}\big(T-t+U^0t, x+W^0_{U^0t}\big)\big)\big]\\&\quad + t\sum_{l=1}^{n-1} \mathbb{E}\big[f\big(T-t+U^0t, x+W^0_{U^0t}, Y^0_{l,j}\big(T-t+U^0t, x+W^0_{U^0t}\big)\big)\\&\qquad\qquad - f\big(T-t+U^0t, x+W^0_{U^0t}, Y^0_{l-1,j}\big(T-t+U^0t, x+W^0_{U^0t}\big)\big)\big]\\&= \mathbb{E}[g(x+W^0_t)] + t\,\mathbb{E}\big[f\big(T-t+U^0t, x+W^0_{U^0t}, Y^0_{n-1,j}\big(T-t+U^0t, x+W^0_{U^0t}\big)\big)\big].\end{aligned}\tag{203}$$

In addition, (136), the fact that $\mathbf{W}$ and W^0 are identically distributed, the fact that W^0 and U^0 are independent, and [68, Lemma 2.4] (with $S = [0,1]$, $U = ([0,1] \times \Omega \ni (u,\omega) \mapsto f(t + (T-t)u, x + W^0_{(T-t)u}(\omega), y(t + (T-t)u, x + W^0_{(T-t)u}(\omega))) \in \mathbb{R})$, $Y = U^0$ for $x \in \mathbb{R}^d$,

$t \in [0,T]$ in the notation of [68, Lemma 2.4]) imply for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$\begin{aligned} y(t,x) &= \mathbb{E}[g(x+\mathbf{W}_{T-t})] + \int_t^T \mathbb{E}\big[f\big(s,x+\mathbf{W}_{s-t},y(s,x+\mathbf{W}_{s-t})\big)\big]\,\mathrm{d}s \\ &= \mathbb{E}[g(x+W^0_{T-t})] + \int_t^T \mathbb{E}\big[f\big(s,x+W^0_{s-t},y(s,x+W^0_{s-t})\big)\big]\,\mathrm{d}s \\ &= \mathbb{E}[g(x+W^0_{T-t})] \\ &\quad + (T-t)\int_0^1 \mathbb{E}\big[f\big(t+(T-t)u,x+W^0_{(T-t)u},y\big(t+(T-t)u,x+W^0_{(T-t)u}\big)\big)\big]\,\mathrm{d}u \\ &= \mathbb{E}[g(x+W^0_{T-t})] \\ &\quad + (T-t)\,\mathbb{E}\big[f\big(t+(T-t)U^0,x+W^0_{(T-t)U^0},y\big(t+(T-t)U^0,x+W^0_{(T-t)U^0}\big)\big)\big]. \end{aligned} \tag{204}$$

This and (203) demonstrate for all $n,j \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$\begin{aligned} &\left|\left[y - \sum_{l=0}^{n-1} \mathbb{E}\big[\Phi_l\big(Y^0_{l,j},Y^1_{l-1,j},Z^0\big)\big]\right](t,x)\right| \\ &\le (T-t)\,\mathbb{E}\big[\big|f\big(t+(T-t)U^0,x+W^0_{(T-t)U^0},y\big(t+(T-t)U^0,x+W^0_{(T-t)U^0}\big)\big) \\ &\qquad\qquad - f\big(t+(T-t)U^0,x+W^0_{(T-t)U^0},Y^0_{n-1,j}\big(t+(T-t)U^0,x+W^0_{(T-t)U^0}\big)\big)\big|\big] \\ &\le L(T-t)\,\mathbb{E}\big[\big|y\big(t+(T-t)U^0,x+W^0_{(T-t)U^0}\big) - Y^0_{n-1,j}\big(t+(T-t)U^0,x+W^0_{(T-t)U^0}\big)\big|\big] \\ &= L(T-t)\,\mathbb{E}\big[\big|\big[Y^0_{n-1,j}-y\big]\big(t+(T-t)U^0,x+W^0_{(T-t)U^0}\big)\big|\big]. \end{aligned} \tag{205}$$

Jensen's inequality and (201) hence ensure for all $n,j \in \mathbb{N}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$\begin{aligned} &\left|\left[y - \sum_{l=0}^{n-1} \mathbb{E}\big[\Phi_l\big(Y^0_{l,j},Y^1_{l-1,j},Z^0\big)\big]\right](t,x)\right|^2 \\ &\le L^2(T-t)^2\big(\mathbb{E}\big[\big|\big[Y^0_{n-1,j}-y\big]\big(t+(T-t)U^0,x+W^0_{(T-t)U^0}\big)\big|\big]\big)^2 \\ &\le L^2(T-t)^2\,\mathbb{E}\Big[\big|\big[Y^0_{n-1,j}-y\big]\big(t+(T-t)U^0,x+W^0_{(T-t)U^0}\big)\big|^2\Big] = L^2\Psi_{n-1,j}(t,x). \end{aligned} \tag{206}$$

Furthermore, (201), the fact that it holds for every $n,j \in \mathbb{N}$ that $\mathbf{W}$, $Y^0_{n-1,j}$, U^0, and W^0 are independent (cf. Lemma 3.11 and (ii)–(iii) in Proposition 2.8), and [68, Lemma 2.3] (with $S = \mathbb{R}^d$, $U = (\mathbb{R}^d \times \Omega \ni (w,\omega) \mapsto |(T-t)\cdot[Y^0_{n-1,j}(\omega)-y](t+(T-t)U^0(\omega),\xi + w + W^0_{(T-t)U^0(\omega)}(\omega))|^2 \in [0,\infty))$, $Y = \mathbf{W}_t$ for $t \in [0,T]$, $j,n \in \mathbb{N}$ in the notation of [68, Lemma 2.3]) prove for all $n,j \in \mathbb{N}$, $t \in [0,T]$ that

$$\begin{aligned} &\mathbb{E}[\Psi_{n-1,j}(t,\xi+\mathbf{W}_t)] = \int_{\mathbb{R}^d} \Psi_{n-1,j}(t,\xi+w)\,\big(\mathbf{W}_t(\mathbb{P})_{\mathscr{B}(\mathbb{R}^d)}\big)(\mathrm{d}w) \\ &= \int_{\mathbb{R}^d} \mathbb{E}\Big[\big|(T-t)\cdot\big[Y^0_{n-1,j}-y\big]\big(t+(T-t)U^0,\xi+w+W^0_{(T-t)U^0}\big)\big|^2\Big]\,\big(\mathbf{W}_t(\mathbb{P})_{\mathscr{B}(\mathbb{R}^d)}\big)(\mathrm{d}w) \\ &= \mathbb{E}\Big[\big|(T-t)\cdot\big[Y^0_{n-1,j}-y\big]\big(t+(T-t)U^0,\xi+\mathbf{W}_t+W^0_{(T-t)U^0}\big)\big|^2\Big] \\ &= \mathbb{E}\Big[\big|(T-t)\cdot\big[Y^0_{n-1,j}-y\big]\big(t+(T-t)U^0,\xi+\mathbf{W}_t+W^0_{t+(T-t)U^0}-W^0_t\big)\big|^2\Big]. \end{aligned} \tag{207}$$

Combining (141) with (206), the fact that $\mathbf{U}$ and $\mathbf{W}$ are independent, [68, Lemma 2.3] (with $S = [0,1]$, $U = ([0,1]\times\Omega \ni (s,\omega) \mapsto {}^{s^{k-1}}\!/_{(k-1)!}\,\Psi_{n-1,j}(sT,\xi+\mathbf{W}_{sT}(\omega)) \in [0,\infty))$, $Y = \mathbf{U}$ for $j,n,k \in \mathbb{N}$ in the notation of [68, Lemma 2.3]), (207), again the fact that it holds for every $n,j \in \mathbb{N}$ that $\mathbf{W}$, $Y^0_{n-1,j}$, U^0, and W^0 are independent, and [68, Lemma 2.10] (with $k = k$, $U = ([0,T] \times \mathbb{R}^d \times \Omega \ni (t,x,\omega) \mapsto [Y^0_{n-1,j}(\omega)](t,x) - y(t,x) \in \mathbb{R})$, $\mathfrak{r} = U^0$, $\mathbb{W} = W^0$ for $j,n,k \in \mathbb{N}$ in the notation of [68, Lemma 2.10]) establishes for all $k,n,j \in \mathbb{N}$ that

$$
\begin{aligned}
&\mathbb{E}\left[\left|\psi_k\left(y - \sum_{l=0}^{n-1} \mathbb{E}\big[\Phi_l\big(Y^0_{l,j}, Y^1_{l-1,j}, Z^0\big)\big]\right)\right|^2\right] \\
&= \mathbb{E}\left[\tfrac{\mathbf{U}^{k-1}}{(k-1)!}\left|\left[y - \sum_{l=0}^{n-1} \mathbb{E}\big[\Phi_l\big(Y^0_{l,j}, Y^1_{l-1,j}, Z^0\big)\big]\right](\mathbf{U}T, \xi + \mathbf{W}_{\mathbf{U}T})\right|^2\right] \\
&\leq L^2\, \mathbb{E}\Big[\tfrac{\mathbf{U}^{k-1}}{(k-1)!}\Psi_{n-1,j}(\mathbf{U}T, \xi + \mathbf{W}_{\mathbf{U}T})\Big] = L^2 \int_0^1 \tfrac{s^{k-1}}{(k-1)!}\, \mathbb{E}[\Psi_{n-1,j}(sT, \xi + \mathbf{W}_{sT})]\, \mathrm{d}s \\
&= \tfrac{L^2}{T^k} \int_0^T \tfrac{t^{k-1}}{(k-1)!}\, \mathbb{E}[\Psi_{n-1,j}(t, \xi + \mathbf{W}_t)]\, \mathrm{d}t \\
&= \tfrac{L^2}{T^k} \int_0^T \tfrac{t^{k-1}}{(k-1)!}\, \mathbb{E}\Big[\big|(T-t)\cdot \big[Y^0_{n-1,j} - y\big]\big(t + (T-t)U^0, \xi + \mathbf{W}_t + W^0_{t+(T-t)U^0} - W^0_t\big)\big|^2\Big]\, \mathrm{d}t \\
&\leq \tfrac{(LT)^2}{T^{k+1}} \int_0^T \tfrac{t^k}{k!}\, \mathbb{E}\Big[\big|\big[Y^0_{n-1,j} - y\big](t, \xi + \mathbf{W}_t)\big|^2\Big]\, \mathrm{d}t.
\end{aligned}
\tag{208}
$$

