Journal of Computational Mathematics
Vol.xx, No.x, 2024, 1–58.
http://www.global-sci.org/jcm
doi:10.4208/jcm.2308-m2021-0266

SPACE-TIME DEEP NEURAL NETWORK APPROXIMATIONS FOR HIGH-DIMENSIONAL PARTIAL DIFFERENTIAL EQUATIONS*

Fabian Hornung
Faculty of Mathematics, Karlsruhe Institute of Technology, Germany
Email: fabianhornung89@gmail.com
Arnulf Jentzen
School of Data Science and Shenzhen Research Institute of Big Data,
The Chinese University of Hong Kong, Shenzhen (CUHK-Shenzhen), China;
Applied Mathematics: Institute for Analysis and Numerics,
Faculty of Mathematics and Computer Science, University of Münster, Germany;
Seminar for Applied Mathematics, Department of Mathematics, ETH Zurich, Switzerland
Emails: ajentzen@cuhk.edu.cn, ajentzen@uni-muenster.de
Diyora Salimova[1)]
Seminar for Applied Mathematics, Department of Mathematics, ETH Zurich, Switzerland;
Department of Information Technology and Electrical Engineering, ETH Zurich, Switzerland;
Department for Applied Mathematics, University of Freiburg, Germany
Email: diyora.salimova@mathematik.uni-freiburg.de

Abstract

It is one of the most challenging issues in applied mathematics to approximately solve high-dimensional partial differential equations (PDEs) and most of the numerical approximation methods for PDEs in the scientific literature suffer from the so-called curse of dimensionality in the sense that the number of computational operations employed in the corresponding approximation scheme to obtain an approximation precision $\varepsilon > 0$ grows exponentially in the PDE dimension and/or the reciprocal of ε. Recently, certain deep learning based methods for PDEs have been proposed and various numerical simulations for such methods suggest that deep artificial neural network (ANN) approximations might have the capacity to indeed overcome the curse of dimensionality in the sense that the number of real parameters used to describe the approximating deep ANNs grows at most polynomially in both the PDE dimension $d \in \mathbb{N}$ and the reciprocal of the prescribed approximation accuracy $\varepsilon > 0$. There are now also a few rigorous mathematical results in the scientific literature which substantiate this conjecture by proving that deep ANNs overcome the curse of dimensionality in approximating solutions of PDEs. Each of these results establishes that deep ANNs overcome the curse of dimensionality in approximating suitable PDE solutions at a fixed time point $T > 0$ and on a compact cube $[a, b]^d$ in space but none of these results provides an answer to the question whether the entire PDE solution on $[0, T] \times [a, b]^d$ can be approximated by deep ANNs without the curse of dimensionality. It is precisely the subject of this article to overcome this issue. More specifically, the main result of this work in particular proves for every $a \in \mathbb{R}$, $b \in (a, \infty)$ that solutions of certain Kolmogorov PDEs can be approximated by deep ANNs on the space-time region $[0, T] \times [a, b]^d$ without the curse of dimensionality.

* Received September 21, 2021 / Revised version received xxx / Accepted August 23, 2023 / Published online June 12, 2024 /
1) Corresponding author

Mathematics subject classification: 65M75, 68T05, 65C30.
Key words: Deep artificial neural network, Curse of dimensionality, Approximation, Partial differential equation, PDE, Stochastic differential equation, SDE, Monte Carlo Euler, Feynman-Kac formula, ANN.

1. Introduction

It is one of the most challenging issues in applied mathematics to approximately solve high-dimensional partial differential equations (PDEs) and most of the numerical approximation methods for PDEs in the scientific literature suffer from the so-called *curse of dimensionality* in the sense that the number of computational operations employed in the corresponding approximation scheme to obtain an approximation precision $\varepsilon > 0$ grows exponentially in the PDE dimension and/or the reciprocal of ε (for such concepts cf., e.g., Bellman [7], Novak & Ritter [52], Novak & Woźniakowski [53, Chapter 1] and Novak & Woźniakowski [54, Chapter 9] and for methods which do not suffer from the curse of dimensionality in the case of some special classes of nonlinear PDEs cf., e.g., [16, 17, 31, 33, 34, 36], [5, Section 4], [15, Sections 2 and 6], and the references therein).

Recently, certain artificial neural networks (ANNs) based approximation methods for PDEs have been proposed and various numerical simulations for such methods suggest (cf., e.g., [9,11, 13,14,19,21,22,24,29,30,32,35,38,40,41,46,47,49–51,55–57,59,60] and the references mentioned therein) that deep ANNs might have the capacity to indeed overcome the curse of dimensionality in the sense that the number of real parameters used to describe the approximating deep ANNs grows at most polynomially in both the PDE dimension $d \in \mathbb{N} = \{1, 2, \ldots\}$ and the reciprocal of the prescribed approximation accuracy $\varepsilon > 0$.

There are now also a few rigorous mathematical results in the scientific literature which substantiate this conjecture by proving that deep ANNs overcome the curse of dimensionality in approximating solutions of PDEs; cf., e.g., [20, 23, 25, 27, 39, 45, 58]. Each of the references mentioned in the previous sentence establishes that deep ANNs overcome the curse of dimensionality in approximating suitable PDE solutions at a fixed time point $T > 0$ and on a compact cube $[a, b]^d$ in space but none of the results in these references provides an answer to the question whether the entire PDE solution on $[0, T] \times [a, b]^d$ can be approximated by deep ANNs without the curse of dimensionality.

It is precisely the subject of this article to overcome this issue. More specifically, the main result of this work, Theorem 4.1 in Subsection 4.6 below, in particular proves for every $a \in \mathbb{R}$, $b \in (a, \infty)$ that solutions of certain Kolmogorov PDEs can be approximated by deep ANNs on the space-time region $[0, T] \times [a, b]^d$ without the curse of dimensionality. To illustrate the findings of this work in more details we now present in Theorem 1.1 below a special case of Theorem 4.1.

Theorem 1.1. *Let* $\mathbf{N} = \cup_{L\in\mathbb{N}}\cup_{l_0,l_1,\ldots,l_L\in\mathbb{N}}(\times_{k=1}^{L}(\mathbb{R}^{l_k\times l_{k-1}}\times\mathbb{R}^{l_k}))$*, let* $A\colon (\cup_{d\in\mathbb{N}}\mathbb{R}^d) \to (\cup_{d\in\mathbb{N}}\mathbb{R}^d)$ *satisfy for all* $d \in \mathbb{N}$*,* $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ *that*

$$A(x) = (\max\{x_1, 0\}, \ldots, \max\{x_d, 0\}), \tag{1.1}$$

let $\mathcal{P}\colon \mathbf{N} \to \mathbb{N}$ *and* $\mathcal{R}\colon \mathbf{N} \to (\cup_{k,l\in\mathbb{N}} C(\mathbb{R}^k, \mathbb{R}^l))$ *satisfy for all* $L \in \mathbb{N}$*,* $l_0, l_1, \ldots, l_L \in \mathbb{N}$*,* $\Phi = ((W_1, B_1), \ldots, (W_L, B_L)) \in (\times_{k=1}^{L}(\mathbb{R}^{l_k\times l_{k-1}} \times \mathbb{R}^{l_k}))$*,* $x_0 \in \mathbb{R}^{l_0}, x_1 \in \mathbb{R}^{l_1}, \ldots, x_L \in \mathbb{R}^{l_L}$ *with*

$\forall\, k \in \mathbb{N} \cap (0, L)\colon x_k = A(W_k x_{k-1} + B_k)$ that

$$\mathcal{P}(\Phi) = \sum_{k=1}^{L} l_k(l_{k-1}+1), \quad \mathcal{R}(\Phi) \in C(\mathbb{R}^{l_0}, \mathbb{R}^{l_L}), \quad (\mathcal{R}(\Phi))(x_0) = W_L x_{L-1} + B_L, \tag{1.2}$$

for every $d \in \mathbb{N}$ let $f_d\colon \mathbb{R}^d \to \mathbb{R}^d$ and $g_d\colon \mathbb{R}^d \to \mathbb{R}$ be functions, let $T, \kappa, p \in (0,\infty)$, assume for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ that there exist $\mathfrak{f}, \mathfrak{g} \in \mathbf{N}$ such that for all $x, y \in \mathbb{R}^d$ it holds that

$$\mathcal{R}(\mathfrak{f}) \in C(\mathbb{R}^d, \mathbb{R}^d), \quad \mathcal{R}(\mathfrak{g}) \in C(\mathbb{R}^d, \mathbb{R}), \quad \|f_d(x) - f_d(y)\| \le \kappa \|x - y\|, \tag{1.3}$$

$$\varepsilon |g_d(x)| + \|f_d(x) - (\mathcal{R}(\mathfrak{f}))(x)\| + |g_d(x) - (\mathcal{R}(\mathfrak{g}))(x)| \le \varepsilon \kappa d^\kappa (1 + \|x\|^\kappa), \tag{1.4}$$

$$|(\mathcal{R}(\mathfrak{g}))(x) - (\mathcal{R}(\mathfrak{g}))(y)| \le \kappa d^\kappa (1 + \|x\|^\kappa + \|y\|^\kappa) \|x - y\|, \tag{1.5}$$

$$\varepsilon^\kappa [\mathcal{P}(\mathfrak{f}) + \mathcal{P}(\mathfrak{g})] + \|(\mathcal{R}(\mathfrak{f}))(x)\| \le \kappa (d^\kappa + \|x\|), \tag{1.6}$$

and for every $d \in \mathbb{N}$ let $u_d \in C([0,T] \times \mathbb{R}^d, \mathbb{R})$ be an at most polynomially growing viscosity solution of

$$\left(\frac{\partial}{\partial t} u_d\right)(t,x) = (\Delta_x u_d)(t,x) + \left(\frac{\partial}{\partial x} u_d\right)(t,x)\, f_d(x) \tag{1.7}$$

with $u_d(0,x) = g_d(x)$ for $(t,x) \in (0,T) \times \mathbb{R}^d$. Then there exists $c \in \mathbb{R}$ such that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ there exists $\mathfrak{u} \in \mathbf{N}$ such that $\mathcal{R}(\mathfrak{u}) \in C(\mathbb{R}^{d+1}, \mathbb{R})$, $\mathcal{P}(\mathfrak{u}) \le c \varepsilon^{-c} d^c$, and

$$\left[\int_{[0,T] \times [0,1]^d} |u_d(y) - (\mathcal{R}(\mathfrak{u}))(y)|^p \, dy \right]^{1/p} \le \varepsilon. \tag{1.8}$$

Theorem 1.1 follows from Corollary 4.4 in Subsection 4.6 below. Corollary 4.4, in turn, is a consequence of Theorem 4.1 which is the main result of this article. In the following we add a few comments on some of the mathematical objects appearing in Theorem 1.1 above.

Note in Theorem 1.1 that $\|\cdot\| \colon (\cup_{d \in \mathbb{N}} \mathbb{R}^d) \to [0,\infty)$ is the function which satisfies for all $d \in \mathbb{N}$, $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ that

$$\|x\| = \left[\sum_{j=1}^{d} |x_j|^2 \right]^{1/2} \tag{1.9}$$

(standard norm, cf. Definition 2.1 below). In Theorem 1.1 we approximate the solution functions $u_d\colon [0,T] \times \mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$, of the PDEs in (1.7) by deep ANNs. We assume that the solution functions $u_d\colon [0,T] \times \mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$, are at most polynomially growing which means that for every $d \in \mathbb{N}$ there exists $q \in (0,\infty)$ such that for all $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that

$$|u_d(t,x)| \le q(1 + \|x\|^q). \tag{1.10}$$

This polynomial growth assumption ensures uniqueness of the solution functions $u_d\colon [0,T] \times \mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$, of the PDEs in (1.7).

The set $\mathbf{N}$ in Theorem 1.1 is a set of tuples of pairs of real matrices and real vectors and we think of $\mathbf{N}$ as the set of all ANNs (cf. Definition 4.1 below). Observe that Theorem 1.1 is an approximation result for ANNs with the rectifier activation function and the corresponding rectifier functions are described through the function $A\colon (\cup_{d \in \mathbb{N}} \mathbb{R}^d) \to (\cup_{d \in \mathbb{N}} \mathbb{R}^d)$ appearing in Theorem 1.1.

For every $\Phi \in \mathbf{N}$ the number $\mathcal{P}(\Phi) \in \mathbb{N}$ in Theorem 1.1 corresponds to the number of real parameters employed to describe the ANN Φ (cf. Definition 4.1 below). For every $\Phi \in \mathbf{N}$ the function

$$\mathcal{R}(\Phi) \in (\cup_{k,l \in \mathbb{N}} C(\mathbb{R}^k, \mathbb{R}^l)) \tag{1.11}$$

corresponds to the realization function associated to the ANN Φ (cf. Definition 4.3 below). We also refer to Fig. 1.1 for a graphical illustration of the architecture of the ANN Φ and its realization function $\mathcal{R}(\Phi)$. The functions $f_d\colon \mathbb{R}^d \to \mathbb{R}^d$, $d \in \mathbb{N}$, describe the drift coefficient functions and the functions $g_d\colon \mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$, describe the initial value functions of the PDEs whose solutions we intend to approximate in Theorem 1.1 (see (1.7) above).

The real number $T \in (0, \infty)$ denotes the time horizon of the PDEs whose solutions we intend to approximate. The real number $\kappa \in (0, \infty)$ is a constant which we employ to formulate the assumptions on the drift coefficient functions $f_d\colon \mathbb{R}^d \to \mathbb{R}^d$, $d \in \mathbb{N}$, and the initial value functions $g_d\colon \mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$, of the PDEs whose solutions we intend to approximate in Theorem 1.1 (see (1.3)-(1.5) above). The real number $p \in (0, \infty)$ is used to describe the way how we measure the error between the exact solutions of the PDEs in (1.7) and the corresponding deep ANN approximations in the sense that we measure the error in the strong L^p-sense (see (1.8) above).