This, the fact that it holds for every $n,j \in \mathbb{N}$ that $Y^0_{n-1,j}$, $\mathbf{W}$, and $\mathbf{U}$ are independent, [68, Lemma 2.3] (with $S = [0,1]$, $U = ([0,1] \times \Omega \ni (s,\omega) \mapsto {}^{s^k}\!/_{k!}|[Y^0_{n-1,j}(\omega) - y](sT, \xi + \mathbf{W}_{sT}(\omega))|^2 \in [0,\infty))$, $Y = \mathbf{U}$ for $j,n,k \in \mathbb{N}$ in the notation of [68, Lemma 2.3]), and (141) show for all $k,n,j \in \mathbb{N}$ that

$$
\begin{aligned}
&\mathbb{E}\left[\left|\psi_k\left(y - \sum_{l=0}^{n-1} \mathbb{E}\big[\Phi_l\big(Y^0_{l,j}, Y^1_{l-1,j}, Z^0\big)\big]\right)\right|^2\right] \\
&\leq (LT)^2 \int_0^1 \tfrac{s^k}{k!}\, \mathbb{E}\Big[\big|\big[Y^0_{n-1,j} - y\big](sT, \xi + \mathbf{W}_{sT})\big|^2\Big]\, \mathrm{d}s \\
&= (LT)^2\, \mathbb{E}\Big[\tfrac{\mathbf{U}^k}{k!}\big|\big[Y^0_{n-1,j} - y\big](\mathbf{U}T, \xi + \mathbf{W}_{\mathbf{U}T})\big|^2\Big] = (LT)^2\, \mathbb{E}\Big[\big|\psi_{k+1}\big(Y^0_{n-1,j} - y\big)\big|^2\Big].
\end{aligned}
\tag{209}
$$

Moreover, (141), (206), and the fact that it holds for all $n,j \in \mathbb{N}$ that $(Y^0_{n-1,j}, U^0, W^0)$ and $(Y^0_{n-1,j}, \mathbf{U}, \mathbf{W})$ are identically distributed demonstrate for all $n,j \in \mathbb{N}$ that

$$
\begin{aligned}
&\mathbb{E}\left[\left|\psi_0\left(y - \sum_{l=0}^{n-1} \mathbb{E}\big[\Phi_l\big(Y^0_{l,j}, Y^1_{l-1,j}, Z^0\big)\big]\right)\right|^2\right] = \left|\left[y - \sum_{l=0}^{n-1} \mathbb{E}\big[\Phi_l\big(Y^0_{l,j}, Y^1_{l-1,j}, Z^0\big)\big]\right](0,\xi)\right|^2 \\
&\leq (LT)^2\, \mathbb{E}\Big[\big|\big[Y^0_{n-1,j} - y\big]\big(U^0T, \xi + W^0_{U^0T}\big)\big|^2\Big] \\
&= (LT)^2\, \mathbb{E}\Big[\big|\big[Y^0_{n-1,j} - y\big]\big(\mathbf{U}T, \xi + \mathbf{W}_{\mathbf{U}T}\big)\big|^2\Big] = (LT)^2\, \mathbb{E}\Big[\big|\psi_1\big(Y^0_{n-1,j} - y\big)\big|^2\Big].
\end{aligned}
\tag{210}
$$

The proof of Lemma 3.15 is thus complete. □

3.3 Complexity analysis

3.3.1 MLP approximations in fixed space dimensions

Proposition 3.16. *Let* $d \in \mathbb{N}$, $\xi \in \mathbb{R}^d$, $T \in (0,\infty)$, $L, p, \mathfrak{B}, \kappa, C \in [0,\infty)$, $\Theta = \cup_{n=1}^\infty \mathbb{Z}^n$, $(M_j)_{j\in\mathbb{N}} \subseteq \mathbb{N}$ *satisfy* $\liminf_{j\to\infty} M_j = \infty$, $\sup_{j\in\mathbb{N}} {}^{M_{j+1}}\!/_{M_j} \leq \mathfrak{B}$, *and* $\sup_{j\in\mathbb{N}} {}^{M_j}\!/_j \leq \kappa$, *let* $f \in C([0,T] \times \mathbb{R}^d \times \mathbb{R}, \mathbb{R})$, $g \in C(\mathbb{R}^d, \mathbb{R})$ *satisfy for all* $t \in [0,T]$, $x \in \mathbb{R}^d$, $v,w \in \mathbb{R}$ *that* $\max\{|f(t,x,0)|, |g(x)|\} \leq L\max\{1, \|x\|^p_{\mathbb{R}^d}\}$ *and* $|f(t,x,v) - f(t,x,w)| \leq L|v-w|$, *let* $(\Omega, \mathscr{F}, \mathbb{P})$ *be a probability space, let* $U^\theta\colon \Omega \to [0,1]$, $\theta \in \Theta$, *be independent on* $[0,1]$ *uniformly distributed random variables, let* $W^\theta\colon [0,T]\times\Omega \to \mathbb{R}^d$, $\theta\in\Theta$, *be independent standard Brownian motions with continuous sample paths, assume that* $(U^\theta)_{\theta\in\Theta}$ *and* $(W^\theta)_{\theta\in\Theta}$ *are independent, assume that*

$$C = \max\Big\{1, e^{LT}\Big[\big(\mathbb{E}\big[|g(\xi+W^0_T)|^2\big]\big)^{1/2} + \sqrt{T}\big(\textstyle\int_0^T \mathbb{E}\big[|f(t,\xi+W^0_t,0)|^2\big]\,dt\big)^{1/2}\Big]\Big\}, \quad (211)$$

let $Y^\theta_{n,j}\colon [0,T]\times\mathbb{R}^d\times\Omega\to\mathbb{R}$, $\theta\in\Theta$, $j\in\mathbb{N}$, $n\in(\mathbb{N}_0\cup\{-1\})$, *satisfy for all* $n,j\in\mathbb{N}$, $\theta\in\Theta$, $t\in[0,T]$, $x\in\mathbb{R}^d$ *that* $Y^\theta_{-1,j}(t,x)=Y^\theta_{0,j}(t,x)=0$ *and*

$$\begin{aligned}
Y^\theta_{n,j}(T-t,x) &= \tfrac{1}{(M_j)^n}\Bigg[\sum_{i=1}^{(M_j)^n} g\big(x+W^{(\theta,0,i)}_t\big)\Bigg] + \sum_{l=0}^{n-1}\tfrac{t}{(M_j)^{n-l}}\Bigg[\sum_{i=1}^{(M_j)^{n-l}}\\
&\Big[f\Big(T-t+U^{(\theta,l,i)}t, x+W^{(\theta,l,i)}_{U^{(\theta,l,i)}t}, Y^{(\theta,l,i)}_{l,j}\big(T-t+U^{(\theta,l,i)}t, x+W^{(\theta,l,i)}_{U^{(\theta,l,i)}t}\big)\Big) \qquad (212)\\
&- \mathbb{1}_{\mathbb{N}}(l) f\Big(T-t+U^{(\theta,l,i)}t, x+W^{(\theta,l,i)}_{U^{(\theta,l,i)}t}, Y^{(\theta,-l,i)}_{l-1,j}\big(T-t+U^{(\theta,l,i)}t, x+W^{(\theta,l,i)}_{U^{(\theta,l,i)}t}\big)\Big)\Big]\Bigg],
\end{aligned}$$

and let $(\mathrm{Cost}_{n,j})_{(n,j)\in(\mathbb{N}_0\cup\{-1\})\times\mathbb{N}}\subseteq\mathbb{N}_0$ *satisfy for all* $n,j\in\mathbb{N}$ *that* $\mathrm{Cost}_{-1,j}=\mathrm{Cost}_{0,j}=0$ *and*

$$\mathrm{Cost}_{n,j} \le (M_j)^n d + \sum_{l=0}^{n-1}\big[(M_j)^{n-l}(\mathrm{Cost}_{l,j}+\mathrm{Cost}_{l-1,j}+d+1)\big]. \quad (213)$$

Then

(i) *there exists a unique at most polynomially growing viscosity solution* $y\in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ *of*

$$\big(\tfrac{\partial y}{\partial t}\big)(t,x) + \tfrac12(\Delta_x y)(t,x) + f(t,x,y(t,x)) = 0 \quad (214)$$

with $y(T,x)=g(x)$ *for* $(t,x)\in(0,T)\times\mathbb{R}^d$,

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

$$\big(\mathbb{E}\big[|Y^0_{n,n}(0,\xi)-y(0,\xi)|^2\big]\big)^{1/2} \le C\left[\frac{e^\kappa(1+(2LT)^2)}{M_n}\right]^{n/2} < \infty, \quad (215)$$

(iii) *it holds for all* $n\in\mathbb{N}$ *that* $\mathrm{Cost}_{n,n}\le(5M_n)^n d$, *and*

(iv) *there exists* $(N_\varepsilon)_{\varepsilon\in(0,1]}\subseteq\mathbb{N}$ *such that it holds for all* $\varepsilon\in(0,1]$, $\delta\in(0,\infty)$ *that* $\sup_{n\in\{N_\varepsilon,N_\varepsilon+1,\ldots\}}\big(\mathbb{E}\big[|Y^0_{n,n}(0,\xi)-y(0,\xi)|^2\big]\big)^{1/2}\le\varepsilon$ *and*

$$\mathrm{Cost}_{N_\varepsilon,N_\varepsilon} \le 5de^\kappa C^{2(1+\delta)}\Big(1+\sup_{n\in\mathbb{N}}\Big[\tfrac{[5\mathfrak{B}e^{2\kappa}(1+(2LT)^2)]^{(1+\delta)}}{(M_n)^\delta}\Big]^n\Big)\varepsilon^{-2(1+\delta)} < \infty. \quad (216)$$

Proof of Proposition 3.16. Throughout this proof assume w.l.o.g. that $L>0$, assume w.l.o.g. that there exist an on $[0,1]$ uniformly distributed random variable $\mathbf{U}\colon\Omega\to[0,1]$ and a standard Brownian motion $\mathbf{W}\colon[0,T]\times\Omega\to\mathbb{R}^d$ with continuous sample paths such that $\mathbf{U}$, $\mathbf{W}$, $(U^\theta)_{\theta\in\Theta}$, and $(W^\theta)_{\theta\in\Theta}$ are independent, let $\mathfrak{z},\gamma\in[0,\infty)$, $c\in(0,\infty)$, $\mathfrak{y}_{-1},\mathfrak{y}_0\in C([0,T]\times\mathbb{R}^d,\mathbb{R})$ be given by $\mathfrak{z}=d$, $\gamma=2$, $c=(LT)^2$, and $\mathfrak{y}_{-1}=\mathfrak{y}_0=0$, let $q\in(p,\infty)$, let $\mathcal{Y}\subseteq C([0,T]\times\mathbb{R}^d,\mathbb{R})$ be the set given by

$$\mathcal{Y} = \left\{v\in C([0,T]\times\mathbb{R}^d,\mathbb{R})\colon \limsup_{\mathbb{N}\ni n\to\infty}\ \sup_{(t,x)\in[0,T]\times\mathbb{R}^d,\,\|x\|_{\mathbb{R}^d}\ge n}\frac{|v(t,x)|}{\|x\|^q_{\mathbb{R}^d}}=0\right\}, \quad (217)$$

let $\|\cdot\|_{\mathcal{Y}}\colon \mathcal{Y} \to [0,\infty)$ satisfy for all $v \in \mathcal{Y}$ that

$$\|v\|_{\mathcal{Y}} = \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \left[\frac{|v(t,x)|}{\max\{1, \|x\|_{\mathbb{R}^d}^q\}}\right], \tag{218}$$

let $(\mathcal{Z}, \mathscr{Z}) = \big([0,1] \times C([0,T],\mathbb{R}^d), \mathscr{B}([0,1]) \otimes \mathscr{B}(C([0,T],\mathbb{R}^d))\big)$, let $Z^\theta\colon \Omega \to \mathcal{Z}$, $\theta \in \Theta$, satisfy for all $\theta \in \Theta$ that $Z^\theta = (U^\theta, W^\theta)$, let $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}}, \|\cdot\|_{\mathcal{H}}) = (\mathbb{R}, \langle\cdot,\cdot\rangle_{\mathbb{R}}, |\cdot|)$, let $\mathscr{S} = \sigma_{L(\mathcal{Y},\mathcal{H})}\big(\{\{\varphi \in L(\mathcal{Y},\mathcal{H})\colon \varphi(v) \in \mathcal{B}\} \subseteq L(\mathcal{Y},\mathcal{H})\colon v \in \mathcal{Y}, \mathcal{B} \in \mathscr{B}(\mathcal{H})\}\big)$, let $\psi_k\colon \Omega \to L(\mathcal{Y},\mathcal{H})$, $k \in \mathbb{N}_0$, satisfy for all $k \in \mathbb{N}_0$, $\omega \in \Omega$, $v \in \mathcal{Y}$ that

$$[\psi_k(\omega)](v) = \begin{cases} v(0,\xi) & : k = 0 \\ \sqrt{\frac{(\mathbf{U}(\omega))^{k-1}}{(k-1)!}}\, v\big(\mathbf{U}(\omega)T, \xi + \mathbf{W}_{\mathbf{U}(\omega)T}(\omega)\big) & : k \in \mathbb{N} \end{cases}, \tag{219}$$

and let $\Phi_l\colon \mathcal{Y} \times \mathcal{Y} \times \mathcal{Z} \to \mathcal{Y}$, $l \in \mathbb{N}_0$, satisfy for all $l \in \mathbb{N}_0$, $v, w \in \mathcal{Y}$, $\mathfrak{z} = (\mathfrak{u}, \mathfrak{w}) \in \mathcal{Z}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$\begin{aligned} &[\Phi_l(v,w,\mathfrak{z})](T-t,x) \\ &= \begin{cases} g(x + \mathfrak{w}_t) + tf\big(T-t+\mathfrak{u}t, x+\mathfrak{w}_{\mathfrak{u}t}, v(T-t+\mathfrak{u}t, x+\mathfrak{w}_{\mathfrak{u}t})\big) & : l = 0 \\ t\big[f\big(T-t+\mathfrak{u}t, x+\mathfrak{w}_{\mathfrak{u}t}, v(T-t+\mathfrak{u}t, x+\mathfrak{w}_{\mathfrak{u}t})\big) & \\ \quad - f\big(T-t+\mathfrak{u}t, x+\mathfrak{w}_{\mathfrak{u}t}, w(T-t+\mathfrak{u}t, x+\mathfrak{w}_{\mathfrak{u}t})\big)\big] & : l \in \mathbb{N} \end{cases} \end{aligned} \tag{220}$$

(cf. Lemma 3.5 and Corollary 3.7). Note that the assumption that $\forall\, t \in [0,T], x \in \mathbb{R}^d, v, w \in \mathbb{R}\colon \big(\max\{|f(t,x,0)|, |g(x)|\} \le L\max\{1, \|x\|_{\mathbb{R}^d}^p\}$ and $|f(t,x,v) - f(t,x,w)| \le L|v-w|\big)$ ensures that there exists a unique at most polynomially growing viscosity solution $y \in C([0,T]\times\mathbb{R}^d, \mathbb{R})$ of

$$\big(\tfrac{\partial y}{\partial t}\big)(t,x) + \tfrac{1}{2}(\Delta_x y)(t,x) + f(t,x,y(t,x)) = 0 \tag{221}$$

with $y(T,x) = g(x)$ for $(t,x) \in (0,T)\times\mathbb{R}^d$ (cf., e.g., Hairer, Hutzenthaler, & Jentzen [57, Section 4], Hutzenthaler et al. [68, Corollary 3.11], and Beck et al. [8, Theorem 1.1]). This shows (i). Moreover, the Feynman–Kac formula proves for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$y(t,x) = \mathbb{E}\bigg[g(x + \mathbf{W}_{T-t}) + \int_t^T f\big(s, x + \mathbf{W}_{s-t}, y(s, x + \mathbf{W}_{s-t})\big)\, \mathrm{d}s\bigg] \tag{222}$$

(cf., e.g., [57, Section 4], [68, Corollary 3.11], and [8, Theorem 1.1]). Combining this with [68, Corollary 3.11] demonstrates that

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \left[\frac{|y(t,x)|}{\max\{1, \|x\|_{\mathbb{R}^d}^p\}}\right] < \infty. \tag{223}$$

Next observe that

- it holds that $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ is a separable $\mathbb{R}$-Banach space (cf. (ii) in Proposition 3.4),
- it holds that $\min\{\mathfrak{B}, \kappa, C\} \ge 1$, $y \in \mathcal{Y}$ (cf. (223) and Lemma 3.5), and $\mathfrak{y}_{-1}, \mathfrak{y}_0 \in \mathcal{Y}$,
- it holds that $(\mathcal{Z}, \mathscr{Z})$ is a measurable space,
- it holds that $Z^\theta\colon \Omega \to \mathcal{Z}$, $\theta \in \Theta$, are i.i.d. $\mathscr{F}/\mathscr{Z}$-measurable functions,
- it holds that $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}}, \|\cdot\|_{\mathcal{H}})$ is a separable $\mathbb{R}$-Hilbert space,

- it holds that $\psi_k\colon \Omega \to L(\mathcal{Y}, \mathcal{H})$, $k \in \mathbb{N}_0$, are $\mathscr{F}/\mathscr{S}$-measurable functions (cf. Lemma 3.9),
- it holds that $(Z^\theta)_{\theta\in\Theta}$ and $(\psi_k)_{k\in\mathbb{N}_0}$ are independent,
- it holds that $\Phi_l\colon \mathcal{Y}\times\mathcal{Y}\times\mathcal{Z} \to \mathcal{Y}$, $l \in \mathbb{N}_0$, are $(\mathscr{B}(\mathcal{Y})\otimes\mathscr{B}(\mathcal{Y})\otimes\mathscr{Z})/\mathscr{B}(\mathcal{Y})$-measurable functions (cf. (ii) in Lemma 3.10),
- it holds for all $n \in (\mathbb{N}_0\cup\{-1\})$, $j \in \mathbb{N}$, $\theta \in \Theta$ that $Y^\theta_{n,j}(\Omega) \subseteq \mathcal{Y}$ (cf. assumption (212) and (i) in Lemma 3.11),
- it holds for all $n, j \in \mathbb{N}$, $\theta \in \Theta$ that $Y^\theta_{-1,j} = \mathfrak{y}_{-1}$, $Y^\theta_{0,j} = \mathfrak{y}_0$, and

$$Y^\theta_{n,j} = \sum_{l=0}^{n-1} \tfrac{1}{(M_j)^{n-l}} \Bigg[\sum_{i=1}^{(M_j)^{n-l}} \Phi_l\big(Y^{(\theta,l,i)}_{l,j}, Y^{(\theta,-l,i)}_{l-1,j}, Z^{(\theta,l,i)}\big)\Bigg] \tag{224}$$