We assume in Theorem 1.1 that the drift coefficient functions $f_d\colon \mathbb{R}^d \to \mathbb{R}^d$, $d \in \mathbb{N}$, and the initial value functions $g_d\colon \mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$, of the PDEs whose solutions we intend to approximate can be approximated by ANNs without the curse of dimensionality (see (1.3) and (1.4) above). We note that according to *the universal approximation type theorems* for every

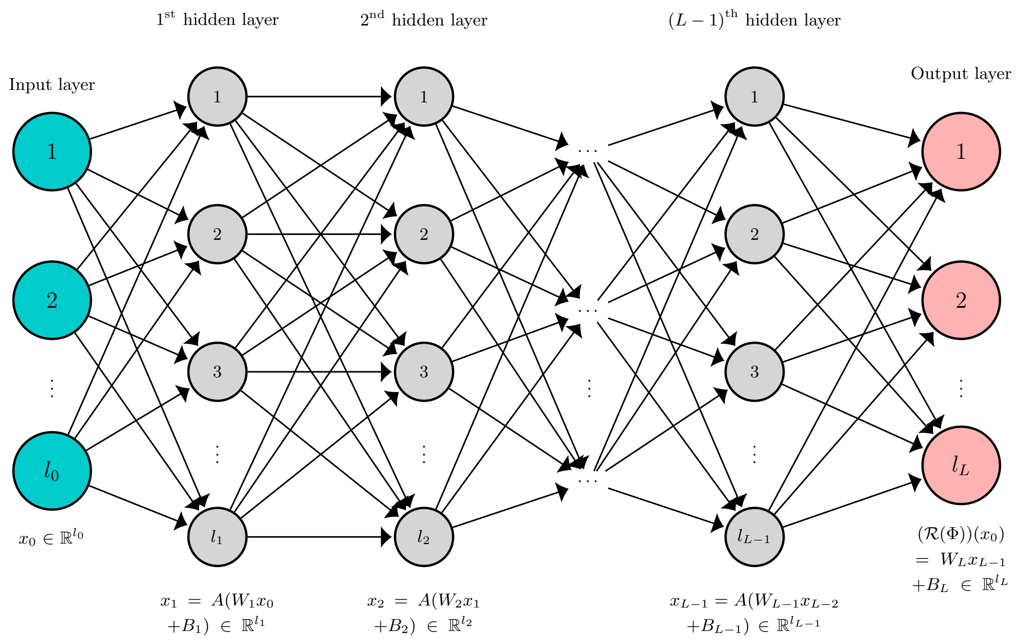

Fig. 1.1. Graphical illustration for the realization function and the architecture of an ANN $\Phi = ((W_1, B_1), \ldots, (W_L, B_L)) \in (\times_{k=1}^{L} (\mathbb{R}^{l_k \times l_{k-1}} \times \mathbb{R}^{l_k})) \subseteq \mathbf{N}$ (see Theorem 1.1) where $L \in \mathbb{N}$ describes the number of affine linear transformations, where $l_0, l_1, \ldots, l_L \in \mathbb{N}$ describe the dimensions of the layers of the ANN, and where the function $A\colon (\cup_{d \in \mathbb{N}} \mathbb{R}^d) \to (\cup_{d \in \mathbb{N}} \mathbb{R}^d)$ represents the activation function (see (1.1)).

$d \in \mathbb{N}$ and every compact set $K \subseteq \mathbb{R}^d$ one can uniformly approximate the functions $f_d \colon \mathbb{R}^d \to \mathbb{R}^d$ and $g_d \colon \mathbb{R}^d \to \mathbb{R}$ on the set K through ANNs with an arbitrary prescribed precision $\varepsilon > 0$ (see, e.g., Kidger & Lyons [42, Theorem 3.2]). However, the universal approximation type theorems do not guarantee that the number of parameters of the approximating ANNs grows at most polynomially in both the PDE dimension $d \in \mathbb{N}$ and the reciprocal of the prescribed approximation accuracy $\varepsilon > 0$, i.e., the universal approximation type theorems do not guarantee that the approximating ANNs do not suffer from the curse of dimensionality.

The functions $u_d \colon [0, T] \times \mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$, in Theorem 1.1 denote the PDE solutions which we intend to approximate by means of deep ANNs. In particular, in (1.8) in Theorem 1.1 we show that there exists a constant $c \in \mathbb{R}$ such that for any dimension $d \in \mathbb{N}$ and any approximation accuracy $\varepsilon \in (0, 1]$ there exists an ANN $\mathfrak{u} \in \mathbf{N}$ such that the number of parameters $\mathcal{P}(\mathfrak{u})$ of the ANN is bounded by $c\varepsilon^{-c}d^c$ and such that the realization function $\mathcal{R}(\mathfrak{u}) \colon \mathbb{R}^{d+1} \to \mathbb{R}$ of the ANN approximates the PDE solution $u_d \colon [0, T] \times \mathbb{R}^d \to \mathbb{R}$ in the $L^p([0, T] \times [0, 1]^d; \mathbb{R})$-sense with the precision ε.

Our proofs of Theorem 1.1 above and Theorem 4.1 below, respectively, are based on an application of Proposition 3.10 in Grohs *et al.* [26] (see (I)-(VI) in the proof of Proposition 4.1 in Subsection 4.4 below for details). More specifically, Proposition 3.10 in Grohs *et al.* [26] allows us to obtain space-time deep ANN approximations for Euler approximations of deterministic equations and Monte-Carlo Euler approximations of stochastic differential equations (SDEs) with desired complexity bounds. Combing this approximation result with the famous *Feynman-Kac theorem* (see Proposition 3.1 below for a special case) and the approximation error estimates for Monte Carlo Euler approximations in Proposition 3.2 below enables us construct space-time deep ANN approximations of certain Kolmogorov PDEs with desired approximation capabilities.

In the following we present concrete examples of PDEs whose coefficient functions satisfy the assumptions of Theorem 1.1 above. For further families of coefficient functions satisfying such kind of regularity assumptions we refer to the arguments, e.g., in Bach [1], E & Wang [18], Cheridito *et al.* [12], Beneventano *et al.* [8], and the references therein.

Example 1.1. For every $d \in \mathbb{N}$ let $f_d \colon \mathbb{R}^d \to \mathbb{R}^d$ and $g_d \colon \mathbb{R}^d \to \mathbb{R}$ satisfy for all $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ that

$$f_d(x) = \begin{cases} |x_1|, & \text{if} \quad d = 1, \\ (|x_d|, |x_1|, |x_2|, \ldots, |x_{d-1}|), & \text{if} \quad d > 1, \end{cases} \tag{1.12}$$

and

$$g_d(x) = \max\{|x_1|, |x_2|, \ldots, |x_d|\}. \tag{1.13}$$

Note that (1.12) and (1.13) ensure that for all $d \in \mathbb{N}$, $x, y \in \mathbb{R}^d$ it holds that

$$\|f_d(x) - f_d(y)\| \leq \|x - y\|, \quad |g_d(x)| \leq \|x\|, \quad |g_d(x) - g_d(y)| \leq \|x - y\| \tag{1.14}$$

(cf., e.g., Beneventano *et al.* [8, Lemma 5.3]). Proposition 3.1 in Section 3.1 below therefore implies that for every $d \in \mathbb{N}$ there exists a unique viscosity solution $u_d \in \{v \in C([0, T] \times \mathbb{R}^d, \mathbb{R}) \colon \inf_{q \in (0,\infty)} \sup_{(t,x) \in [0,T] \times \mathbb{R}^d} \frac{|v(t,x)|}{1 + \|x\|^q} < \infty\}$ of

$$\left(\frac{\partial}{\partial t} u_d\right)(t, x) = (\Delta_x u_d)(t, x) + \left(\frac{\partial}{\partial x} u_d\right)(t, x)\, f_d(x) \tag{1.15}$$

with $u_d(0,x) = g_d(x)$ for $(t,x) \in (0,T) \times \mathbb{R}^d$. Moreover, we observe that the families $(f_d)_{d\in\mathbb{N}}$ and $(g_d)_{d\in\mathbb{N}}$ can be exactly represented as realizations of ANNs with the number of parameters growing at most quadratically in the input dimension. More specifically, using the notation of Theorem 1.1 we note that there exists $\kappa \in (0,\infty)$ such that for all $d \in \mathbb{N}$ there exist $\mathfrak{f}, \mathfrak{g} \in \mathbf{N}$ such that

$$\mathcal{P}(\mathfrak{f}) + \mathcal{P}(\mathfrak{g}) \le \kappa d^2, \quad \mathcal{R}(\mathfrak{f}) = f_d, \quad \mathcal{R}(\mathfrak{g}) = g_d \tag{1.16}$$

(cf., e.g., Beneventano *et al.* [8, Proposition 5.4]). Theorem 1.1 hence proves that there exist $c \in \mathbb{R}$ such that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ there exists $\mathfrak{u} \in \mathbf{N}$ such that $\mathcal{R}(\mathfrak{u}) \in C(\mathbb{R}^{d+1},\mathbb{R})$, $\mathcal{P}(\mathfrak{u}) \le c\varepsilon^{-c}d^c$, and

$$\left[\int_{[0,T]\times[0,1]^d} |u_d(y) - (\mathcal{R}(\mathfrak{u}))(y)|^p \, dy\right]^{1/p} \le \varepsilon. \tag{1.17}$$

Theorem 1.1 is a purely theoretical result which asserts the existence of ANNs that can approximate the solutions of the PDEs without the curse of dimensionality. However, the proof of Theorem 1.1 and the earlier results, e.g., in [25,26,39] on which this work is partially based on, respectively, in some way suggest a concrete class of algorithms, specifically, a concrete class of ANN architectures with which the PDEs could be solved numerically and in this regard we refer to Becker *et al.* [6] for details and concrete numerical simulations. For standard feedforward fully connected ANN architectures used to approximatively solve Kolmogorov PDEs we refer, e.g., to Beck *et al.* [3] and Berner *et al.* [10].

The remainder of this article is structured in the following way. In Section 2 we establish in Proposition 2.1 and Lemma 2.3 suitable weak and strong error estimates for Euler-Maruyama approximations for a certain class of SDEs. In Section 3 we use these weak and strong error estimates for Euler-Maruyama approximations to establish in Proposition 3.2 below suitable error estimates for Monte Carlo Euler approximations for a class of SDEs with perturbed drift coefficient functions. In Section 4 we use these error estimates for Monte Carlo Euler approximations to establish in Theorem 4.1 below that for every $T \in (0,\infty)$, $a \in \mathbb{R}$, $b \in (a,\infty)$ it holds that solutions of certain Kolmogorov PDEs can be approximated by deep ANNs on the space-time region $[0,T] \times [a,b]^d$ without the curse of dimensionality.

2. Numerical Approximations for Stochastic Differential Equations (SDEs)

In this section we establish in Proposition 2.1 and Lemma 2.3 below suitable weak and strong error estimates for Euler-Maruyama approximations for a certain class of SDEs (see, e.g., Kloeden & Platen [44] for an extensive introduction to numerical approximations for SDEs). Our proofs of Proposition 2.1 and Lemma 2.3 are based on the elementary a priori moment estimates in Lemmas 2.1-2.2 below. Lemma 2.1 is, e.g., proved as Gonon *et al.* [23, Lemma 3.1] (see also, e.g., Jentzen *et al.* [39, Lemma 4.2]) and a slightly modified version of Lemma 2.2 is, e.g., proved as Gonon *et al.* [23, Lemma 3.4] (see also, e.g., Jentzen *et al.* [39, Lemma 4.1]).

2.1. A priori moment bounds for Gaussian random variables

Definition 2.1 (Standard Norms). *We denote by $\|\cdot\| : (\cup_{d\in\mathbb{N}}\mathbb{R}^d) \to [0,\infty)$ the function which satisfies for all $d \in \mathbb{N}$, $x = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$ that*

$$\|x\| = \left[\sum_{j=1}^{d} |x_j|^2 \right]^{1/2}. \tag{2.1}$$

Lemma 2.1. *Let $d \in \mathbb{N}$, $p \in (0,\infty)$, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $X\colon \Omega \to \mathbb{R}^d$ be a centered Gaussian random variable. Then*

$$\big(\mathbb{E}\big[\|X\|^p\big]\big)^{1/p} \leq \sqrt{\max\{1, p-1\}\operatorname{Trace}(\operatorname{Cov}(X))} \tag{2.2}$$

(cf. Definition 2.1).

2.2. A priori moment bounds for solutions of SDEs

Lemma 2.2. *Let $d \in \mathbb{N}$, $\xi \in \mathbb{R}^d$, $p \in [1,\infty)$, $c, C, T \in [0,\infty)$, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\mu\colon \mathbb{R}^d \to \mathbb{R}^d$ and $\chi\colon [0,T] \to [0,T]$ be measurable functions, assume for all $x \in \mathbb{R}^d$, $t \in [0,T]$ that $\|\mu(x)\| \leq C + c\|x\|$ and $\chi(t) \leq t$, and let $X, \beta\colon [0,T] \times \Omega \to \mathbb{R}^d$ be stochastic processes with continuous sample paths which satisfy for all $t \in [0,T]$ that*

$$\mathbb{P}\left(X_t = \xi + \int_0^t \mu\left(X_{\chi(s)}\right) ds + \beta_t \right) = 1 \tag{2.3}$$

(cf. Definition 2.1). Then

$$\sup_{t\in[0,T]} \big(\mathbb{E}\big[\|X_t\|^p\big]\big)^{1/p} \leq \Big(\|\xi\| + CT + \sup_{t\in[0,T]} \big(\mathbb{E}\big[\|\beta_t\|^p\big]\big)^{1/p} \Big) e^{cT}. \tag{2.4}$$