  (cf. assumption (212) and (ii) in Lemma 3.11),
- it holds for all $n, j \in \mathbb{N}$ that $\mathrm{Cost}_{-1,j} = \mathrm{Cost}_{0,j} = 0$ and

$$\begin{aligned} \mathrm{Cost}_{n,j} &\leq (M_j)^n d + \sum_{l=0}^{n-1}\big[(M_j)^{n-l}(\mathrm{Cost}_{l,j} + \mathrm{Cost}_{l-1,j} + d + 1)\big] \\ &\leq (M_j)^n \mathfrak{z} + \sum_{l=0}^{n-1}\big[(M_j)^{n-l}(\mathrm{Cost}_{l,j} + \mathrm{Cost}_{l-1,j} + \gamma\mathfrak{z})\big] \end{aligned} \tag{225}$$

  (cf. assumption (213)),
- it holds for all $k \in \mathbb{N}_0$, $j \in \mathbb{N}$ that $\mathbb{E}\big[\|\Phi_k(Y^0_{k,j}, Y^1_{k-1,j}, Z^0)\|_{\mathcal{Y}}\big] < \infty$ (cf. Lemma 3.12), and
- it holds for all $k \in \mathbb{N}_0$, $n, j \in \mathbb{N}$, $u, v \in \mathcal{Y}$ that

$$\max\big\{\mathbb{E}\big[\|\psi_k(\Phi_0(\mathfrak{y}_0, \mathfrak{y}_{-1}, Z^0))\|^2_{\mathcal{H}}\big], \mathbb{1}_{\mathbb{N}}(k)\,\mathbb{E}\big[\|\psi_k(\mathfrak{y}_0 - y)\|^2_{\mathcal{H}}\big]\big\} \leq \tfrac{C^2}{k!}, \tag{226}$$

$$\mathbb{E}\big[\|\psi_k(\Phi_n(u, v, Z^0))\|^2_{\mathcal{H}}\big] \leq c\,\mathbb{E}\big[\|\psi_{k+1}(u - v)\|^2_{\mathcal{H}}\big], \tag{227}$$

$$\mathbb{E}\Bigg[\Bigg\|\psi_k\Bigg(y - \sum_{l=0}^{n-1}\mathbb{E}\big[\Phi_l\big(Y^0_{l,j}, Y^1_{l-1,j}, Z^0\big)\big]\Bigg)\Bigg\|^2_{\mathcal{H}}\Bigg] \leq 2c\,\mathbb{E}\Big[\big\|\psi_{k+1}\big(Y^0_{n-1,j} - y\big)\big\|^2_{\mathcal{H}}\Big] \tag{228}$$

  (cf. (222), assumption (211), Lemma 3.13, Lemma 3.14, and Lemma 3.15).

Corollary 2.15 hence establishes (ii)–(iv). The proof of Proposition 3.16 is thus complete. □

3.3.2 MLP approximations in variable space dimensions

Theorem 3.17. *Let* $T \in (0, \infty)$, $K, L, p, \mathfrak{B}, \kappa \in [0, \infty)$, $\Theta = \cup_{n=1}^\infty \mathbb{Z}^n$, $(M_j)_{j\in\mathbb{N}} \subseteq \mathbb{N}$ *satisfy* $\liminf_{j\to\infty} M_j = \infty$, $\sup_{j\in\mathbb{N}} {}^{M_{j+1}}/_{M_j} \leq \mathfrak{B}$, *and* $\sup_{j\in\mathbb{N}} {}^{M_j}/_j \leq \kappa$, *let* $\xi_d \in \mathbb{R}^d$, $d \in \mathbb{N}$, *satisfy* $\sup_{d\in\mathbb{N}}\|\xi_d\|_{\mathbb{R}^d} \leq K$, *for every* $d \in \mathbb{N}$ *let* $f_d \in C([0, T] \times \mathbb{R}^d \times \mathbb{R}, \mathbb{R})$, $g_d \in C(\mathbb{R}^d, \mathbb{R})$ *satisfy for all* $t \in [0, T]$, $x \in \mathbb{R}^d$, $v, w \in \mathbb{R}$ *that* $\max\{|f_d(t, x, 0)|, |g_d(x)|\} \leq L\max\{1, \|x\|^p_{\mathbb{R}^d}\}$ *and* $|f_d(t, x, v) - f_d(t, x, w)| \leq L|v - w|$, *let* $(\Omega, \mathscr{F}, \mathbb{P})$ *be a probability space, let* $U^\theta\colon \Omega \to [0, 1]$,

$\theta \in \Theta$, be independent on $[0,1]$ uniformly distributed random variables, for every $d \in \mathbb{N}$ let $W^{d,\theta}\colon [0,T] \times \Omega \to \mathbb{R}^d$, $\theta \in \Theta$, be independent standard Brownian motions, assume for every $d \in \mathbb{N}$ that $(U^\theta)_{\theta\in\Theta}$ and $(W^{d,\theta})_{\theta\in\Theta}$ are independent, let $Y^{d,\theta}_{n,j}\colon [0,T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}$, $\theta \in \Theta$, $d, j \in \mathbb{N}$, $n \in (\mathbb{N}_0 \cup \{-1\})$, satisfy for all $n, j, d \in \mathbb{N}$, $\theta \in \Theta$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that $Y^{d,\theta}_{-1,j}(t,x) = Y^{d,\theta}_{0,j}(t,x) = 0$ and

$$
\begin{aligned}
Y^{d,\theta}_{n,j}(T-t,x) = \tfrac{1}{(M_j)^n}\Bigg[\sum_{i=1}^{(M_j)^n} g_d\big(x + W^{d,(\theta,0,i)}_t\big)\Bigg] + \sum_{l=0}^{n-1} \tfrac{t}{(M_j)^{n-l}}\Bigg[\sum_{i=1}^{(M_j)^{n-l}} \\
\Big[f_d\Big(T-t+U^{(\theta,l,i)}t, x + W^{d,(\theta,l,i)}_{U^{(\theta,l,i)}t}, Y^{d,(\theta,l,i)}_{l,j}\big(T-t+U^{(\theta,l,i)}t, x+W^{d,(\theta,l,i)}_{U^{(\theta,l,i)}t}\big)\Big) \\
- \mathbb{1}_{\mathbb{N}}(l) f_d\Big(T-t+U^{d,(\theta,l,i)}t, x + W^{d,(\theta,l,i)}_{U^{(\theta,l,i)}t}, Y^{d,(\theta,-l,i)}_{l-1,j}\big(T-t+U^{(\theta,l,i)}t, x+W^{d,(\theta,l,i)}_{U^{(\theta,l,i)}t}\big)\Big)\Big]\Bigg],
\end{aligned}
\tag{229}
$$

and let $(\mathrm{Cost}_{d,n,j})_{(d,n,j)\in\mathbb{N}\times(\mathbb{N}_0\cup\{-1\})\times\mathbb{N}} \subseteq \mathbb{N}_0$ satisfy for all $d, n, j \in \mathbb{N}$ that $\mathrm{Cost}_{d,-1,j} = \mathrm{Cost}_{d,0,j} = 0$ and

$$
\mathrm{Cost}_{d,n,j} \le (M_j)^n d + \sum_{l=0}^{n-1} \big[(M_j)^{n-l}(\mathrm{Cost}_{d,l,j} + \mathrm{Cost}_{d,l-1,j} + d + 1)\big]. \tag{230}
$$

Then

(i) for every $d \in \mathbb{N}$ there exists a unique at most polynomially growing viscosity solution $y_d \in C([0,T] \times \mathbb{R}^d, \mathbb{R})$ of

$$
\big(\tfrac{\partial y_d}{\partial t}\big)(t,x) + \tfrac{1}{2}(\Delta_x y_d)(t,x) + f_d(t,x,y_d(t,x)) = 0 \tag{231}
$$

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

(ii) there exists $(N_{d,\varepsilon})_{(d,\varepsilon)\in\mathbb{N}\times(0,1]} \subseteq \mathbb{N}$ such that it holds for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$, $\delta \in (0,\infty)$ that $\sup_{n\in\{N_{d,\varepsilon},N_{d,\varepsilon}+1,\dots\}}\big(\mathbb{E}\big[|Y^{d,0}_{n,n}(0,\xi_d) - y_d(0,\xi_d)|^2\big]\big)^{1/2} \le \varepsilon$ and

$$
\begin{aligned}
\mathrm{Cost}_{d,N_{d,\varepsilon},N_{d,\varepsilon}} \le \Big[[4^{p+2}\max\{L,1\}(1+T)^{p/2+1}e^{LT}(\max\{K,p,1\})^p]^{2(1+\delta)}e^{\kappa} \\
\cdot \Big(1 + \sup\nolimits_{n\in\mathbb{N}}\Big[\tfrac{[5\mathfrak{B}e^{2\kappa}(1+(2LT)^2)]^{(1+\delta)}}{(M_n)^\delta}\Big]^n\Big)\Big] d^{1+p(1+\delta)}\varepsilon^{-2(1+\delta)} < \infty.
\end{aligned}
\tag{232}
$$

Proof of Theorem 3.17. Throughout this proof assume w.l.o.g. for every $d \in \mathbb{N}$ that $W^{d,\theta}\colon [0,T] \times \Omega \to \mathbb{R}^d$, $\theta \in \Theta$, are independent standard Brownian motions with continuous sample paths (cf., e.g., Klenke [74, Definition 21.8]) and throughout this proof let $C_d \in [1,\infty)$, $d \in \mathbb{N}$, be the real numbers which satisfy for all $d \in \mathbb{N}$ that

$$
C_d = \max\Big\{1, e^{LT}\Big[\big(\mathbb{E}\big[|g_d(\xi_d + W^{d,0}_T)|^2\big]\big)^{1/2} + \sqrt{T}\big(\textstyle\int_0^T \mathbb{E}\big[|f_d(t,\xi_d + W^{d,0}_t, 0)|^2\big]\,\mathrm{d}t\big)^{1/2}\Big]\Big\}. \tag{233}
$$