2.3. Weak error estimates for Euler-Maruyama approximations

Proposition 2.1. *Let $d, m \in \mathbb{N}$, $\xi \in \mathbb{R}^d$, $T \in (0,\infty)$, $c, C, \varepsilon_0, \varepsilon_1, \varepsilon_2, \varsigma_0, \varsigma_1, \varsigma_2, L_0, L_1, \ell \in [0,\infty)$, $h \in [0,T]$, $p \in [2,\infty)$, $q \in (1,2]$ satisfy $1/p + 1/q = 1$, let $B \in \mathbb{R}^{d\times m}$, $(\varpi_r)_{r\in(0,\infty)} \subseteq \mathbb{R}$ satisfy for all $r \in (0,\infty)$ that*

$$\varpi_r = \max\big\{1, \sqrt{\max\{1, r-1\}\operatorname{Trace}(B^*B)}\big\}, \tag{2.5}$$

let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $W\colon [0,T]\times\Omega \to \mathbb{R}^m$ be a standard Brownian motion, let $F_0\colon \mathbb{R}^d \to \mathbb{R}$, $f_1\colon \mathbb{R}^d \to \mathbb{R}^d$, $F_2\colon \mathbb{R}^d \to \mathbb{R}^d$, and $\chi\colon [0,T] \to [0,T]$ be functions, let $f_0\colon \mathbb{R}^d \to \mathbb{R}$ and $F_1\colon \mathbb{R}^d \to \mathbb{R}^d$ be measurable functions, assume for all $t \in [0,T]$, $x, y \in \mathbb{R}^d$ that

$$|f_0(x) - F_0(x)| \leq \varepsilon_0(1 + \|x\|^{\varsigma_0}), \quad \|f_1(x) - F_1(x)\| \leq \varepsilon_1(1 + \|x\|^{\varsigma_1}), \tag{2.6}$$

$$|F_0(x) - F_0(y)| \leq L_0 \left[1 + \int_0^1 \big[r\|x\| + (1-r)\|y\| \big]^{\ell} dr \right] \|x - y\|, \tag{2.7}$$

$$\|f_1(x) - f_1(y)\| \leq L_1\|x - y\|, \quad \|F_1(x)\| \leq C + c\|x\|, \quad \|\xi - F_2(\xi)\| \leq \varepsilon_2(1 + \|\xi\|^{\varsigma_2}), \tag{2.8}$$

and $\chi(t) = \max\left(\{0, h, 2h, \dots\} \cap [0,t]\right)$, and let $X, Y\colon [0,T]\times\Omega \to \mathbb{R}^d$ be stochastic processes with continuous sample paths which satisfy for all $t \in [0,T]$ that

$$X_t = \xi + \int_0^t f_1(X_s)\, ds + BW_t, \quad Y_t = F_2(\xi) + \int_0^t F_1\big(Y_{\chi(s)}\big)\, ds + BW_t \tag{2.9}$$

(cf. Definition 2.1). Then it holds for all $t \in [0,T]$ that

$$
\begin{aligned}
&\big|\mathbb{E}\big[f_0(X_t)\big] - \mathbb{E}\big[F_0(Y_t)\big]\big| \\
&\leq \Big(\varepsilon_2(1+\|\xi\|^{\varsigma_2}) + \varepsilon_0 + \varepsilon_1 + h + h^{1/2}\Big) \\
&\quad \times e^{[\max\{\varsigma_0,1\}L_1 + 1 - 1/p + \ell\max\{L_1,c\} + \max\{\varsigma_1,1\}c]T} (\varpi_{\max\{\varsigma_0,\ell q, p\varsigma_1, p\}})^{\max\{\varsigma_0,\ell+\max\{1,\varsigma_1\}\}} \\
&\quad \times (\max\{T,1\})^{\max\{\varsigma_0,\ell+\max\{\varsigma_1,1\}+1/p\}} \max\{L_0,1\} \max\{L_1,1\}\, 2^{\max\{\ell-1,0\}} \\
&\quad \times \Big[\max\{C,1\} + 5\max\{C,c,1\}\Big(\|\xi\| + \varepsilon_2(1+\|\xi\|^{\varsigma_2}) \\
&\qquad\qquad + 2\max\{\|f_1(0)\|, C, 1\}\Big)^{\max\{\varsigma_0,\ell+\max\{\varsigma_1,1\}\}}\Big]. \qquad (2.10)
\end{aligned}
$$

Proof. Observe that (2.5), Lemma 2.1, the fact that for all $t \in [0,T]$ it holds that BW_t is a centered Gaussian random variable, and the fact that for all $t \in [0,T]$ it holds that $\operatorname{Cov}(BW_t) = BB^* t$ assure that for all $r \in (0,\infty)$, $t \in [0,T]$ it holds that

$$
\begin{aligned}
\big(\mathbb{E}\big[\|BW_t\|^r\big]\big)^{1/r} &\leq \sqrt{\max\{1,r-1\}\operatorname{Trace}(\operatorname{Cov}(BW_t))} \\
&= \sqrt{\max\{1,r-1\}\operatorname{Trace}(BB^*)t} \\
&\leq \max\big\{t^{1/2}, \sqrt{\max\{1,r-1\}\operatorname{Trace}(B^*B)t}\big\} = \varpi_r t^{1/2}. \qquad (2.11)
\end{aligned}
$$

In addition, note that (2.8) shows that for all $x \in \mathbb{R}^d$ it holds that

$$
\|f_1(x)\| \leq \|f_1(x) - f_1(0)\| + \|f_1(0)\| \leq \|f_1(0)\| + L_1\|x\|. \qquad (2.12)
$$

Hölder's inequality, (2.8), Lemma 2.2, and (2.11) hence demonstrate that for all $r \in (0,\infty)$, $t \in [0,T]$ it holds that

$$
\begin{aligned}
\sup_{s\in[0,t]} \big(\mathbb{E}\big[\|Y_s\|^r\big]\big)^{1/r} &\leq \sup_{s\in[0,t]} \big(\mathbb{E}\big[\|Y_s\|^{\max\{r,1\}}\big]\big)^{1/\max\{r,1\}} \\
&\leq \Big[\|F_2(\xi)\| + Ct + \sup_{s\in[0,t]} \big(\mathbb{E}\big[\|BW_s\|^{\max\{r,1\}}\big]\big)^{1/\max\{r,1\}}\Big] e^{ct} \\
&\leq \Big[\|F_2(\xi)\| + Ct + \varpi_{\max\{r,1\}} t^{1/2}\Big] e^{ct} \qquad (2.13)
\end{aligned}
$$

and

$$
\begin{aligned}
\sup_{s\in[0,t]} \big(\mathbb{E}\big[\|X_s\|^r\big]\big)^{1/r} &\leq \sup_{s\in[0,t]} \big(\mathbb{E}\big[\|X_s\|^{\max\{r,1\}}\big]\big)^{1/\max\{r,1\}} \\
&\leq \Big[\|\xi\| + \|f_1(0)\|t + \sup_{s\in[0,t]} \big(\mathbb{E}\big[\|BW_s\|^{\max\{r,1\}}\big]\big)^{1/\max\{r,1\}}\Big] e^{L_1 t} \\
&\leq \Big[\|\xi\| + \|f_1(0)\|t + \varpi_{\max\{r,1\}} t^{1/2}\Big] e^{L_1 t}. \qquad (2.14)
\end{aligned}
$$

Hölder's inequality, (2.8), and the triangle inequality therefore imply that for all $r \in (0,\infty)$, $t \in [0,T]$ it holds that

$$
\begin{aligned}
\sup_{s\in[0,t]} \big(\mathbb{E}\big[\|F_1(Y_s)\|^r\big]\big)^{1/r} &\leq \sup_{s\in[0,t]} \big(\mathbb{E}\big[\|F_1(Y_s)\|^{\max\{r,1\}}\big]\big)^{1/\max\{r,1\}} \\
&\leq C + c \sup_{s\in[0,t]} \big(\mathbb{E}\big[\|Y_s\|^{\max\{r,1\}}\big]\big)^{1/\max\{r,1\}} \\
&\leq C + c\Big(\|F_2(\xi)\| + Ct + \varpi_{\max\{r,1\}} t^{1/2}\Big) e^{ct}. \qquad (2.15)
\end{aligned}
$$

Moreover, note that [39, Lemma 4.3] assures for all $t \in [0,T]$ that

$$\begin{aligned}
&\big|\mathbb{E}\big[f_0(X_t)\big] - \mathbb{E}\big[F_0(Y_t)\big]\big| \\
&\leq \varepsilon_0 \left(1 + \mathbb{E}\left[\|X_t\|^{\varsigma_0}\right]\right) + L_0\, 2^{\max\{\ell-1,0\}}\, e^{[L_1+1-1/p]t} \left[1 + \left(\mathbb{E}\big[\|X_t\|^{\ell q}\big]\right)^{1/q} + \left(\mathbb{E}\big[\|Y_t\|^{\ell q}\big]\right)^{1/q}\right] \\
&\quad \times \Bigg[\|\xi - F_2(\xi)\| + \varepsilon_1 t^{1/p}\Big[1 + \sup_{s\in[0,t]} \left(\mathbb{E}\big[\|Y_s\|^{p\varsigma_1}\big]\right)^{1/p}\Big] + h t^{1/p} L_1 \Big[\sup_{s\in[0,t]} \left(\mathbb{E}\big[\|F_1(Y_s)\|^{p}\big]\right)^{1/p}\Big] \\
&\qquad + t^{1/p} L_1 \left(\mathbb{E}\big[\|BW_h\|^p\big]\right)^{1/p}\Bigg]. \qquad (2.16)
\end{aligned}$$

This, (2.11), (2.13), (2.14), and (2.15) prove that for all $t \in [0,T]$ it holds that

$$\begin{aligned}
&\big|\mathbb{E}\big[f_0(X_t)\big] - \mathbb{E}\big[F_0(Y_t)\big]\big| \\
&\leq \varepsilon_0 \left[1 + \left(\|\xi\| + \|f_1(0)\| t + \varpi_{\max\{\varsigma_0,1\}} t^{1/2}\right)^{\varsigma_0} e^{\varsigma_0 L_1 t}\right] + L_0\, 2^{\max\{\ell-1,0\}}\, e^{[L_1+1-1/p]t} \\
&\quad \times \left[1 + \left(\|\xi\| + \|f_1(0)\| t + \varpi_{\max\{\ell q,1\}} t^{1/2}\right)^{\ell} e^{\ell L_1 t} + \left(\|F_2(\xi)\| + Ct + \varpi_{\max\{\ell q,1\}} t^{1/2}\right)^{\ell} e^{\ell c t}\right] \\
&\quad \times \Bigg[\|\xi - F_2(\xi)\| + \varepsilon_1 t^{1/p} \left[1 + \left(\|F_2(\xi)\| + Ct + \varpi_{\max\{p\varsigma_1,1\}} t^{1/2}\right)^{\varsigma_1} e^{\varsigma_1 c t}\right] \\
&\qquad + h t^{1/p} L_1 \left[C + c\left(\|F_2(\xi)\| + Ct + \varpi_p t^{1/2}\right) e^{ct}\right] + h^{1/2} t^{1/p} L_1 \varpi_p\Bigg]. \qquad (2.17)
\end{aligned}$$

In addition, observe that the fact that $(0,\infty) \ni r \mapsto \varpi_r \in (0,\infty)$ is non-decreasing and the hypothesis that $p \in [2,\infty)$ imply that

$$\varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}} \geq \max\{\varpi_{\max\{\varsigma_0,1\}}, \varpi_{\max\{\ell q,1\}}, \varpi_{\max\{p\varsigma_1,1\}}, \varpi_p\}. \qquad (2.18)$$

This and (2.17) ensure that for all $t \in [0,T]$ it holds that

$$\begin{aligned}
&\big|\mathbb{E}\big[f_0(X_t)\big] - \mathbb{E}\big[F_0(Y_t)\big]\big| \\
&\leq \varepsilon_0 e^{\varsigma_0 L_1 T} \left[1 + \left(\|\xi\| + \|f_1(0)\| T + \varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}} T^{1/2}\right)^{\varsigma_0}\right] \\
&\quad + L_0\, 2^{\max\{\ell-1,0\}}\, e^{[L_1+1-1/p+\ell\max\{L_1,c\}+\max\{\varsigma_1,1\}c]T} \\
&\quad \times \Bigg[1 + \left(\|\xi\| + \|f_1(0)\| T + \varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}} T^{1/2}\right)^{\ell} \\
&\qquad + \left(\|F_2(\xi)\| + CT + \varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}} T^{1/2}\right)^{\ell}\Bigg] \\
&\quad \times \Bigg[\|\xi - F_2(\xi)\| + \varepsilon_1 T^{1/p} \left[1 + \left(\|F_2(\xi)\| + CT + \varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}} T^{1/2}\right)^{\varsigma_1}\right] \\
&\qquad + h T^{1/p} L_1 \left[C + c\left(\|F_2(\xi)\| + CT + \varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}} T^{1/2}\right)\right] \\
&\qquad + h^{1/2} T^{1/p} L_1 \varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}}\Bigg]. \qquad (2.19)
\end{aligned}$$

Combining this with the fact that $\varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}} \geq 1$ and the fact that $T^{1/2} \leq (\max\{T,1\})^{1/2} \leq \max\{T,1\}$ assures that for all $t \in [0,T]$ it holds that

$$\begin{aligned}
&\big|\mathbb{E}\big[f_0(X_t)\big] - \mathbb{E}\big[F_0(Y_t)\big]\big| \\
&\leq \varepsilon_0 e^{\varsigma_0 L_1 T} (\varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}})^{\varsigma_0} \left[1 + \left(\|\xi\| + \|f_1(0)\| T + T^{1/2}\right)^{\varsigma_0}\right]
\end{aligned}$$

$$
\begin{aligned}
&+ L_0\, 2^{\max\{\ell-1,0\}}\, e^{[L_1+1-1/p+\ell\max\{L_1,c\}+\max\{\varsigma_1,1\}c]T} (\varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}})^{\ell+\max\{1,\varsigma_1\}} \\
&\times \left[1 + \left(\|\xi\| + \|f_1(0)\|T + T^{1/2}\right)^{\ell} + \left(\|F_2(\xi)\| + CT + T^{1/2}\right)^{\ell}\right] \\
&\times \Bigg[\|\xi - F_2(\xi)\| + \varepsilon_1 T^{1/p} \left[1 + \left(\|F_2(\xi)\| + CT + T^{1/2}\right)^{\varsigma_1}\right] \\
&\qquad + hT^{1/p} L_1 \left[C + c\left(\|F_2(\xi)\| + CT + T^{1/2}\right)\right] + h^{1/2} T^{1/p} L_1\Bigg]. \qquad (2.20)
\end{aligned}
$$