First of all, observe that the Burkholder–Davis–Gundy-type inequality in Da Prato & Zabczyk [35, Lemma 7.7] establishes for all $r \in [2,\infty)$, $d \in \mathbb{N}$, $t \in [0,T]$ that

$$
\big(\mathbb{E}\big[\|W^{d,0}_t\|^r_{\mathbb{R}^d}\big]\big)^{1/r} \le \sqrt{\tfrac{1}{2}r(r-1)td} \le r\sqrt{\tfrac{Td}{2}}. \tag{234}
$$

Jensen's inequality and the fact that $\forall\, a, b, r \in [0, \infty)\colon (a+b)^r \le 2^{\max\{r-1,0\}}(a^r + b^r)$ hence prove for all $d \in \mathbb{N}$ that

$$
\begin{aligned}
&\big(\mathbb{E}\big[|g_d(\xi_d + W_T^{d,0})|^2\big]\big)^{1/2} + \sqrt{T}\big(\textstyle\int_0^T \mathbb{E}\big[|f_d(t, \xi_d + W_t^{d,0}, 0)|^2\big]\,\mathrm{d}t\big)^{1/2} \\
&\le L\big(\mathbb{E}\big[\max\{1, \|\xi_d + W_T^{d,0}\|_{\mathbb{R}^d}^{2p}\}\big]\big)^{1/2} + L\sqrt{T}\big(\textstyle\int_0^T \mathbb{E}\big[\max\{1, \|\xi_d + W_t^{d,0}\|_{\mathbb{R}^d}^{2p}\}\big]\,\mathrm{d}t\big)^{1/2} \\
&\le L\Big(1 + \big(\mathbb{E}\big[\|\xi_d + W_T^{d,0}\|_{\mathbb{R}^d}^{2p}\big]\big)^{1/2}\Big) + LT\big(1 + \tfrac{1}{T}\textstyle\int_0^T \mathbb{E}\big[\|\xi_d + W_t^{d,0}\|_{\mathbb{R}^d}^{2p}\big]\,\mathrm{d}t\big)^{1/2} \\
&\le L(1+T) + L\Big(\big(\mathbb{E}\big[\|\xi_d + W_T^{d,0}\|_{\mathbb{R}^d}^{2\max\{p,1\}}\big]\big)^{\frac{1}{2\max\{p,1\}}}\Big)^p \\
&\quad + LT\Big(\big(\tfrac{1}{T}\textstyle\int_0^T \mathbb{E}\big[\|\xi_d + W_t^{d,0}\|_{\mathbb{R}^d}^{2\max\{p,1\}}\big]\,\mathrm{d}t\big)^{\frac{1}{2\max\{p,1\}}}\Big)^p \\
&\le L(1+T) + L\Big(\|\xi_d\|_{\mathbb{R}^d} + \big(\mathbb{E}\big[\|W_T^{d,0}\|_{\mathbb{R}^d}^{2\max\{p,1\}}\big]\big)^{\frac{1}{2\max\{p,1\}}}\Big)^p \\
&\quad + LT\Big(\|\xi_d\|_{\mathbb{R}^d} + \big(\tfrac{1}{T}\textstyle\int_0^T \mathbb{E}\big[\|W_t^{d,0}\|_{\mathbb{R}^d}^{2\max\{p,1\}}\big]\,\mathrm{d}t\big)^{\frac{1}{2\max\{p,1\}}}\Big)^p \\
&\le L(1+T) + L\big(\|\xi_d\|_{\mathbb{R}^d} + \max\{p,1\}\sqrt{2Td}\big)^p + LT\big(\|\xi_d\|_{\mathbb{R}^d} + \max\{p,1\}\sqrt{2Td}\big)^p \\
&\le L(1+T) + L(1+T)2^{\max\{p-1,0\}}\big(\|\xi_d\|_{\mathbb{R}^d}^p + \max\{p^p,1\}(2Td)^{p/2}\big) \\
&\le d^{p/2}2^{\max\{p,1\}+p/2}L(1+T)^{p/2+1}\big(K^p + \max\{p^p,1\}\big) \\
&\le d^{p/2}4^{p+1}L(1+T)^{p/2+1}(\max\{K,p,1\})^p.
\end{aligned}
\tag{235}
$$

This and (233) show for all $d \in \mathbb{N}$, $\delta \in (0,\infty)$ that

$$
\begin{aligned}
5(C_d)^{2(1+\delta)} &\le 5[d^{p/2}4^{p+1}\max\{L,1\}(1+T)^{p/2+1}e^{LT}(\max\{K,p,1\})^p]^{2(1+\delta)} \\
&\le [4^{p+2}\max\{L,1\}(1+T)^{p/2+1}e^{LT}(\max\{K,p,1\})^p]^{2(1+\delta)}d^{p(1+\delta)}.
\end{aligned}
\tag{236}
$$

Combining this with Proposition 3.16 completes the proof of Theorem 3.17. □

Corollary 3.18. *Let $T \in (0,\infty)$, $p \in [0,\infty)$, $\Theta = \cup_{n=1}^\infty \mathbb{Z}^n$, $(M_j)_{j\in\mathbb{N}} \subseteq \mathbb{N}$, $(\xi_d)_{d\in\mathbb{N}} \subseteq \mathbb{R}^d$ satisfy $\sup_{j\in\mathbb{N}}(M_{j+1}/M_j + M_j/j + \|\xi_j\|_{\mathbb{R}^j}) < \infty = \liminf_{j\to\infty} M_j$, let $f\colon \mathbb{R} \to \mathbb{R}$ be a Lipschitz continuous function, let $g_d \in C(\mathbb{R}^d, \mathbb{R})$, $d \in \mathbb{N}$, satisfy $\sup_{d\in\mathbb{N}, x\in\mathbb{R}^d} |g_d(x)|/\max\{1,\|x\|_{\mathbb{R}^d}^p\} < \infty$, let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, let $U^\theta\colon \Omega \to [0,1]$, $\theta \in \Theta$, be independent on $[0,1]$ uniformly distributed random variables, 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, assume that $(U^\theta)_{\theta\in\Theta}$ and $(W^{d,\theta})_{(d,\theta)\in\mathbb{N}\times\Theta}$ are independent, let $Y_{n,j}^{d,\theta}\colon [0,T]\times\mathbb{R}^d\times\Omega \to \mathbb{R}$, $\theta \in \Theta$, $d, j \in \mathbb{N}$, $n \in (\mathbb{N}_0 \cup \{-1\})$, satisfy for all $n, j, d \in \mathbb{N}$, $\theta \in \Theta$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that $Y_{-1,j}^{d,\theta}(t,x) = Y_{0,j}^{d,\theta}(t,x) = 0$ and*

$$
\begin{aligned}
Y_{n,j}^{d,\theta}(T-t,x) &= \sum_{l=0}^{n-1} \tfrac{t}{(M_j)^{n-l}}\Bigg[\sum_{i=1}^{(M_j)^{n-l}} \Big[f\Big(Y_{l,j}^{d,(\theta,l,i)}\big(T-t+U^{(\theta,l,i)}t, x + W_{U^{(\theta,l,i)}t}^{d,(\theta,l,i)}\big)\Big) \\
&\quad - \mathbb{1}_{\mathbb{N}}(l)f\Big(Y_{l-1,j}^{d,(\theta,-l,i)}\big(T-t+U^{(\theta,l,i)}t, x + W_{U^{(\theta,l,i)}t}^{d,(\theta,l,i)}\big)\Big)\Big]\Bigg] + \tfrac{1}{(M_j)^n}\Bigg[\sum_{i=1}^{(M_j)^n} g_d\big(x + W_t^{d,(\theta,0,i)}\big)\Bigg],
\end{aligned}
\tag{237}
$$

and let $(\mathrm{Cost}_{d,n,j})_{(d,n,j)\in\mathbb{N}\times(\mathbb{N}_0\cup\{-1\})\times\mathbb{N}} \subseteq \mathbb{N}_0$ satisfy for all $d, n, j \in \mathbb{N}$ that $\mathrm{Cost}_{d,-1,j} = \mathrm{Cost}_{d,0,j} = 0$ and

$$
\mathrm{Cost}_{d,n,j} \le (M_j)^n d + \sum_{l=0}^{n-1}\big[(M_j)^{n-l}(\mathrm{Cost}_{d,l,j} + \mathrm{Cost}_{d,l-1,j} + d + 1)\big].
\tag{238}
$$

Then

(i) for every $d \in \mathbb{N}$ there exists a unique at most polynomially growing viscosity solution $y_d \in C([0,T] \times \mathbb{R}^d, \mathbb{R})$ of

$$\big(\tfrac{\partial y_d}{\partial t}\big)(t,x) + \tfrac{1}{2}(\Delta_x y_d)(t,x) + f(y_d(t,x)) = 0 \tag{239}$$

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

(ii) there exist $(N_{d,\varepsilon})_{(d,\varepsilon)\in\mathbb{N}\times(0,1]} \subseteq \mathbb{N}$ and $(C_\delta)_{\delta\in(0,\infty)} \subseteq (0,\infty)$ such that it holds for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$, $\delta \in (0,\infty)$ that $\operatorname{Cost}_{d,N_{d,\varepsilon},N_{d,\varepsilon}} \leq C_\delta d^{1+p(1+\delta)}\varepsilon^{-2(1+\delta)}$ and

$$\sup_{n\in\{N_{d,\varepsilon},N_{d,\varepsilon}+1,\ldots\}} \big(\mathbb{E}\big[|Y^{d,0}_{n,n}(0,\xi_d) - y_d(0,\xi_d)|^2\big]\big)^{1/2} \leq \varepsilon. \tag{240}$$

Acknowledgements

This project has been partially supported through the ETH Research Grant ETH-47 15-2 "Mild stochastic calculus and numerical approximations for nonlinear stochastic evolution equations with Lévy noise" and by the project "Deep artificial neural network approximations for stochastic partial differential equations: Algorithms and convergence proofs" (project number 184220) funded by the Swiss National Science Foundation (SNSF).