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

$$
\begin{aligned}
&\left|\mathbb{E}\big[f_0(X_t)\big] - \mathbb{E}\big[F_0(Y_t)\big]\right| \\
&\le \varepsilon_0 e^{\varsigma_0 L_1 T} (\varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}})^{\varsigma_0} \left(1 + (\max\{T,1\})^{\varsigma_0} \left(\|\xi\| + \|f_1(0)\| + 1\right)^{\varsigma_0}\right) \\
&\quad + L_0\, 2^{\max\{\ell-1,0\}}\, e^{[L_1+1-1/p+\ell\max\{L_1,c\}+\max\{\varsigma_1,1\}c]T} (\varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}})^{\ell+\max\{1,\varsigma_1\}} \\
&\quad \times \left[1 + (\max\{T,1\})^{\ell} \left[(\|\xi\| + \|f_1(0)\| + 1)^{\ell} + (\|F_2(\xi)\| + C + 1)^{\ell}\right]\right] \\
&\quad \times \Bigg[\|\xi - F_2(\xi)\| + \varepsilon_1 T^{1/p} \left[1 + (\max\{T,1\})^{\varsigma_1} (\|F_2(\xi)\| + C + 1)^{\varsigma_1}\right] \\
&\qquad + hT^{1/p} \max\{T,1\} L_1 \left[C + c\,(\|F_2(\xi)\| + C + 1)\right] + h^{1/2} T^{1/p} L_1\Bigg]. \qquad (2.21)
\end{aligned}
$$

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

$$
\begin{aligned}
&\left|\mathbb{E}\big[f_0(X_t)\big] - \mathbb{E}\big[F_0(Y_t)\big]\right| \\
&\le \varepsilon_0 e^{\varsigma_0 L_1 T} (\varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}})^{\varsigma_0} (\max\{T,1\})^{\varsigma_0} \left(1 + (\|\xi\| + \|f_1(0)\| + 1)^{\varsigma_0}\right) \\
&\quad + L_0\, 2^{\max\{\ell-1,0\}}\, e^{[L_1+1-1/p+\ell\max\{L_1,c\}+\max\{\varsigma_1,1\}c]T} (\varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}})^{\ell+\max\{1,\varsigma_1\}} \\
&\quad \times (\max\{T,1\})^{\ell+\max\{\varsigma_1,1\}+1/p} \left[1 + (\|\xi\| + \|f_1(0)\| + 1)^{\ell} + (\|F_2(\xi)\| + C + 1)^{\ell}\right] \\
&\quad \times \Bigg[\|\xi - F_2(\xi)\| + \varepsilon_1 \left[1 + (\|F_2(\xi)\| + C + 1)^{\varsigma_1}\right] \\
&\qquad + hL_1 \left[C + c\,(\|F_2(\xi)\| + C + 1)\right] + h^{1/2} L_1\Bigg]. \qquad (2.22)
\end{aligned}
$$

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

$$
\begin{aligned}
&\left|\mathbb{E}\big[f_0(X_t)\big] - \mathbb{E}\big[F_0(Y_t)\big]\right| \\
&\le \varepsilon_0 e^{\varsigma_0 L_1 T} (\varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}})^{\varsigma_0} (\max\{T,1\})^{\varsigma_0} \\
&\quad \times \left(1 + \left(\max\{\|F_2(\xi)\|, \|\xi\|\} + 2\max\{\|f_1(0)\|, C, 1\}\right)^{\varsigma_0}\right) \\
&\quad + L_0\, 2^{\max\{\ell-1,0\}}\, e^{[L_1+1-1/p+\ell\max\{L_1,c\}+\max\{\varsigma_1,1\}c]T} (\varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}})^{\ell+\max\{1,\varsigma_1\}} \\
&\quad \times (\max\{T,1\})^{\ell+\max\{\varsigma_1,1\}+1/p} \left[1 + 2\left(\max\{\|F_2(\xi)\|, \|\xi\|\} + 2\max\{\|f_1(0)\|, C, 1\}\right)^{\ell}\right] \\
&\quad \times \Bigg[\|\xi - F_2(\xi)\| + \varepsilon_1 \left[1 + (\max\{\|F_2(\xi)\|, \|\xi\|\} + 2\max\{\|f_1(0)\|, C, 1\})^{\varsigma_1}\right] \\
&\qquad + hL_1 \left[C + c\,(\max\{\|F_2(\xi)\|, \|\xi\|\} + 2\max\{\|f_1(0)\|, C, 1\})\right] + h^{1/2} L_1\Bigg]. \qquad (2.23)
\end{aligned}
$$

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

$$
\begin{aligned}
&\left|\mathbb{E}\big[f_0(X_t)\big] - \mathbb{E}\big[F_0(Y_t)\big]\right| \\
&\le \varepsilon_0 e^{\varsigma_0 L_1 T} (\varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}})^{\varsigma_0} (\max\{T,1\})^{\varsigma_0}
\end{aligned}
$$

$$\begin{aligned}
&\times \big(1 + \big(\max\{\|F_2(\xi)\|, \|\xi\|\} + 2 \max\{\|f_1(0)\|, C, 1\}\big)^{\varsigma_0}\big) \\
&+ L_0\, 2^{\max\{\ell-1,0\}}\, e^{[L_1+1-1/p+\ell \max\{L_1,c\}+\max\{\varsigma_1,1\}c]T} (\varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}})^{\ell+\max\{1,\varsigma_1\}} \\
&\times (\max\{T,1\})^{\ell+\max\{\varsigma_1,1\}+1/p} \Big[1 + 2\big(\max\{\|F_2(\xi)\|, \|\xi\|\} + 2 \max\{\|f_1(0)\|, C, 1\}\big)^{\ell}\Big] \\
&\times \Big(\|\xi - F_2(\xi)\| + \varepsilon_1 + h + h^{1/2}\Big) \max\{L_1, 1\} \\
&\times \Big[\max\{C,1\} + \max\{c,1\}\big(\max\{\|F_2(\xi)\|, \|\xi\|\} + 2 \max\{\|f_1(0)\|, C, 1\}\big)^{\max\{\varsigma_1,1\}}\Big]. \qquad (2.24)
\end{aligned}$$

Combining this with the fact that for all $a, b \in [0, \infty)$, $z \in [1, \infty)$ it holds that

$$\begin{aligned}
&\big(1 + 2z^{\ell}\big)\big(a + bz^{\max\{\varsigma_1,1\}}\big) \\
&= a + bz^{\max\{\varsigma_1,1\}} + 2az^{\ell} + 2bz^{\ell+\max\{\varsigma_1,1\}} \\
&\le a + (3b + 2a)z^{\ell+\max\{\varsigma_1,1\}} \\
&\le a + 5 \max\{a, b\} z^{\ell+\max\{\varsigma_1,1\}} \qquad (2.25)
\end{aligned}$$

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

$$\begin{aligned}
&\big|\mathbb{E}\big[f_0(X_t)\big] - \mathbb{E}\big[F_0(Y_t)\big]\big| \\
&\le \varepsilon_0 e^{\varsigma_0 L_1 T} (\varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}})^{\varsigma_0} (\max\{T,1\})^{\varsigma_0} \\
&\quad \times \big(1 + \big(\max\{\|F_2(\xi)\|, \|\xi\|\} + 2 \max\{\|f_1(0)\|, C, 1\}\big)^{\varsigma_0}\big) \\
&\quad + L_0\, 2^{\max\{\ell-1,0\}}\, e^{[L_1+1-1/p+\ell \max\{L_1,c\}+\max\{\varsigma_1,1\}c]T} (\varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}})^{\ell+\max\{1,\varsigma_1\}} \\
&\quad \times (\max\{T,1\})^{\ell+\max\{\varsigma_1,1\}+1/p} \Big(\|\xi - F_2(\xi)\| + \varepsilon_1 + h + h^{1/2}\Big) \max\{L_1, 1\} \\
&\quad \times \Bigg[\max\{C,1\} + 5 \max\{C, c, 1\} \Big(\max\{\|F_2(\xi)\|, \|\xi\|\} \\
&\qquad\qquad + 2 \max\{\|f_1(0)\|, C, 1\}\Big)^{\ell+\max\{\varsigma_1,1\}}\Bigg]. \qquad (2.26)
\end{aligned}$$

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

$$\begin{aligned}
&\big|\mathbb{E}\big[f_0(X_t)\big] - \mathbb{E}\big[F_0(Y_t)\big]\big| \\
&\le \Big(\|\xi - F_2(\xi)\| + \varepsilon_0 + \varepsilon_1 + h + h^{1/2}\Big) \\
&\quad \times e^{[\max\{\varsigma_0,1\}L_1+1-1/p+\ell \max\{L_1,c\}+\max\{\varsigma_1,1\}c]T} (\varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}})^{\max\{\varsigma_0,\ell+\max\{1,\varsigma_1\}\}} \\
&\quad \times (\max\{T,1\})^{\max\{\varsigma_0,\ell+\max\{\varsigma_1,1\}+1/p\}} \max\{L_0, 1\} \max\{L_1, 1\}\, 2^{\max\{\ell-1,0\}} \\
&\quad \times \Bigg[\max\{C,1\} + 5 \max\{C, c, 1\} \Big(\max\{\|F_2(\xi)\|, \|\xi\|\} \\
&\qquad\qquad + 2 \max\{\|f_1(0)\|, C, 1\}\Big)^{\max\{\varsigma_0,\ell+\max\{\varsigma_1,1\}\}}\Bigg]. \qquad (2.27)
\end{aligned}$$

The hypothesis that $\|F_2(\xi) - \xi\| \le \varepsilon_2(1 + \|\xi\|^{\varsigma_2})$ and the fact that

$$\max\{\|\xi\|, \|F_2(\xi)\|\} \le \|\xi\| + \|F_2(\xi) - \xi\| \le \|\xi\| + \varepsilon_2(1 + \|\xi\|^{\varsigma_2}) \qquad (2.28)$$

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

$$\begin{aligned}
&\big|\mathbb{E}\big[f_0(X_t)\big] - \mathbb{E}\big[F_0(Y_t)\big]\big| \\
&\le \Big(\varepsilon_2(1 + \|\xi\|^{\varsigma_2}) + \varepsilon_0 + \varepsilon_1 + h + h^{1/2}\Big)
\end{aligned}$$

$$\times\, e^{[\max\{\varsigma_0,1\}L_1+1-1/p+\ell\max\{L_1,c\}+\max\{\varsigma_1,1\}c]T}\big(\varpi_{\max\{\varsigma_0,\ell q,p\varsigma_1,p\}}\big)^{\max\{\varsigma_0,\ell+\max\{1,\varsigma_1\}\}}$$
$$\times\,(\max\{T,1\})^{\max\{\varsigma_0,\ell+\max\{\varsigma_1,1\}+1/p\}}\max\{L_0,1\}\max\{L_1,1\}\,2^{\max\{\ell-1,0\}}$$
$$\times\Bigg[\max\{C,1\}+5\max\{C,c,1\}\Big(\|\xi\|+\varepsilon_2(1+\|\xi\|^{\varsigma_2})$$
$$+\,2\max\{\|f_1(0)\|,C,1\}\Big)^{\max\{\varsigma_0,\ell+\max\{\varsigma_1,1\}\}}\Bigg]. \tag{2.29}$$

This completes the proof of Proposition 2.1. □

2.4. Strong error estimates for linearly interpolated Euler-Maruyama approximations

Lemma 2.3. *Let $d,m,N\in\mathbb{N}$, $T,p\in(0,\infty)$, $C,c\in[0,\infty)$, $q\in[1,\infty)$, $x\in\mathbb{R}^d$, $B\in\mathbb{R}^{d\times m}$, $\tau_0,\tau_1,\ldots,\tau_N\in[0,T]$ satisfy that $0=\tau_0<\tau_1<\cdots<\tau_{N-1}<\tau_N=T$, let $\mu\colon\mathbb{R}^d\to\mathbb{R}^d$ be a measurable function, assume for all $x\in\mathbb{R}^d$ that $\|\mu(x)\|\leq C+c\|x\|$, let $\lfloor\cdot\rfloor\colon[0,T]\to[0,T]$ satisfy for all $t\in[0,T]$ that $\lfloor t\rfloor=\max(\{\tau_0,\tau_1,\ldots,\tau_N\}\cap[0,t])$, let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $W\colon[0,T]\times\Omega\to\mathbb{R}^m$ be a standard Brownian motion, let $\mathcal{Y}\colon[0,T]\times\Omega\to\mathbb{R}^d$ satisfy for all $t\in[0,T]$ that*

$$\mathcal{Y}_t=x+\int_0^t\mu(\mathcal{Y}_{\lfloor s\rfloor})\,ds+BW_t, \tag{2.30}$$

and let $Y\colon[0,T]\times\Omega\to\mathbb{R}^d$ satisfy for all $n\in\{0,1,\ldots,N-1\}$, $t\in[\tau_n,\tau_{n+1}]$ that $Y_0=x$ and

$$Y_t=Y_{\tau_n}+\frac{t-\tau_n}{\tau_{n+1}-\tau_n}\big[\mu(Y_{\tau_n})(\tau_{n+1}-\tau_n)+B(W_{\tau_{n+1}}-W_{\tau_n})\big] \tag{2.31}$$

(cf. Definition 2.1). Then

(i) it holds that $\mathcal{Y}$ and Y are stochastic processes,

(ii) it holds for all $n\in\{0,1,\ldots,N\}$ that $Y_{\tau_n}=\mathcal{Y}_{\tau_n}$,

(iii) it holds for all $n\in\{0,1,\ldots,N-1\}$, $t\in(\tau_n,\tau_{n+1})$ that

$$\big(\mathbb{E}\big[\|Y_t-\mathcal{Y}_t\|^p\big]\big)^{1/p}\leq\frac{1}{2}\sqrt{\max\{1,p-1\}(\tau_{n+1}-\tau_n)\operatorname{Trace}(BB^*)}, \tag{2.32}$$

(iv) it holds for all $t\in[0,T]$ that

$$\max\Big\{\big(\mathbb{E}\big[\|\mathcal{Y}_t\|^q\big]\big)^{1/q},\big(\mathbb{E}\big[\|Y_t\|^q\big]\big)^{1/q}\Big\}$$
$$\leq\Big[\|x\|+CT+\sqrt{\max\{1,q-1\}T\operatorname{Trace}(BB^*)}\Big]e^{cT}. \tag{2.33}$$

Proof. Throughout this proof for every $\mathfrak{d}\in\mathbb{N}$ let $\mathrm{I}_\mathfrak{d}\in\mathbb{R}^{\mathfrak{d}\times\mathfrak{d}}$ be the identity matrix in $\mathbb{R}^{\mathfrak{d}\times\mathfrak{d}}$, let $\lceil\cdot\rceil\colon[0,T]\to[0,T]$ satisfy for all $t\in[0,T]$ that $\lceil t\rceil=\min(\{\tau_0,\tau_1,\ldots,\tau_N\}\cap[t,T])$, and let $\rho\colon[0,T]\to[0,1]$ satisfy for all $t\in[0,T]$ that

$$\rho(t)=\begin{cases}\dfrac{t-\lfloor t\rfloor}{\lceil t\rceil-\lfloor t\rfloor}, & \text{if}\quad t\notin\{\tau_0,\tau_1,\ldots,\tau_N\},\\ 0, & \text{if}\quad t\in\{\tau_0,\tau_1,\ldots,\tau_N\}.\end{cases} \tag{2.34}$$

Observe that (2.30), the fact that for all $t \in [0,T]$ it holds that $\Omega \ni \omega \mapsto W_t(\omega) \in \mathbb{R}^m$ is measurable, and induction imply that for all $t \in [0,T]$ it holds that $\Omega \ni \omega \mapsto \mathcal{Y}_t(\omega) \in \mathbb{R}^d$ is measurable. Moreover, note that (2.31), the fact that for all $t \in [0,T]$ it holds that $\Omega \ni \omega \mapsto W_t(\omega) \in \mathbb{R}^m$ is measurable, and induction prove that for all $t \in [0,T]$ it holds that $\Omega \ni \omega \mapsto Y_t(\omega) \in \mathbb{R}^d$ is measurable. Combining this with the fact that for all $t \in [0,T]$ it holds that $\Omega \ni \omega \mapsto \mathcal{Y}_t(\omega) \in \mathbb{R}^d$ is measurable establishes item (i). Next we claim that for all $n \in \{0,1,\ldots,N\}$, $t \in [\tau_{\max\{n-1,0\}}, \tau_n]$ it holds that $Y_{\tau_n} = \mathcal{Y}_{\tau_n}$ and

$$Y_t = x + \int_0^t \mu(Y_{\lfloor s \rfloor})\,ds + BW_{\lfloor t \rfloor} + \rho(t)B(W_{\lceil t \rceil} - W_{\lfloor t \rfloor}). \tag{2.35}$$

We prove (2.35) by induction on $n \in \{0,1,\ldots,N\}$. Note that the fact that $Y_{\tau_0} = x$, the fact that $\rho(\tau_0) = 0$, and the fact that $W_{\tau_0} = 0$ demonstrate that

$$Y_{\tau_0} = x = x + \int_0^{\tau_0} \mu(Y_{\lfloor s \rfloor})\,ds + BW_{\lfloor \tau_0 \rfloor} + \rho(\tau_0)B(W_{\lceil \tau_0 \rceil} - W_{\lfloor \tau_0 \rfloor}). \tag{2.36}$$

This and the fact that $\mathcal{Y}_{\tau_0} = x$ prove (2.35) in the base case $n = 0$. For the induction step $\{0,1,\ldots,N-1\} \ni n \to n+1 \in \{1,2,\ldots,N\}$ assume that there exists $n \in \{0,1,\ldots,N-1\}$ which satisfies that for all $m \in \{0,1,\ldots,n\}$, $t \in [\tau_{\max\{n-1,0\}}, \tau_n]$ it holds that $Y_{\tau_m} = \mathcal{Y}_{\tau_m}$ and

$$Y_t = x + \int_0^t \mu(Y_{\lfloor s \rfloor})\,ds + BW_{\lfloor t \rfloor} + \rho(t)B(W_{\lceil t \rceil} - W_{\lfloor t \rfloor}). \tag{2.37}$$

Note that (2.30) and (2.37) imply that

$$Y_{\tau_n} = \mathcal{Y}_{\tau_n} = x + \int_0^{\tau_n} \mu(\mathcal{Y}_{\lfloor s \rfloor})\,ds + BW_{\tau_n} = x + \int_0^{\tau_n} \mu(Y_{\lfloor s \rfloor})\,ds + BW_{\tau_n}. \tag{2.38}$$

Combining this with (2.31) ensures that for all $t \in [\tau_n, \tau_{n+1}] = [\tau_{\max\{n,0\}}, \tau_{n+1}]$ it holds that

$$\begin{aligned} Y_t &= Y_{\tau_n} + (t-\tau_n)\mu(Y_{\tau_n}) + \frac{t-\tau_n}{\tau_{n+1}-\tau_n} B(W_{\tau_{n+1}} - W_{\tau_n}) \\ &= x + \int_0^{\tau_n} \mu(Y_{\lfloor s \rfloor})\,ds + BW_{\tau_n} + (t-\tau_n)\mu(Y_{\tau_n}) + \frac{t-\tau_n}{\tau_{n+1}-\tau_n} B(W_{\tau_{n+1}} - W_{\tau_n}) \\ &= x + \int_0^t \mu(Y_{\lfloor s \rfloor})\,ds + BW_{\tau_n} + \frac{t-\tau_n}{\tau_{n+1}-\tau_n} B(W_{\tau_{n+1}} - W_{\tau_n}). \end{aligned} \tag{2.39}$$

Therefore, we obtain that for all $t \in [\tau_{\max\{n,0\}}, \tau_{n+1}]$ it holds that

$$Y_t = x + \int_0^t \mu(Y_{\lfloor s \rfloor})\,ds + BW_{\lfloor t \rfloor} + \rho(t)B(W_{\lceil t \rceil} - W_{\lfloor t \rfloor}). \tag{2.40}$$

This, (2.37), and (2.30) assure that

$$\begin{aligned} Y_{\tau_{n+1}} &= x + \int_0^{\tau_{n+1}} \mu(Y_{\lfloor s \rfloor})\,ds + BW_{\tau_{n+1}} \\ &= x + \int_0^{\tau_{n+1}} \mu(\mathcal{Y}_{\lfloor s \rfloor})\,ds + BW_{\tau_{n+1}} = \mathcal{Y}_{\tau_{n+1}}. \end{aligned} \tag{2.41}$$

Combining this with (2.40) implies that for all $t \in [\tau_{\max\{n,0\}}, \tau_{n+1}]$ it holds that $Y_{\tau_{n+1}} = \mathcal{Y}_{\tau_{n+1}}$ and

$$Y_t = x + \int_0^t \mu(Y_{\lfloor s \rfloor})\,ds + BW_{\lfloor t \rfloor} + \rho(t)B(W_{\lceil t \rceil} - W_{\lfloor t \rfloor}). \tag{2.42}$$

Induction thus proves (2.35). Next observe that (2.35) establishes item (ii). Moreover, note that (2.30), (2.31), (2.34), and (2.35) demonstrate that for all $n \in \{0,1,\dots,N-1\}$, $t \in (\tau_n,\tau_{n+1})$ it holds that

$$\begin{aligned} Y_t - \mathcal{Y}_t &= Y_{\tau_n} + \rho(t)\left[(\tau_{n+1}-\tau_n)\mu(Y_{\tau_n}) + B(W_{\tau_{n+1}} - W_{\tau_n})\right] \\ &\quad - \left[x + \int_0^t \mu(\mathcal{Y}_{\lfloor s \rfloor})\,ds + BW_t\right] \\ &= Y_{\tau_n} + (t-\tau_n)\mu(Y_{\tau_n}) + \rho(t)B(W_{\tau_{n+1}} - W_{\tau_n}) \\ &\quad - \left[\mathcal{Y}_{\tau_n} + \int_{\tau_n}^t \mu(\mathcal{Y}_{\lfloor s \rfloor})\,ds + B(W_t - W_{\tau_n})\right] \\ &= (t-\tau_n)\mu(Y_{\tau_n}) - (t-\tau_n)\mu(\mathcal{Y}_{\tau_n}) + \rho(t)B(W_{\tau_{n+1}} - W_{\tau_n}) - B(W_t - W_{\tau_n}). \end{aligned} \tag{2.43}$$

This and (2.35) prove that for all $n \in \{0,1,\dots,N-1\}$, $t \in (\tau_n,\tau_{n+1})$ it holds that

$$\begin{aligned} Y_t - \mathcal{Y}_t &= \rho(t)B\big(W_{\tau_{n+1}} - W_{\tau_n}\big) + BW_{\tau_n} - BW_t \\ &= -[\rho(t)-1]BW_{\tau_n} + \big[(\rho(t)-1) - \rho(t)\big]BW_t + \rho(t)BW_{\tau_{n+1}} \\ &= [\rho(t)-1]B\big(W_t - W_{\tau_n}\big) + \rho(t)B\big(W_{\tau_{n+1}} - W_t\big). \end{aligned} \tag{2.44}$$

In addition, note that the hypothesis that $W\colon [0,T]\times\Omega \to \mathbb{R}^m$ is a standard Brownian motion ensures that

(A) it holds for all $a, \mathfrak{a} \in \mathbb{R}$, $r,s,t \in [0,T]$ with $r \le s \le t$ that $aB(W_t - W_s) + \mathfrak{a}B(W_s - W_r)$ is a centered Gaussian random variable,

(B) it holds for all $a, \mathfrak{a} \in \mathbb{R}$, $r,s,t \in [0,T]$ with $r \le s \le t$ that

$$\operatorname{Cov}\big(aB(W_t - W_s) + \mathfrak{a}B(W_s - W_r)\big) = \big[a^2(t-s) + \mathfrak{a}^2(s-r)\big]\,BB^*. \tag{2.45}$$

Combining this with (2.44) ensures that for all $n \in \{0,1,\dots,N-1\}$, $t \in (\tau_n,\tau_{n+1})$ it holds that $Y_t - \mathcal{Y}_t$ is a centered Gaussian random variable. Moreover, note that (2.44) and (2.45) demonstrate that for all $n \in \{0,1,\dots,N-1\}$, $t \in (\tau_n,\tau_{n+1})$ it holds that

$$\operatorname{Cov}\big(Y_t - \mathcal{Y}_t\big) = \big([\rho(t)-1]^2(t-\tau_n) + [\rho(t)]^2(\tau_{n+1}-t)\big)\,BB^*. \tag{2.46}$$

In addition, observe that (2.34) implies that for all $n \in \{0,1,\dots,N-1\}$, $t \in (\tau_n,\tau_{n+1})$ it holds that

$$\begin{aligned} &[\rho(t)-1]^2(t-\tau_n) + [\rho(t)]^2(\tau_{n+1}-t) \\ &= [\rho(t)]^2(\tau_{n+1}-\tau_n) + \big[1-2\rho(t)\big](t-\tau_n) \\ &= \frac{(t-\tau_n)^2}{(\tau_{n+1}-\tau_n)} + (t-\tau_n) - \frac{2(t-\tau_n)^2}{(\tau_{n+1}-\tau_n)} \\ &= (t-\tau_n)\left[1 - \frac{(t-\tau_n)}{(\tau_{n+1}-\tau_n)}\right] \\ &= \frac{(t-\tau_n)(\tau_{n+1}-t)}{(\tau_{n+1}-\tau_n)}. \end{aligned} \tag{2.47}$$

This and the fact that for all $a \in \mathbb{R}$, $b \in (a,\infty)$, $r \in [a,b]$ it holds that

$$(r-a)(b-r) \le \left(\frac{1}{2}(b+a) - a\right)\left(b - \frac{1}{2}(b+a)\right) = \frac{1}{4}(b-a)^2 \tag{2.48}$$

show that for all $n \in \{0, 1, \dots, N-1\}$, $t \in (\tau_n, \tau_{n+1})$ it holds that

$$[\rho(t) - 1]^2 (t - \tau_n) + [\rho(t)]^2 (\tau_{n+1} - t) \leq \frac{1}{4}\big(\tau_{n+1} - \tau_n\big). \tag{2.49}$$

The fact that BB^* is a symmetric positive semidefinite matrix and (2.46) therefore imply that for all $t \in [0, T]$ it holds that

$$\operatorname{Trace}\big(\operatorname{Cov}(Y_t - \mathcal{Y}_t)\big) \leq \frac{1}{4}\big(\tau_{n+1} - \tau_n\big) \operatorname{Trace}(BB^*). \tag{2.50}$$

Lemma 2.1 hence demonstrates that for all $n \in \{0, 1, \dots, N-1\}$, $t \in (\tau_n, \tau_{n+1})$ it holds that

$$\begin{aligned} \big(\mathbb{E}\big[\|Y_t - \mathcal{Y}_t\|^p\big]\big)^{1/p} &\leq \sqrt{\max\{1, p-1\} \operatorname{Trace}\big(\operatorname{Cov}(Y_t - \mathcal{Y}_t)\big)} \\ &\leq \frac{1}{2}\sqrt{\max\{1, p-1\}\big(\tau_{n+1} - \tau_n\big) \operatorname{Trace}(BB^*)}. \end{aligned} \tag{2.51}$$

This establishes item (iii). Next note that Lemmas 2.1 and 2.2, the fact that for all $t \in [0, T]$ it holds that BW_t is a centered Gaussian random variable, and the fact that for all $t \in [0, T]$ it holds that $\operatorname{Cov}(BW_t) = BB^* \, t$ ensure that for all $t \in [0, T]$ it holds that

$$\begin{aligned} \big(\mathbb{E}\left[\|\mathcal{Y}_t\|^q\right]\big)^{1/q} &\leq \left[\|x\| + CT + \sup_{t \in [0,T]} \big(\mathbb{E}\big[\|BW_t\|^q\big]\big)^{1/q}\right] e^{cT} \\ &\leq \left[\|x\| + CT + \sup_{t \in [0,T]} \sqrt{\max\{1, q-1\} \operatorname{Trace}(BB^*)t}\right] e^{cT} \\ &= \left[\|x\| + CT + \sqrt{\max\{1, q-1\} \operatorname{Trace}(BB^*)T}\right] e^{cT}. \end{aligned} \tag{2.52}$$

Next note that (2.45), the fact that $W_0 = 0$, the fact that BB^* is a symmetric positive semidefinite matrix, and the fact that $\forall \, t \in [0, T] \colon 0 \leq \rho(t) \leq 1$ imply that

a) it holds for all $t \in [0, T]$ that $BW_{\lfloor t \rfloor} + \rho(t)B\big(W_{\lceil t \rceil} - W_{\lfloor t \rfloor}\big)$ is a centered Gaussian random variable,

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

$$\begin{aligned} &\operatorname{Trace}\big(\operatorname{Cov}\big(BW_{\lfloor t \rfloor} + \rho(t)B\big(W_{\lceil t \rceil} - W_{\lfloor t \rfloor}\big)\big)\big) \\ &= \operatorname{Trace}(BB^*) \left[\lfloor t \rfloor + [\rho(t)]^2 \big(\lceil t \rceil - \lfloor t \rfloor\big)\right] \leq \operatorname{Trace}(BB^*) \lceil t \rceil. \end{aligned} \tag{2.53}$$

Combining this with (2.35), Lemma 2.1, and Lemma 2.2 demonstrates that for all $t \in [0, T]$ it holds that

$$\begin{aligned} \big(\mathbb{E}\left[\|Y_t\|^q\right]\big)^{1/q} &\leq \left[\|x\| + CT + \sup_{t \in [0,T]} \big(\mathbb{E}\big[\|BW_{\lfloor t \rfloor} + \rho(t)B\big(W_{\lceil t \rceil} - W_{\lfloor t \rfloor}\big)\|^q\big]\big)^{1/q}\right] e^{cT} \\ &\leq \left[\|x\| + CT + \sup_{t \in [0,T]} \sqrt{\max\{1, q-1\} \operatorname{Trace}(BB^*)\lceil t \rceil}\right] e^{cT} \\ &= \left[\|x\| + CT + \sqrt{\max\{1, q-1\} \operatorname{Trace}(BB^*)T}\right] e^{cT}. \end{aligned} \tag{2.54}$$

This and (2.52) establish item (iv). This completes the proof of Lemma 2.3. □

3. Numerical Approximations for Partial Differential Equations (PDEs)

In this section we use the weak and strong error estimates which we have presented in Proposition 2.1 and Lemma 2.3 in Section 2 above to establish in Proposition 3.2 below suitable error estimates for Monte Carlo Euler approximations for a class of SDEs with perturbed drift coefficient functions.