References

[1] Agarwal, A., and Claisse, J. Branching diffusion representation of semi-linear elliptic PDEs and estimation using Monte Carlo method. *ArXiv e-prints* (2017), 24 pages. arXiv: `1704.00328 [math.PR]`.

[2] Agarwal, R. P. *Difference equations and inequalities*. Second. Vol. 228. Monographs and Textbooks in Pure and Applied Mathematics. Theory, methods, and applications. Marcel Dekker, Inc., New York, 2000, xvi+971. isbn: 0-8247-9007-3.

[3] Aliprantis, C. D., and Border, K. C. *Infinite dimensional analysis*. Third. A hitchhiker's guide. Springer, Berlin, 2006, xxii+703. isbn: 978-3-540-32696-0.

[4] Bally, V., and Pagès, G. A quantization algorithm for solving multi-dimensional discrete-time optimal stopping problems. *Bernoulli* 9, 6 (2003), 1003–1049. issn: 1350-7265. doi: `10.3150/bj/1072215199`. url: `https://doi.org/10.3150/bj/1072215199`.

[5] Beck, C., Becker, S., Cheridito, P., Jentzen, A., and Neufeld, A. Deep splitting method for parabolic PDEs. *ArXiv e-prints* (2019), 40 pages. arXiv: `1907.03452 [math.NA]`.

[6] Beck, C., Becker, S., Grohs, P., Jaafari, N., and Jentzen, A. Solving stochastic differential equations and Kolmogorov equations by means of deep learning. *ArXiv e-prints* (2018), 56 pages. arXiv: `1806.00421 [math.NA]`.

[7] 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. *J. Nonlinear Sci.* 29, 4 (2019), 1563–1619. issn: 0938-8974. doi: `10.1007/s00332-018-9525-3`. url: `https://doi.org/10.1007/s00332-018-9525-3`.

[8] Beck, C., Gonon, L., Hutzenthaler, M., and Jentzen, A. On existence and uniqueness properties for solutions of stochastic fixed point equations. *ArXiv e-prints* (2019), 33 pages. arXiv: 1908.03382 [math.PR].

[9] Beck, C., Hornung, F., Hutzenthaler, M., Jentzen, A., and Kruse, T. Overcoming the curse of dimensionality in the numerical approximation of Allen-Cahn partial differential equations via truncated full-history recursive multilevel Picard approximations. *ArXiv e-prints* (2019), 31 pages. arXiv: 1907.06729 [math.NA].

[10] Becker, S., Cheridito, P., and Jentzen, A. Deep optimal stopping. *J. Mach. Learn. Res.* 20 (2019), Paper No. 74, 25. issn: 1532-4435.

[11] Becker, S., Cheridito, P., Jentzen, A., and Welti, T. Solving high-dimensional optimal stopping problems using deep learning. *ArXiv e-prints* (2019), 42 pages. arXiv: 1908.01602 [cs.CE].

[12] Belak, C., Hoffmann, D., and Seifried, F. T. Probabilistic Representations of Nonlocal Nonlinear PDEs via Branching Diffusions with Jumps (2019), 28 pages. Available at SSRN: http://dx.doi.org/10.2139/ssrn.3451280.

[13] Bellman, R. *Dynamic programming.* Princeton University Press, Princeton, N. J., 1957, xxv+342.

[14] Bender, C., and Denk, R. A forward scheme for backward SDEs. *Stochastic Process. Appl.* 117, 12 (2007), 1793–1812. issn: 0304-4149. doi: 10.1016/j.spa.2007.03.005. url: https://doi.org/10.1016/j.spa.2007.03.005.

[15] Bender, C., Schweizer, N., and Zhuo, J. A primal-dual algorithm for BSDEs. *Math. Finance* 27, 3 (2017), 866–901. issn: 0960-1627. doi: 10.1111/mafi.12100. url: https://doi.org/10.1111/mafi.12100.

[16] Berg, J., and Nyström, K. A unified deep artificial neural network approach to partial differential equations in complex geometries. *Neurocomputing* 317 (Nov. 2018), 28–41. issn: 0925-2312. doi: 10.1016/j.neucom.2018.06.056. url: http://dx.doi.org/10.1016/j.neucom.2018.06.056.

[17] Bernal, F., dos Reis, G., and Smith, G. Hybrid PDE solver for data-driven problems and modern branching. *European J. Appl. Math.* 28, 6 (2017), 949–972. issn: 0956-7925. doi: 10.1017/S0956792517000109. url: https://doi.org/10.1017/S0956792517000109.

[18] Berner, J., Grohs, P., and Jentzen, A. Analysis of the generalization error: Empirical risk minimization over deep artificial neural networks overcomes the curse of dimensionality in the numerical approximation of Black-Scholes partial differential equations. *ArXiv e-prints* (2018), 35 pages. arXiv: 1809.03062 [cs.LG].

[19] Bouchard, B., Tan, X., and Warin, X. Numerical approximation of general Lipschitz BSDEs with branching processes. In: *CEMRACS 2017—numerical methods for stochastic models: control, uncertainty quantification, mean-field.* Vol. 65. ESAIM Proc. Surveys. EDP Sci., Les Ulis, 2019, 309–329.

[20] Bouchard, B., Tan, X., Warin, X., and Zou, Y. Numerical approximation of BSDEs using local polynomial drivers and branching processes. *Monte Carlo Methods Appl.* 23, 4 (2017), 241–263. issn: 0929-9629. doi: 10.1515/mcma-2017-0116. url: https://doi.org/10.1515/mcma-2017-0116.

[21] Bouchard, B., and Touzi, N. Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Process. Appl.* 111, 2 (2004), 175–206. issn: 0304-4149. doi: 10.1016/j.spa.2004.01.001. url: https://doi.org/10.1016/j.spa.2004.01.001.

[22] Briand, P., and Labart, C. Simulation of BSDEs by Wiener chaos expansion. *Ann. Appl. Probab.* 24, 3 (2014), 1129–1171. issn: 1050-5164. doi: 10.1214/13-AAP943. url: https://doi.org/10.1214/13-AAP943.

[23] Chang, D., Liu, H., and Xiong, J. A branching particle system approximation for a class of FBSDEs. *Probab. Uncertain. Quant. Risk* 1 (2016), Paper No. 9, 34. issn: 2367-0126. doi: 10.1186/s41546-016-0007-y. url: https://doi.org/10.1186/s41546-016-0007-y.

[24] Chan-Wai-Nam, Q., Mikael, J., and Warin, X. Machine learning for semi linear PDEs. *J. Sci. Comput.* 79, 3 (2019), 1667–1712. issn: 0885-7474. doi: 10.1007/s10915-019-00908-3. url: https://doi.org/10.1007/s10915-019-00908-3.

[25] Chassagneux, J.-F. Linear multistep schemes for BSDEs. *SIAM J. Numer. Anal.* 52, 6 (2014), 2815–2836. issn: 0036-1429. doi: 10.1137/120902951. url: https://doi.org/10.1137/120902951.

[26] Chassagneux, J.-F., and Crisan, D. Runge-Kutta schemes for backward stochastic differential equations. *Ann. Appl. Probab.* 24, 2 (2014), 679–720. issn: 1050-5164. doi: 10.1214/13-AAP933. url: https://doi.org/10.1214/13-AAP933.

[27] Chassagneux, J.-F., and Richou, A. Numerical stability analysis of the Euler scheme for BSDEs. *SIAM J. Numer. Anal.* 53, 2 (2015), 1172–1193. issn: 0036-1429. doi: 10.1137/140977047. url: https://doi.org/10.1137/140977047.

[28] Chassagneux, J.-F., and Richou, A. Numerical simulation of quadratic BSDEs. *Ann. Appl. Probab.* 26, 1 (2016), 262–304. issn: 1050-5164. doi: 10.1214/14-AAP1090. url: https://doi.org/10.1214/14-AAP1090.

[29] Chen, Y., and Wan, J. W. L. Deep Neural Network Framework Based on Backward Stochastic Differential Equations for Pricing and Hedging American Options in High Dimensions. *ArXiv e-prints* (2019), 35 pages. arXiv: 1909.11532 [q-fin.CP].

[30] Conway, J. B. *A course in functional analysis*. Second. Vol. 96. Graduate Texts in Mathematics. Springer-Verlag, New York, 1990, xvi+399. isbn: 0-387-97245-5.

[31] Crisan, D., and Manolarakis, K. Probabilistic methods for semilinear partial differential equations. Applications to finance. *M2AN Math. Model. Numer. Anal.* 44, 5 (2010), 1107–1133. issn: 0764-583X. doi: 10.1051/m2an/2010054. url: https://doi.org/10.1051/m2an/2010054.

[32] Crisan, D., and Manolarakis, K. Solving backward stochastic differential equations using the cubature method: application to nonlinear pricing. *SIAM J. Financial Math.* 3, 1 (2012), 534–571. issn: 1945-497X. doi: 10.1137/090765766. url: https://doi.org/10.1137/090765766.

[33] Crisan, D., and Manolarakis, K. Second order discretization of backward SDEs and simulation with the cubature method. *Ann. Appl. Probab.* 24, 2 (2014), 652–678. issn: 1050-5164. doi: 10.1214/13-AAP932. url: https://doi.org/10.1214/13-AAP932.