Besides Proposition 2.1 and Lemma 2.3, our proof of Proposition 3.2 also employs a special case of the famous Feynman-Kac formula, which provides a connection between solutions of SDEs and solutions of deterministic Kolmogorov PDEs. For completeness we briefly recall in Proposition 3.1 below this special case of the Feynman-Kac formula. Proposition 3.1 is well-known in the literature, cf., e.g., Hairer *et al.* [28, Subsection 4.4], Jentzen *et al.* [39, Theorem 3.1], and Beck *et al.* [4, Theorem 1.1].

3.1. On the Feynman-Kac formula for additive noise driven SDEs

Proposition 3.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $T \in (0,\infty)$, $d, m \in \mathbb{N}$, $B \in \mathbb{R}^{d\times m}$, $\varphi \in C(\mathbb{R}^d, \mathbb{R})$, let $W\colon [0,T]\times\Omega \to \mathbb{R}^m$ be a standard Brownian motion, let $\langle\cdot,\cdot\rangle\colon \mathbb{R}^d\times\mathbb{R}^d \to \mathbb{R}$ be the standard scalar product on $\mathbb{R}^d$, let $\mu\colon \mathbb{R}^d \to \mathbb{R}^d$ be a locally Lipschitz continuous function, and assume that*

$$\inf_{p\in(0,\infty)} \sup_{x\in\mathbb{R}^d} \left[\frac{|\varphi(x)|}{(1+\|x\|^p)} + \frac{\|\mu(x)\|}{(1+\|x\|)} \right] < \infty \tag{3.1}$$

(cf. Definition 2.1). Then

(i) there exist unique stochastic processes $X^x\colon [0,T]\times\Omega \to \mathbb{R}^d$, $x \in \mathbb{R}^d$, with continuous sample paths which satisfy for all $x \in \mathbb{R}^d$, $t \in [0,T]$ that

$$X_t^x = x + \int_0^t \mu(X_s^x)\, ds + BW_t, \tag{3.2}$$

(ii) there exists a unique viscosity solution $u \in \{v \in C([0,T]\times\mathbb{R}^d, \mathbb{R}) \colon \inf_{p\in(0,\infty)} \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \frac{|v(t,x)|}{1+\|x\|^p} < \infty\}$ of

$$\left(\frac{\partial}{\partial t}u\right)(t,x) = \big\langle (\nabla_x u)(t,x), \mu(x) \big\rangle + \frac{1}{2}\operatorname{Trace}\big(BB^*(\operatorname{Hess}_x u)(t,x)\big) \tag{3.3}$$

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

(iii) it holds for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that $\mathbb{E}[|\varphi(X_t^x)|] < \infty$ and $u(t,x) = \mathbb{E}[\varphi(X_t^x)]$.

3.2. Approximation error estimates for Monte Carlo Euler approximations

Proposition 3.2. *Let $T, \kappa \in (0,\infty)$, $\eta \in [1,\infty)$, $p \in [2,\infty)$, let $A_d = (a_{d,i,j})_{(i,j)\in\{1,2,\dots,d\}^2} \in \mathbb{R}^{d\times d}$, $d \in \mathbb{N}$, be symmetric positive semidefinite matrices, let $\nu_d\colon \mathcal{B}([0,T]\times\mathbb{R}^d) \to [0,\infty)$, $d \in \mathbb{N}$, be finite measures which satisfy for all $d \in \mathbb{N}$ that*

$$\left[\int_{[0,T]\times\mathbb{R}^d} \|x\|^{2p\max\{2\kappa,3\}}\, \nu_d(dt,dx)\right]^{1/p} \leq \eta d^\eta, \tag{3.4}$$

let $f_d^m \in C(\mathbb{R}^d, \mathbb{R}^{md-m+1})$*,* $m \in \{0,1\}$*,* $d \in \mathbb{N}$*, and* $F_{d,\varepsilon}^m \in C(\mathbb{R}^d, \mathbb{R}^{md-m+1})$*,* $m \in \{0,1\}$*,* $d \in \mathbb{N}$*,* $\varepsilon \in (0,1]$*, satisfy for all* $d \in \mathbb{N}$*,* $\varepsilon \in (0,1]$*,* $m \in \{0,1\}$*,* $x, y \in \mathbb{R}^d$ *that*

$$|f_d^0(x)| + \mathrm{Trace}(A_d) \le \kappa d^\kappa (1 + \|x\|^\kappa), \quad \|f_d^1(x) - f_d^1(y)\| \le \kappa \|x - y\|, \tag{3.5}$$

$$\|f_d^m(x) - F_{d,\varepsilon}^m(x)\| \le \varepsilon \kappa d^\kappa (1 + \|x\|^\kappa), \quad \|F_{d,\varepsilon}^1(x)\| \le \kappa (d^\kappa + \|x\|), \tag{3.6}$$

$$|F_{d,\varepsilon}^0(x) - F_{d,\varepsilon}^0(y)| \le \kappa d^\kappa (1 + \|x\|^\kappa + \|y\|^\kappa) \|x - y\|, \tag{3.7}$$

let $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space, let* $W^{d,m} \colon [0,T] \times \Omega \to \mathbb{R}^d$*,* $d, m \in \mathbb{N}$*, be independent standard Brownian motions, and let* $Y^{N,d,m,x} \colon [0,T] \times \Omega \to \mathbb{R}^d$*,* $x \in \mathbb{R}^d$*,* $N, d, m \in \mathbb{N}$*, be stochastic processes which satisfy for all* $N, d, m \in \mathbb{N}$*,* $x \in \mathbb{R}^d$*,* $n \in \{0, 1, \dots, N-1\}$*,* $t \in [\frac{nT}{N}, \frac{(n+1)T}{N}]$ *that* $Y_0^{N,d,m,x} = x$ *and*

$$\begin{aligned} Y_t^{N,d,m,x} &= Y_{\frac{nT}{N}}^{N,d,m,x} + \left(\frac{tN}{T} - n\right) \\ &\quad \times \left[\frac{T}{N} F_{d,\min\{(T/N)^{1/2},1\}}^1 \big(Y_{\frac{nT}{N}}^{N,d,m,x}\big) + \sqrt{2A_d}\big(W_{\frac{(n+1)T}{N}}^{d,m} - W_{\frac{nT}{N}}^{d,m}\big)\right] \end{aligned} \tag{3.8}$$

(cf. Definition 2.1*). Then*

(i) for every $d \in \mathbb{N}$ *there exists a unique viscosity solution* $u_d \in \{v \in C([0,T] \times \mathbb{R}^d, \mathbb{R}) \colon \inf_{q \in (0,\infty)} \sup_{(t,x) \in [0,T] \times \mathbb{R}^d} \frac{|v(t,x)|}{1 + \|x\|^q} < \infty\}$ *of*

$$\left(\frac{\partial}{\partial t} u_d\right)(t,x) = \left(\frac{\partial}{\partial x} u_d\right)(t,x)\, f_d^1(x) + \sum_{i,j=1}^d a_{d,i,j} \left(\frac{\partial^2}{\partial x_i \partial x_j} u_d\right)(t,x) \tag{3.9}$$

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

(ii) there exists $\mathcal{C} \in \mathbb{R}$ *such that for all* $d, N, M \in \mathbb{N}$ *it holds that*

$$\begin{aligned} &\left(\mathbb{E}\left[\int_{[0,T]\times\mathbb{R}^d} \left| u_d(t,x) - \frac{1}{M}\left[\sum_{m=1}^M F_{d,\min\{(T/N)^{1/2},1\}}^0 \big(Y_t^{N,d,m,x}\big)\right]\right|^p \nu_d(dt,dx)\right]\right)^{1/p} \\ &\le \mathcal{C}\left[\frac{d^{\kappa(\kappa+4)+\max\{\eta,\kappa(2\kappa+1)\}}}{N^{1/2}} + \frac{d^{\kappa+\max\{\eta,\kappa^2\}}}{M^{1/2}}\right]\left[\max\left\{1, \nu_d([0,T]\times\mathbb{R}^d)\right\}\right]^{1/p}. \end{aligned} \tag{3.10}$$

Proof. Throughout this proof let $\iota \in \mathbb{R}$ satisfy that $\iota = \max\{\kappa, 1\}$, let $C, \mathcal{C}_1, \mathcal{C}_2, \mathcal{C} \in (0,\infty)$ satisfy that

$$C = e^{\kappa^2 T} 2^{\max\{0,\kappa-1\}} \left(\eta + \left[\kappa T + \max\left\{1, \sqrt{2(p\iota - 1)\kappa}\right\} T^{1/2}\right]^\kappa\right), \tag{3.11}$$

$$\begin{aligned} \mathcal{C}_1 &= \iota^2 2^\iota (\kappa+1) \left[\max\left\{1, \sqrt{2\max\{1, 2\kappa-1\}\kappa}\right\}\right]^{2\iota} e^{[3\iota^2 + 1/2]T} (\max\{T,1\})^{\kappa+\iota+3/2} \\ &\quad \times \big[\max\{2\kappa(\kappa+1), 1\}\big]\big[1 + 5\eta\, 2^{\kappa+\iota-1} + 5(4\iota)^{\kappa+\iota}\, 2^{\kappa+\iota-1}\big], \end{aligned} \tag{3.12}$$

$$\begin{aligned} \mathcal{C}_2 &= \frac{1}{\sqrt{2}} \kappa^{3/2} e^{\kappa^2 T} 2^\iota (\max\{T,1\})^{\kappa+1/2} \\ &\quad \times \left(\eta + 1 + \left[\kappa + \max\left\{1, \sqrt{2\max\{1, 2\kappa-1\}\kappa}\right\}\right]^\kappa\right), \end{aligned} \tag{3.13}$$

and

$$\mathcal{C} = \max\{\mathcal{C}_1 + \mathcal{C}_2, 8\kappa(1+C)\sqrt{p-1}\}, \tag{3.14}$$

let $N, M \in \mathbb{N}$, $\delta \in (0,\infty)$ satisfy that $\delta = \sqrt{\frac{T}{N}}$, let $\mathcal{A}_d \in \mathbb{R}^{d\times d}$, $d \in \mathbb{N}$, satisfy for all $d \in \mathbb{N}$ that $\mathcal{A}_d = \sqrt{2A_d}$, let $\varpi_{d,q} \in \mathbb{R}$, $d \in \mathbb{N}$, $q \in (0,\infty)$, satisfy for all $q \in (0,\infty)$, $d \in \mathbb{N}$ that

$$\varpi_{d,q} = \max\Big\{1, \sqrt{\max\{1, q-1\}\operatorname{Trace}((\mathcal{A}_d)^*\mathcal{A}_d)}\Big\}, \tag{3.15}$$

let $X^{d,x}\colon [0,T]\times\Omega \to \mathbb{R}^d$, $x \in \mathbb{R}^d$, $d \in \mathbb{N}$, be stochastic processes with continuous sample paths which satisfy for all $d \in \mathbb{N}$, $x \in \mathbb{R}^d$, $t \in [0,T]$ that

$$X_t^{d,x} = x + \int_0^t f_d^1(X_s^{d,x})\, ds + \mathcal{A}_d W_t^{d,1} \tag{3.16}$$

(cf. item (i) in Proposition 3.1), let $\lfloor\cdot\rfloor\colon [0,T]\to[0,T]$ satisfy for all $t \in [0,T]$ that

$$\lfloor t\rfloor = \max\left(\{0, \delta^2, 2\delta^2, \dots\} \cap [0,t]\right), \tag{3.17}$$

let $\lceil\cdot\rceil\colon [0,T]\to[0,T]$ satisfy for all $t \in [0,T]$ that

$$\lceil t\rceil = \min\left(\{0, \delta^2, 2\delta^2, \dots\} \cap [t,T]\right), \tag{3.18}$$

and let $\mathcal{Y}^{d,x}\colon [0,T]\times\Omega \to \mathbb{R}^d$, $x \in \mathbb{R}^d$, $d \in \mathbb{N}$, be stochastic processes with continuous sample paths which satisfy for all $d \in \mathbb{N}$, $x \in \mathbb{R}^d$, $t \in [0,T]$ that

$$\mathcal{Y}_t^{d,x} = x + \int_0^t F^1_{d,\min\{\delta,1\}}\big(\mathcal{Y}^{d,x}_{\lfloor s\rfloor}\big)\, ds + \mathcal{A}_d W_t^{d,1}. \tag{3.19}$$

Note that Hölder's inequality and (3.4) imply that for all $d \in \mathbb{N}$, $r \in (0, 2\max\{2\kappa,3\})$ it holds that

$$\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} \|x\|^{pr}\, \nu_d(dt,dx)\right]^{1/p} \\
&\leq \left[\int_{[0,T]\times\mathbb{R}^d} \|x\|^{2p\max\{2\kappa,3\}}\, \nu_d(dt,dx)\right]^{r/(2p\max\{2\kappa,3\})} \left[\nu_d([0,T]\times\mathbb{R}^d)\right]^{(1-r/(2\max\{2\kappa,3\}))/p} \\
&\leq (\eta d^\eta)^{r/(2\max\{2\kappa,3\})} \max\left\{1, [\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\right\} \\
&\leq \eta d^\eta \max\left\{1, [\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\right\}.
\end{aligned} \tag{3.20}$$

Furthermore, observe that (3.5) and Proposition 3.1 establish item (i). It thus remains to prove item (ii). For this note that the triangle inequality and Proposition 3.1 ensure that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\left[\left|u_d(t,x) - \frac{1}{M}\left[\sum_{m=1}^M F^0_{d,\min\{\delta,1\}}\big(Y_t^{N,d,m,x}\big)\right]\right|^p\right] \nu_d(dt,dx)\right]^{1/p} \\
&\leq \left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\left[\left|u_d(t,x) - \mathbb{E}\left[F^0_{d,\min\{\delta,1\}}\big(Y_t^{N,d,1,x}\big)\right]\right|^p\right] \nu_d(dt,dx)\right]^{1/p} \\
&\quad + \left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\left[\left|\mathbb{E}\left[F^0_{d,\min\{\delta,1\}}(Y_t^{N,d,1,x})\right]\right.\right.\right.
\end{aligned}$$

$$
\begin{aligned}
&\qquad\qquad -\frac{1}{M}\left[\sum_{m=1}^{M} F^0_{d,\min\{\delta,1\}}(Y^{N,d,m,x}_t)\right]\Bigg|^p\Bigg]\,\nu_d(dt,dx)\Bigg]^{1/p} \\
&= \left[\int_{[0,T]\times\mathbb{R}^d}\left|\mathbb{E}\left[f^0_d(X^{d,x}_t)\right]-\mathbb{E}\left[F^0_{d,\min\{\delta,1\}}(Y^{N,d,1,x}_t)\right]\right|^p\,\nu_d(dt,dx)\right]^{1/p} \\
&\quad + \Bigg[\int_{[0,T]\times\mathbb{R}^d}\mathbb{E}\Bigg[\Bigg|\mathbb{E}\left[F^0_{d,\min\{\delta,1\}}(Y^{N,d,1,x}_t)\right] \\
&\qquad\qquad -\frac{1}{M}\left[\sum_{m=1}^{M} F^0_{d,\min\{\delta,1\}}(Y^{N,d,m,x}_t)\right]\Bigg|^p\Bigg]\,\nu_d(dt,dx)\Bigg]^{1/p}. \qquad (3.21)
\end{aligned}
$$

This and, e.g., [25, Corollary 2.5] prove that for all $d\in\mathbb{N}$ it holds that

$$
\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d}\mathbb{E}\left[\left|u_d(t,x)-\frac{1}{M}\left[\sum_{m=1}^{M}F^0_{d,\min\{\delta,1\}}(Y^{N,d,m,x}_t)\right]\right|^p\right]\nu_d(dt,dx)\right]^{1/p} \\
&\leq \left[\int_{[0,T]\times\mathbb{R}^d}\left|\mathbb{E}\big[f^0_d(X^{d,x}_t)\big]-\mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(Y^{N,d,1,x}_t)\big]\right|^p\nu_d(dt,dx)\right]^{1/p} \\
&\quad+\frac{2\sqrt{p-1}}{M^{1/2}}\Bigg[\int_{[0,T]\times\mathbb{R}^d}\mathbb{E}\Big[\Big|F^0_{d,\min\{\delta,1\}}(Y^{N,d,1,x}_t) \\
&\qquad\qquad -\mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(Y^{N,d,1,x}_t)\big]\Big|^p\Big]\,\nu_d(dt,dx)\Bigg]^{1/p}. \qquad (3.22)
\end{aligned}
$$

Next note that the triangle inequality and Hölder's inequality imply for all $d\in\mathbb{N}$ that

$$
\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d}\mathbb{E}\Big[\big|F^0_{d,\min\{\delta,1\}}(Y^{N,d,1,x}_t)-\mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(Y^{N,d,1,x}_t)\big]\big|^p\Big]\,\nu_d(dt,dx)\right]^{1/p} \\
&\leq \left[\int_{[0,T]\times\mathbb{R}^d}\mathbb{E}\Big[\big|F^0_{d,\min\{\delta,1\}}(Y^{N,d,1,x}_t)\big|^p\Big]\,\nu_d(dt,dx)\right]^{1/p} \\
&\quad+\left[\int_{[0,T]\times\mathbb{R}^d}\big|\mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(Y^{N,d,1,x}_t)\big]\big|^p\,\nu_d(dt,dx)\right]^{1/p} \\
&\leq 2\left[\int_{[0,T]\times\mathbb{R}^d}\mathbb{E}\Big[\big|F^0_{d,\min\{\delta,1\}}(Y^{N,d,1,x}_t)\big|^p\Big]\,\nu_d(dt,dx)\right]^{1/p}. \qquad (3.23)
\end{aligned}
$$

Combining this and (3.22) demonstrates for all $d\in\mathbb{N}$ that

$$
\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d}\mathbb{E}\left[\left|u_d(t,x)-\frac{1}{M}\left[\sum_{m=1}^{M}F^0_{d,\min\{\delta,1\}}(Y^{N,d,m,x}_t)\right]\right|^p\right]\nu_d(dt,dx)\right]^{1/p} \\
&\leq \left[\int_{[0,T]\times\mathbb{R}^d}\left|\mathbb{E}\left[f^0_d(X^{d,x}_t)\right]-\mathbb{E}\left[F^0_{d,\min\{\delta,1\}}(Y^{N,d,1,x}_t)\right]\right|^p\,\nu_d(dt,dx)\right]^{1/p} \\
&\quad+\frac{4\sqrt{p-1}}{M^{1/2}}\left[\int_{[0,T]\times\mathbb{R}^d}\mathbb{E}\big[|F^0_{d,\min\{\delta,1\}}(Y^{N,d,1,x}_t)|^p\big]\,\nu_d(dt,dx)\right]^{1/p}. \qquad (3.24)
\end{aligned}
$$

Next observe that the fact that $\forall\, a,b,q \in [0,\infty)\colon a^q + b^q \le (a+b)^q + (a+b)^q = 2(a+b)^q$ proves that for all $d \in \mathbb{N}$, $x, y \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
&2 \int_0^1 \big[r\|x\| + (1-r)\|y\|\big]^\kappa \, dr \\
&\ge \int_0^1 \big[r^\kappa \|x\|^\kappa + (1-r)^\kappa \|y\|^\kappa\big] \, dr \\
&= \big[\|x\|^\kappa + \|y\|^\kappa\big] \int_0^1 r^\kappa \, dr = \frac{\big[\|x\|^\kappa + \|y\|^\kappa\big]}{\kappa+1}.
\end{aligned} \tag{3.25}$$

This and (3.7) show that for all $d \in \mathbb{N}$, $x, y \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
&\big|F^0_{d,\min\{\delta,1\}}(x) - F^0_{d,\min\{\delta,1\}}(y)\big| \\
&\le \kappa d^\kappa (1 + \|x\|^\kappa + \|y\|^\kappa)\|x-y\| \\
&\le \kappa d^\kappa \left[1 + 2(\kappa+1) \int_0^1 \big[r\|x\| + (1-r)\|y\|\big]^\kappa \, dr\right] \|x-y\| \\
&\le 2\kappa(\kappa+1) d^\kappa \left[1 + \int_0^1 \big[r\|x\| + (1-r)\|y\|\big]^\kappa \, dr\right] \|x-y\|.
\end{aligned} \tag{3.26}$$

Proposition 2.1 (applied with $d \leftarrow d$, $m \leftarrow d$, $\xi \leftarrow x$, $T \leftarrow T$, $c \leftarrow \kappa$, $C \leftarrow \kappa d^\kappa$, $\varepsilon_0 \leftarrow \min\{\delta,1\}\kappa d^\kappa$, $\varepsilon_1 \leftarrow \min\{\delta,1\}\kappa d^\kappa$, $\varepsilon_2 \leftarrow 0$, $\varsigma_0 \leftarrow \kappa$, $\varsigma_1 \leftarrow \kappa$, $\varsigma_2 \leftarrow 0$, $L_0 \leftarrow 2\kappa(\kappa+1)d^\kappa$, $L_1 \leftarrow \kappa$, $\ell \leftarrow \kappa$, $h \leftarrow \delta^2$, $p \leftarrow 2$, $q \leftarrow 2$, $B \leftarrow \mathcal{A}_d$, $(\varpi_r)_{r\in(0,\infty)} \leftarrow (\varpi_{d,r})_{r\in(0,\infty)}$, $\|\cdot\| \leftarrow \|\cdot\|$, $(\Omega,\mathcal{F},\mathbb{P}) \leftarrow (\Omega,\mathcal{F},\mathbb{P})$, $W \leftarrow W^{d,1}$, $F_0 \leftarrow F^0_{d,\min\{\delta,1\}}$, $f_1 \leftarrow f^1_d$, $F_2 \leftarrow \mathrm{id}_{\mathbb{R}^d}$, $\chi(s) \leftarrow \lfloor s \rfloor$, $f_0 \leftarrow f^0_d$, $F_1 \leftarrow F^1_{d,\min\{\delta,1\}}$, $X \leftarrow X^{d,x}$, $Y \leftarrow \mathcal{Y}^{d,x}$ for $d \in \mathbb{N}$, $x \in \mathbb{R}^d$, $s \in [0,T]$ in the notation of Proposition 2.1) hence ensures that for all $d \in \mathbb{N}$, $x \in \mathbb{R}^d$, $t \in [0,T]$ it holds that

$$\begin{aligned}
&\big|\mathbb{E}\big[f^0_d(X^{d,x}_t)\big] - \mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(\mathcal{Y}^{d,x}_t)\big]\big| \\
&\le \big(\min\{\delta,1\}\kappa d^\kappa + \min\{\delta,1\}\kappa d^\kappa + \delta^2 + \delta\big) \\
&\quad \times e^{[\max\{\kappa,1\}\kappa + 1 - 1/2 + \kappa\max\{\kappa,\kappa\} + \max\{\kappa,1\}\kappa]T} (\varpi_{d,\max\{\kappa,2\kappa,2\kappa,2\}})^{\max\{\kappa,\kappa+\max\{1,\kappa\}\}} \\
&\quad \times (\max\{T,1\})^{\max\{\kappa,\kappa+\max\{\kappa,1\}+1/2\}} \max\{2\kappa(\kappa+1)d^\kappa, 1\} \max\{\kappa,1\}\, 2^{\max\{\kappa-1,0\}} \\
&\quad \times \Big[\max\{\kappa d^\kappa, 1\} + 5\max\{\kappa d^\kappa, \kappa, 1\} \\
&\qquad\qquad \times \big(\|x\| + 2\max\{\|f^1_d(0)\|, \kappa d^\kappa, 1\}\big)^{\max\{\kappa,\kappa+\max\{\kappa,1\}\}}\Big].
\end{aligned} \tag{3.27}$$

Next note that (3.5) demonstrates that for all $d \in \mathbb{N}$, $q \in (0,\infty)$ it holds that

$$\begin{aligned}
\varpi_{d,q} &= \max\Big\{1, \sqrt{\max\{1,q-1\}\operatorname{Trace}((\mathcal{A}_d)^*\mathcal{A}_d)}\Big\} \\
&= \max\Big\{1, \sqrt{2\max\{1,q-1\}\operatorname{Trace}(A_d)}\Big\} \\
&\le \max\Big\{1, \sqrt{2\max\{1,q-1\}\kappa d^\kappa}\Big\} \\
&\le d^{\kappa/2} \max\Big\{1, \sqrt{2\max\{1,q-1\}\kappa}\Big\}.
\end{aligned} \tag{3.28}$$

Moreover, observe that (3.6) implies that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned}
\|f^1_d(0)\| &\le \|f^1_d(0) - F^1_{d,\min\{\delta,1\}}(0)\| + \|F^1_{d,\min\{\delta,1\}}(0)\| \\
&\le \min\{\delta,1\}\kappa d^\kappa + \kappa d^\kappa \le 2\kappa d^\kappa.
\end{aligned} \tag{3.29}$$

This, (3.27), (3.28) and the fact that $\iota = \max\{\kappa, 1\}$ ensure that for all $d \in \mathbb{N}$, $x \in \mathbb{R}^d$, $t \in [0,T]$ it holds that

$$\begin{aligned}
&\big|\mathbb{E}\big[f_d^0(X_t^{d,x})\big] - \mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(\mathcal{Y}_t^{d,x})\big]\big| \\
&\leq \big(2\delta\kappa d^\kappa + \delta^2 + \delta\big)\, e^{[3\iota^2+1/2]T} \left[d^{\kappa/2} \max\left\{1, \sqrt{2\max\{1, 2\kappa-1\}\kappa}\right\}\right]^{2\iota} \\
&\quad \times (\max\{T,1\})^{\kappa+\iota+1/2} \max\{2\kappa(\kappa+1)d^\kappa, 1\}\, \iota\, 2^{\iota-1} \\
&\quad \times \Big[\max\{\kappa d^\kappa, 1\} + 5\max\{\kappa d^\kappa, 1\}\big(\|x\| + 2\max\{2\kappa d^\kappa, 1\}\big)^{\kappa+\iota}\Big]. \qquad (3.30)
\end{aligned}$$

The triangle inequality hence ensures that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} \big|\mathbb{E}\big[f_d^0(X_t^{d,x})\big] - \mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(\mathcal{Y}_t^{d,x})\big]\big|^p \,\nu_d(dt,dx)\right]^{1/p} \\
&\leq \big(2\delta\kappa d^\kappa + \delta^2 + \delta\big)\, e^{[3\iota^2+1/2]T} \left[d^{\kappa/2} \max\left\{1, \sqrt{2\max\{1, 2\kappa-1\}\kappa}\right\}\right]^{2\iota} \\
&\quad \times (\max\{T,1\})^{\kappa+\iota+1/2} \max\{2\kappa(\kappa+1)d^\kappa, 1\}\, \iota 2^{\iota-1} \\
&\quad \times \Bigg[\max\{\kappa d^\kappa, 1\}[\nu_d([0,T]\times\mathbb{R}^d)]^{1/p} \\
&\qquad +5\max\{\kappa d^\kappa, 1\}\left[\int_{[0,T]\times\mathbb{R}^d} \big(\|x\| + 2\max\{2\kappa d^\kappa, 1\}\big)^{p(\kappa+\iota)}\,\nu_d(dt,dx)\right]^{1/p}\Bigg]. \qquad (3.31)
\end{aligned}$$