[34] Crisan, D., Manolarakis, K., and Touzi, N. On the Monte Carlo simulation of BSDEs: an improvement on the Malliavin weights. *Stochastic Process. Appl.* 120, 7 (2010), 1133–1158. issn: 0304-4149. doi: 10.1016/j.spa.2010.03.015. url: https://doi.org/10.1016/j.spa.2010.03.015.

[35] Da Prato, G., and Zabczyk, J. *Stochastic equations in infinite dimensions.* Vol. 44. Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 1992, xviii+454. isbn: 0-521-38529-6. doi: 10.1017/CBO9780511666223. url: https://doi.org/10.1017/CBO9780511666223.

[36] Delarue, F., and Menozzi, S. A forward-backward stochastic algorithm for quasi-linear PDEs. *Ann. Appl. Probab.* 16, 1 (2006), 140–184. issn: 1050-5164. doi: 10.1214/105051605000000674. url: https://doi.org/10.1214/105051605000000674.

[37] Dockhorn, T. A Discussion on Solving Partial Differential Equations using Neural Networks. *ArXiv e-prints* (2019), 9 pages. arXiv: 1904.07200 [cs.LG].

[38] Douglas Jr., J., Ma, J., and Protter, P. Numerical methods for forward-backward stochastic differential equations. *Ann. Appl. Probab.* 6, 3 (1996), 940–968. issn: 1050-5164. doi: 10.1214/aoap/1034968235. url: https://doi.org/10.1214/aoap/1034968235.

[39] 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. issn: 2194-6701. doi: 10.1007/s40304-017-0117-6. url: https://doi.org/10.1007/s40304-017-0117-6.

[40] E, W., Hutzenthaler, M., Jentzen, A., and Kruse, T. Multilevel Picard iterations for solving smooth semilinear parabolic heat equations. *ArXiv e-prints* (2016), 19 pages. arXiv: 1607.03295 [math.NA].

[41] 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. *J. Sci. Comput.* 79, 3 (2019), 1534–1571. issn: 0885-7474. doi: 10.1007/s10915-018-00903-0. url: https://doi.org/10.1007/s10915-018-00903-0.

[42] E, W., and Yu, B. The deep Ritz method: a deep learning-based numerical algorithm for solving variational problems. *Commun. Math. Stat.* 6, 1 (2018), 1–12. issn: 2194-6701. doi: 10.1007/s40304-018-0127-z. url: https://doi.org/10.1007/s40304-018-0127-z.

[43] Elbrächter, D., Grohs, P., Jentzen, A., and Schwab, C. DNN Expression Rate Analysis of High-dimensional PDEs: Application to Option Pricing. *ArXiv e-prints* (2018), 50 pages. arXiv: 1809.07669 [math.FA].

[44] Farahmand, A.-m., Nabi, S., and Nikovski, D. N. Deep reinforcement learning for partial differential equation control. In: *2017 American Control Conference (ACC).* May 2017, 3120–3127. doi: 10.23919/ACC.2017.7963427.

[45] Fujii, M., Takahashi, A., and Takahashi, M. Asymptotic Expansion as Prior Knowledge in Deep Learning Method for High dimensional BSDEs. *Asia-Pacific Financial Markets* (2019). issn: 1573-6946. doi: 10.1007/s10690-019-09271-7. url: https://doi.org/10.1007/s10690-019-09271-7.

[46] Geiss, C., and Labart, C. Simulation of BSDEs with jumps by Wiener chaos expansion. *Stochastic Process. Appl.* 126, 7 (2016), 2123–2162. issn: 0304-4149. doi: 10.1016/j.spa.2016.01.006. url: https://doi.org/10.1016/j.spa.2016.01.006.

[47] Gobet, E., and Labart, C. Solving BSDE with adaptive control variate. *SIAM J. Numer. Anal.* 48, 1 (2010), 257–277. issn: 0036-1429. doi: 10.1137/090755060. url: https://doi.org/10.1137/090755060.

[48] Gobet, E., and Lemor, J.-P. Numerical simulation of BSDEs using empirical regression methods: theory and practice. *ArXiv e-prints* (2008), 17 pages. arXiv: 0806.4447 [math.PR].

[49] Gobet, E., Lemor, J.-P., and Warin, X. A regression-based Monte Carlo method to solve backward stochastic differential equations. *Ann. Appl. Probab.* 15, 3 (2005), 2172–2202. issn: 1050-5164. doi: 10.1214/105051605000000412. url: https://doi.org/10.1214/105051605000000412.

[50] Gobet, E., López-Salas, J. G., Turkedjiev, P., and Vázquez, C. Stratified regression Monte-Carlo scheme for semilinear PDEs and BSDEs with large scale parallelization on GPUs. *SIAM J. Sci. Comput.* 38, 6 (2016), C652–C677. issn: 1064-8275. doi: 10.1137/16M106371X. url: https://doi.org/10.1137/16M106371X.

[51] Gobet, E., and Turkedjiev, P. Approximation of backward stochastic differential equations using Malliavin weights and least-squares regression. *Bernoulli* 22, 1 (2016), 530–562. issn: 1350-7265. doi: 10.3150/14-BEJ667. url: https://doi.org/10.3150/14-BEJ667.

[52] Gobet, E., and Turkedjiev, P. Linear regression MDP scheme for discrete backward stochastic differential equations under general conditions. *Math. Comp.* 85, 299 (2016), 1359–1391. issn: 0025-5718. doi: 10.1090/mcom/3013. url: https://doi.org/10.1090/mcom/3013.

[53] Goudenège, L., Molent, A., and Zanette, A. Machine Learning for Pricing American Options in High Dimension. *ArXiv e-prints* (2019), 11 pages. arXiv: 1903.11275 [q-fin.CP].

[54] Grohs, P., Hornung, F., Jentzen, A., and von Wurstemberger, P. A proof that artificial neural networks overcome the curse of dimensionality in the numerical approximation of Black-Scholes partial differential equations. *ArXiv e-prints* (2018), 124 pages. arXiv: 1809.02362 [math.NA]. Revision requested from *Memoires of the American Mathematical Society.*

[55] Grohs, P., Hornung, F., Jentzen, A., and Zimmermann, P. Space-time error estimates for deep neural network approximations for differential equations. *ArXiv e-prints* (2019), 86 pages. arXiv: 1908.03833 [math.NA].

[56] Grohs, P., Jentzen, A., and Salimova, D. Deep neural network approximations for Monte Carlo algorithms. *ArXiv e-prints* (2019), 45 pages. arXiv: 1908.10828 [math.NA].

[57] Hairer, M., Hutzenthaler, M., and Jentzen, A. Loss of regularity for Kolmogorov equations. *Ann. Probab.* 43, 2 (2015), 468–527. issn: 0091-1798. doi: 10.1214/13-AOP838. url: https://doi.org/10.1214/13-AOP838.

[58] Han, J., Jentzen, A., and E, W. Solving high-dimensional partial differential equations using deep learning. *Proc. Natl. Acad. Sci. USA* 115, 34 (2018), 8505–8510. issn: 0027-8424. doi: 10.1073/pnas.1718942115. url: https://doi.org/10.1073/pnas.1718942115.

[59] Han, J., and Long, J. Convergence of the Deep BSDE Method for Coupled FBSDEs. *ArXiv e-prints* (2018), 27 pages. arXiv: 1811.01165 [math.PR].

[60] Henry-Labordère, P. Counterparty Risk Valuation: A Marked Branching Diffusion Approach. *ArXiv e-prints* (2012), 17 pages. arXiv: 1203.2369 [q-fin.CP].

[61] Henry-Labordère, P. Deep Primal-Dual Algorithm for BSDEs: Applications of Machine Learning to CVA and IM (2017), 16 pages. Available at SSRN: http://dx.doi.org/10.2139/ssrn.3071506.

[62] Henry-Labordère, P., Oudjane, N., Tan, X., Touzi, N., and Warin, X. Branching diffusion representation of semilinear PDEs and Monte Carlo approximation. *Ann. Inst. Henri Poincaré Probab. Stat.* 55, 1 (2019), 184–210. issn: 0246-0203. doi: 10.1214/17-aihp880. url: https://doi.org/10.1214/17-aihp880.

[63] 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. issn: 0304-4149. doi: 10.1016/j.spa.2013.10.005. url: https://doi.org/10.1016/j.spa.2013.10.005.

[64] Henry-Labordère, P., and Touzi, N. Branching diffusion representation for nonlinear Cauchyproblems and Monte Carlo approximation. *ArXiv e-prints* (2018), 29 pages. arXiv: 1801.08794 [math.PR].

[65] Huijskens, T. P., Ruijter, M. J., and Oosterlee, C. W. Efficient numerical Fourier methods for coupled forward-backward SDEs. *J. Comput. Appl. Math.* 296 (2016), 593–612. issn: 0377-0427. doi: 10.1016/j.cam.2015.10.019. url: https://doi.org/10.1016/j.cam.2015.10.019.

[66] Huré, C., Pham, H., and Warin, X. Some machine learning schemes for high-dimensional nonlinear PDEs. *ArXiv e-prints* (2019), 33 pages. arXiv: 1902.01599 [math.PR].

[67] Hutzenthaler, M., Jentzen, A., Kruse, T., and Nguyen, T. A. A proof that rectified deep neural networks overcome the curse of dimensionality in the numerical approximation of semilinear heat equations. *ArXiv e-prints* (2019), 29 pages. arXiv: 1901.10854 [math.NA].

[68] Hutzenthaler, M., Jentzen, A., Kruse, T., Nguyen, T. A., and von Wurstemberger, P. Overcoming the curse of dimensionality in the numerical approximation of semilinear parabolic partial differential equations. *ArXiv e-prints* (2018), 30 pages. arXiv: 1807.01212 [math.PR].

[69] Hutzenthaler, M., Jentzen, A., and von Wurstemberger, P. Overcoming the curse of dimensionality in the approximative pricing of financial derivatives with default risks. *ArXiv e-prints* (2019), 71 pages. arXiv: 1903.05985 [math.NA].

[70] Hutzenthaler, M., and Kruse, T. Multi-level Picard approximations of high-dimensional semilinear parabolic differential equations with gradient-dependent nonlinearities. *ArXiv e-prints* (2017), 19 pages. arXiv: 1711.01080 [math.NA].

[71] Jacquier, A., and Oumgari, M. Deep PPDEs for rough local stochastic volatility. *ArXiv e-prints* (2019), 21 pages. arXiv: 1906.02551 [q-fin.PR].

[72] Jentzen, A., Mazzonetto, S., and Salimova, D. Existence and uniqueness properties for solutions of a class of Banach space valued evolution equations. *ArXiv e-prints* (2018), 28 pages. arXiv: 1812.06859 [math.CA].

[73] Jentzen, A., Salimova, D., and Welti, T. A proof that deep artificial neural networks overcome the curse of dimensionality in the numerical approximation of Kolmogorov partial differential equations with constant diffusion and nonlinear drift coefficients. *ArXiv e-prints* (2018), 48 pages. arXiv: 1809.07321 [math.NA].

[74] Klenke, A. *Probability theory.* Second. Universitext. A comprehensive course. Springer, London, 2014, xii+638. isbn: 978-1-4471-5361-0. doi: 10.1007/978-1-4471-5361-0. url: https://doi.org/10.1007/978-1-4471-5361-0.

[75] Kutyniok, G., Petersen, P., Raslan, M., and Schneider, R. A Theoretical Analysis of Deep Neural Networks and Parametric PDEs. *ArXiv e-prints* (2019), 43 pages. arXiv: 1904.00377 [math.NA].

[76] Labart, C., and Lelong, J. A parallel algorithm for solving BSDEs. *Monte Carlo Methods Appl.* 19, 1 (2013), 11–39. issn: 0929-9629. doi: 10.1515/mcma-2013-0001. url: https://doi.org/10.1515/mcma-2013-0001.

[77] Lemor, J.-P., Gobet, E., and Warin, X. Rate of convergence of an empirical regression method for solving generalized backward stochastic differential equations. *Bernoulli* 12, 5 (2006), 889–916. issn: 1350-7265. doi: 10.3150/bj/1161614951. url: https://doi.org/10.3150/bj/1161614951.

[78] Lionnet, A., dos Reis, G., and Szpruch, L. Time discretization of FBSDE with polynomial growth drivers and reaction-diffusion PDEs. *Ann. Appl. Probab.* 25, 5 (2015), 2563–2625. issn: 1050-5164. doi: 10.1214/14-AAP1056. url: https://doi.org/10.1214/14-AAP1056.

[79] Long, Z., Lu, Y., Ma, X., and Dong, B. PDE-Net: Learning PDEs from Data. In: *Proceedings of the 35th International Conference on Machine Learning.* Ed. by Dy, J., and Krause, A. Vol. 80. Proceedings of Machine Learning Research. PMLR, July 2018, 3208–3216. url: http://proceedings.mlr.press/v80/long18a.html.

[80] Lye, K. O., Mishra, S., and Ray, D. Deep learning observables in computational fluid dynamics. *ArXiv e-prints* (2019), 57 pages. arXiv: 1903.03040 [physics.comp-ph].

[81] Ma, J., Protter, P., San Martín, J., and Torres, S. Numerical method for backward stochastic differential equations. *Ann. Appl. Probab.* 12, 1 (2002), 302–316. issn: 1050-5164. doi: 10.1214/aoap/1015961165. url: https://doi.org/10.1214/aoap/1015961165.

[82] Ma, J., Protter, P., and Yong, J. Solving forward-backward stochastic differential equations explicitly—a four step scheme. *Probab. Theory Related Fields* 98, 3 (1994), 339–359. issn: 0178-8051. doi: 10.1007/BF01192258. url: https://doi.org/10.1007/BF01192258.

[83] Ma, J., and Yong, J. *Forward-backward stochastic differential equations and their applications.* Vol. 1702. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1999, xiv+270. isbn: 3-540-65960-9.

[84] Magill, M., Qureshi, F., and de Haan, H. Neural Networks Trained to Solve Differential Equations Learn General Representations. In: *Advances in Neural Information Processing Systems 31*. Ed. by Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., and Garnett, R. Curran Associates, Inc., 2018, 4071–4081. url: http://papers.nips.cc/paper/7662-neural-networks-trained-to-solve-differential-equations-learn-general-representations.pdf.

[85] McKean, H. P. Application of Brownian motion to the equation of Kolmogorov-Petrovskii-Piskunov. *Comm. Pure Appl. Math.* 28, 3 (1975), 323–331. issn: 0010-3640. doi: 10.1002/cpa.3160280302. url: https://doi.org/10.1002/cpa.3160280302.

[86] Milstein, G. N., and Tretyakov, M. V. Numerical algorithms for forward-backward stochastic differential equations. *SIAM J. Sci. Comput.* 28, 2 (2006), 561–582. issn: 1064-8275. doi: 10.1137/040614426. url: https://doi.org/10.1137/040614426.

[87] Milstein, G. N., and Tretyakov, M. V. Discretization of forward-backward stochastic differential equations and related quasi-linear parabolic equations. *IMA J. Numer. Anal.* 27, 1 (2007), 24–44. issn: 0272-4979. doi: 10.1093/imanum/drl019. url: https://doi.org/10.1093/imanum/drl019.

[88] Novak, E., and Woźniakowski, H. *Tractability of multivariate problems. Vol. 1: Linear information*. Vol. 6. EMS Tracts in Mathematics. European Mathematical Society (EMS), Zürich, 2008, xii+384. isbn: 978-3-03719-026-5. doi: 10.4171/026. url: https://doi.org/10.4171/026.

[89] Novak, E., and Woźniakowski, H. *Tractability of multivariate problems. Volume II: Standard information for functionals*. Vol. 12. EMS Tracts in Mathematics. European Mathematical Society (EMS), Zürich, 2010, xviii+657. isbn: 978-3-03719-084-5. doi: 10.4171/084. url: https://doi.org/10.4171/084.

[90] Pham, H. Feynman-Kac representation of fully nonlinear PDEs and applications. *Acta Math. Vietnam.* 40, 2 (2015), 255–269. issn: 0251-4184. doi: 10.1007/s40306-015-0128-x. url: https://doi.org/10.1007/s40306-015-0128-x.

[91] Pham, H., and Warin, X. Neural networks-based backward scheme for fully nonlinear PDEs. *ArXiv e-prints* (2019), 15 pages. arXiv: 1908.00412 [math.OC].

[92] Raissi, M. Deep Hidden Physics Models: Deep Learning of Nonlinear Partial Differential Equations. *J. Mach. Learn. Res.* 19, 25 (2018), 1–24. issn: 1532-4435. url: http://jmlr.org/papers/v19/18-046.html.

[93] Rasulov, A., Raimova, G., and Mascagni, M. Monte Carlo solution of Cauchy problem for a nonlinear parabolic equation. *Math. Comput. Simulation* 80, 6 (2010), 1118–1123. issn: 0378-4754. doi: 10.1016/j.matcom.2009.12.009. url: https://doi.org/10.1016/j.matcom.2009.12.009.

[94] Reisinger, C., and Zhang, Y. Rectified deep neural networks overcome the curse of dimensionality for nonsmooth value functions in zero-sum games of nonlinear stiff systems. *ArXiv e-prints* (2019), 34 pages. arXiv: 1903.06652 [math.NA].

[95] Ruijter, M. J., and Oosterlee, C. W. A Fourier cosine method for an efficient computation of solutions to BSDEs. *SIAM J. Sci. Comput.* 37, 2 (2015), A859–A889. issn: 1064-8275. doi: 10.1137/130913183. url: https://doi.org/10.1137/130913183.

[96] Ruijter, M. J., and Oosterlee, C. W. Numerical Fourier method and second-order Taylor scheme for backward SDEs in finance. *Appl. Numer. Math.* 103 (2016), 1–26. issn: 0168-9274. doi: 10.1016/j.apnum.2015.12.003. url: https://doi.org/10.1016/j.apnum.2015.12.003.

[97] Ruszczynski, A., and Yao, J. A Dual Method For Backward Stochastic Differential Equations with Application to Risk Valuation. *ArXiv e-prints* (2017), 25 pages. arXiv: 1701.06234 [math.OC].

[98] Sirignano, J., and Spiliopoulos, K. DGM: A deep learning algorithm for solving partial differential equations. *J. Comput. Phys.* 375 (2018), 1339–1364. issn: 0021-9991. doi: 10.1016/j.jcp.2018.08.029. url: https://doi.org/10.1016/j.jcp.2018.08.029.

[99] Skorohod, A. V. Branching diffusion processes. *Teor. Verojatnost. i Primenen.* 9 (1964), 492–497. issn: 0040-361x.

[100] Turkedjiev, P. Two algorithms for the discrete time approximation of Markovian backward stochastic differential equations under local conditions. *Electron. J. Probab.* 20 (2015), no. 50, 49. doi: 10.1214/EJP.v20-3022. url: https://doi.org/10.1214/EJP.v20-3022.

[101] Warin, X. Variations on branching methods for non linear PDEs. *ArXiv e-prints* (2017), 25 pages. arXiv: 1701.07660 [math.PR].

[102] Watanabe, S. On the branching process for Brownian particles with an absorbing boundary. *J. Math. Kyoto Univ.* 4 (1965), 385–398. issn: 0023-608X. doi: 10.1215/kjm/1250524667. url: https://doi.org/10.1215/kjm/1250524667.

[103] Zhang, J. A numerical scheme for BSDEs. *Ann. Appl. Probab.* 14, 1 (2004), 459–488. issn: 1050-5164. doi: 10.1214/aoap/1075828058. url: https://doi.org/10.1214/aoap/1075828058.