Moreover, observe that the fact that $\forall\, y, z \in \mathbb{R}, \alpha \in [1,\infty)\colon |y+z|^\alpha \leq 2^{\alpha-1}(|y|^\alpha + |z|^\alpha)$ and the triangle inequality demonstrate that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} \big(\|x\| + 2\max\{2\kappa d^\kappa, 1\}\big)^{p(\kappa+\iota)}\,\nu_d(dt,dx)\right]^{1/p} \\
&\leq 2^{\kappa+\iota-1}\left[\int_{[0,T]\times\mathbb{R}^d} \Big[\|x\|^{\kappa+\iota} + (2\max\{2\kappa d^\kappa, 1\})^{\kappa+\iota}\Big]^p\,\nu_d(dt,dx)\right]^{1/p} \\
&\leq 2^{\kappa+\iota-1}\left[\int_{[0,T]\times\mathbb{R}^d} \|x\|^{p(\kappa+\iota)}\,\nu_d(dt,dx)\right]^{1/p} \\
&\quad + 2^{\kappa+\iota-1}[\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}(2\max\{2\kappa d^\kappa, 1\})^{\kappa+\iota}. \qquad (3.32)
\end{aligned}$$

This, the fact that $\kappa + \iota = \max\{2\kappa, \kappa+1\} < 2\max\{2\kappa, 3\}$, and (3.20) ensure that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} \big(\|x\| + 2\max\{2\kappa d^\kappa, 1\}\big)^{p(\kappa+\iota)}\,\nu_d(dt,dx)\right]^{1/p} \\
&\leq 2^{\kappa+\iota-1}\eta d^\eta \max\left\{1, [\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\right\} \\
&\quad + 2^{\kappa+\iota-1}[\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}(2\max\{2\kappa d^\kappa, 1\})^{\kappa+\iota} \\
&\leq 2^{\kappa+\iota-1}\max\left\{1, [\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\right\}\big[\eta d^\eta + (2\max\{2\kappa d^\kappa, 1\})^{\kappa+\iota}\big]. \qquad (3.33)
\end{aligned}$$

In addition, note that the fact that $\delta = \sqrt{\frac{T}{N}}$ proves that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned} &2\delta\kappa d^\kappa + \delta^2 + \delta \leq 2\delta\kappa d^\kappa + \delta\sqrt{T} + \delta \\ &\leq 2\delta\kappa d^\kappa + 2\max\{\sqrt{T}, 1\}\delta \leq 2\max\{\sqrt{T}, 1\}(\delta\kappa d^\kappa + \delta). \end{aligned} \tag{3.34}$$

Combining this, (3.31), and (3.33) shows that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned} &\left[\int_{[0,T]\times\mathbb{R}^d} \big|\mathbb{E}\big[f_d^0(X_t^{d,x})\big] - \mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(\mathcal{Y}_t^{d,x})\big]\big|^p \, \nu_d(dt, dx)\right]^{1/p} \\ &\leq \iota 2^\iota \max\{\sqrt{T}, 1\}\,(\delta\kappa d^\kappa + \delta)\,\Big[d^{\kappa/2}\max\Big\{1, \sqrt{2\max\{1, 2\kappa-1\}\kappa}\Big\}\Big]^{2\iota} \\ &\quad \times e^{[3\iota^2+1/2]T}(\max\{T,1\})^{\kappa+\iota+1/2}\max\{2\kappa(\kappa+1)d^\kappa, 1\}\max\{\kappa d^\kappa, 1\} \\ &\quad \times \max\Big\{1, [\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\Big\}\,\Big(1 + 5\,\big[\eta d^\eta + (2\max\{2\kappa d^\kappa, 1\})^{\kappa+\iota}\big]\,2^{\kappa+\iota-1}\Big). \end{aligned} \tag{3.35}$$

Hence, we obtain that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned} &\left[\int_{[0,T]\times\mathbb{R}^d} \big|\mathbb{E}\big[f_d^0(X_t^{d,x})\big] - \mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(\mathcal{Y}_t^{d,x})\big]\big|^p \, \nu_d(dt, dx)\right]^{1/p} \\ &\leq \iota 2^\iota d^{3\kappa+\kappa\iota}\delta(\kappa+1)\,\Big[\max\Big\{1, \sqrt{2\max\{1, 2\kappa-1\}\kappa}\Big\}\Big]^{2\iota} \\ &\quad \times e^{[3\iota^2+1/2]T}(\max\{T,1\})^{\kappa+\iota+1}\max\{2\kappa(\kappa+1), 1\}\max\{\kappa, 1\} \\ &\quad \times \max\Big\{1, [\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\Big\}\,\Big(1 + 5\,\big[\eta d^\eta + (2\max\{2\kappa d^\kappa, 1\})^{\kappa+\iota}\big]\,2^{\kappa+\iota-1}\Big) \\ &\leq \iota^2 2^\iota d^{3\kappa+\kappa\iota+\max\{\eta,\kappa(\kappa+\iota)\}}\delta(\kappa+1)\,\Big[\max\Big\{1, \sqrt{2\max\{1, 2\kappa-1\}\kappa}\Big\}\Big]^{2\iota} \\ &\quad \times e^{[3\iota^2+1/2]T}(\max\{T,1\})^{\kappa+\iota+1}\max\{2\kappa(\kappa+1), 1\}\max\Big\{1, [\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\Big\} \\ &\quad \times \big[1 + 5\eta\,2^{\kappa+\iota-1} + 5(4\iota)^{\kappa+\iota}\,2^{\kappa+\iota-1}\big]. \end{aligned} \tag{3.36}$$

The fact that $\delta = \sqrt{\frac{T}{N}}$ and (3.12) hence prove that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned} &\left[\int_{[0,T]\times\mathbb{R}^d} \big|\mathbb{E}\big[f_d^0(X_t^{d,x})\big] - \mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(\mathcal{Y}_t^{d,x})\big]\big|^p \, \nu_d(dt, dx)\right]^{1/p} \\ &\leq N^{-1/2}\iota^2 2^\iota d^{3\kappa+\kappa\iota+\max\{\eta,\kappa(\kappa+\iota)\}}(\kappa+1)\,\Big[\max\Big\{1, \sqrt{2\max\{1, 2\kappa-1\}\kappa}\Big\}\Big]^{2\iota} \\ &\quad \times e^{[3\iota^2+1/2]T}(\max\{T,1\})^{\kappa+\iota+3/2}\max\{2\kappa(\kappa+1), 1\} \\ &\quad \times \big[1 + 5\eta\,2^{\kappa+\iota-1} + 5(4\iota)^{\kappa+\iota}\,2^{\kappa+\iota-1}\big]\max\Big\{1, [\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\Big\} \\ &= N^{-1/2}\,\mathcal{C}_1\,d^{3\kappa+\kappa\iota+\max\{\eta,\kappa(\kappa+\iota)\}}\max\Big\{1, [\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\Big\}. \end{aligned} \tag{3.37}$$

Furthermore, observe that (3.7), the Cauchy-Schwarz inequality, and the triangle inequality ensure that for all $d \in \mathbb{N}$, $x \in \mathbb{R}^d$, $t \in [0,T]$ it holds that

$$\begin{aligned} &\big|\mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(\mathcal{Y}_t^{d,x})\big] - \mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(Y_t^{N,d,1,x})\big]\big| \\ &\leq \mathbb{E}\big[\big|F^0_{d,\min\{\delta,1\}}(\mathcal{Y}_t^{d,x}) - F^0_{d,\min\{\delta,1\}}(Y_t^{N,d,1,x})\big|\big] \end{aligned}$$

$$
\begin{aligned}
&\le \kappa d^\kappa \mathbb{E}\big[\big(1 + \|\mathcal{Y}_t^{d,x}\|^\kappa + \|Y_t^{N,d,1,x}\|^\kappa\big)\|\mathcal{Y}_t^{d,x} - Y_t^{N,d,1,x}\|\big] \\
&\le \kappa d^\kappa \big(\mathbb{E}\big[\big(1 + \|\mathcal{Y}_t^{d,x}\|^\kappa + \|Y_t^{N,d,1,x}\|^\kappa\big)^2\big]\big)^{1/2}\big(\mathbb{E}\big[\|\mathcal{Y}_t^{d,x} - Y_t^{N,d,1,x}\|^2\big]\big)^{1/2} \\
&\le \kappa d^\kappa \left[1 + \big(\mathbb{E}\big[\|\mathcal{Y}_t^{d,x}\|^{2\kappa}\big]\big)^{1/2} + \big(\mathbb{E}\big[\|Y_t^{N,d,1,x}\|^{2\kappa}\big]\big)^{1/2}\right] \left(\mathbb{E}\big[\|\mathcal{Y}_t^{d,x} - Y_t^{N,d,1,x}\|^2\big]\right)^{1/2}. \qquad (3.38)
\end{aligned}
$$

Next note that (3.6) and items (ii)-(iii) in Lemma 2.3 prove that for all $d \in \mathbb{N}$, $x \in \mathbb{R}^d$, $t \in [0,T]$ it holds that

$$
\begin{aligned}
&\big(\mathbb{E}\big[\|\mathcal{Y}_t^{d,x} - Y_t^{N,d,1,x}\|^2\big]\big)^{1/2} \\
&\le \frac{1}{2}\sqrt{\big(\lceil t\rceil - \lfloor t\rfloor\big)\operatorname{Trace}(\mathcal{A}_d\mathcal{A}_d^*)} \\
&= \frac{1}{2}\sqrt{2\big(\lceil t\rceil - \lfloor t\rfloor\big)\operatorname{Trace}(A_d)} \le \frac{1}{\sqrt{2}}\delta\sqrt{\kappa d^\kappa}. \qquad (3.39)
\end{aligned}
$$

Moreover, observe that Hölder's inequality, (3.6), and item (iv) in Lemma 2.3 show that for all $q \in (0,\infty)$, $d, m \in \mathbb{N}$, $x \in \mathbb{R}^d$, $t \in [0,T]$ it holds that

$$
\begin{aligned}
&\max\left\{\big(\mathbb{E}\big[\|\mathcal{Y}_t^{d,x}\|^q\big]\big)^{1/q}, \big(\mathbb{E}\big[\|Y_t^{N,d,m,x}\|^q\big]\big)^{1/q}\right\} \\
&\le \max\left\{\big(\mathbb{E}\big[\|\mathcal{Y}_t^{d,x}\|^{\max\{q,1\}}\big]\big)^{1/\max\{q,1\}}, \big(\mathbb{E}\big[\|Y_t^{N,d,m,x}\|^{\max\{q,1\}}\big]\big)^{1/\max\{q,1\}}\right\} \\
&\le \left[\|x\| + \kappa d^\kappa T + \sqrt{\max\{1, \max\{1,q\} - 1\}T\operatorname{Trace}(\mathcal{A}_d\mathcal{A}_d^*)}\right] e^{\kappa T}. \qquad (3.40)
\end{aligned}
$$

Combining this with (3.15) ensures that for all $q \in (0,\infty)$, $d, m \in \mathbb{N}$, $x \in \mathbb{R}^d$, $t \in [0,T]$ it holds that

$$
\begin{aligned}
&\max\left\{\big(\mathbb{E}\big[\|\mathcal{Y}_t^{d,x}\|^q\big]\big)^{1/q}, \big(\mathbb{E}\big[\|Y_t^{N,d,m,x}\|^q\big]\big)^{1/q}\right\} \\
&\le \left[\|x\| + \kappa d^\kappa T + \varpi_{d,\max\{q,1\}}T^{1/2}\right] e^{\kappa T}. \qquad (3.41)
\end{aligned}
$$

The fact that $\forall\, y, z \in \mathbb{R}$, $\alpha \in (0,\infty)$: $|y+z|^\alpha \le 2^{\max\{0,\alpha-1\}}(|y|^\alpha + |z|^\alpha)$, (3.38), and (3.39) hence demonstrate that for all $d \in \mathbb{N}$, $x \in \mathbb{R}^d$, $t \in [0,T]$ it holds that

$$
\begin{aligned}
&\big|\mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(\mathcal{Y}_t^{d,x})\big] - \mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(Y_t^{N,d,1,x})\big]\big| \\
&\le \kappa d^\kappa \left[1 + 2\left(\|x\| + \kappa d^\kappa T + \varpi_{d,\max\{2\kappa,1\}}T^{1/2}\right)^\kappa e^{\kappa^2 T}\right]\frac{1}{\sqrt{2}}\delta\sqrt{\kappa d^\kappa} \\
&\le \frac{1}{\sqrt{2}}\delta\big(\kappa d^\kappa\big)^{3/2} e^{\kappa^2 T}\left[1 + 2^{1+\max\{0,\kappa-1\}}\left(\|x\|^\kappa + \big[\kappa d^\kappa T + \varpi_{d,\max\{2\kappa,1\}}T^{1/2}\big]^\kappa\right)\right] \\
&\le \frac{1}{\sqrt{2}}\delta\big(\kappa d^\kappa\big)^{3/2} e^{\kappa^2 T} 2^{\max\{1,\kappa\}}\left[\|x\|^\kappa + \left(1 + \big[\kappa d^\kappa T + \varpi_{d,\max\{2\kappa,1\}}T^{1/2}\big]^\kappa\right)\right]. \qquad (3.42)
\end{aligned}
$$

Combining this with the triangle inequality assures that for all $d \in \mathbb{N}$ it holds that

$$
\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} \big|\mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(\mathcal{Y}_t^{d,x})\big] - \mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(Y_t^{N,d,1,x})\big]\big|^p \,\nu_d(dt,dx)\right]^{1/p} \\
&\le \frac{1}{\sqrt{2}}\delta\big(\kappa d^\kappa\big)^{3/2} e^{\kappa^2 T} 2^\iota \left[\int_{[0,T]\times\mathbb{R}^d} \left[\|x\|^\kappa + \left(1 + \big[\kappa d^\kappa T + \varpi_{d,\max\{2\kappa,1\}}T^{1/2}\big]^\kappa\right)\right]^p \nu_d(dt,dx)\right]^{1/p} \\
&\le \frac{1}{\sqrt{2}}\delta\big(\kappa d^\kappa\big)^{3/2} e^{\kappa^2 T} 2^\iota \left[\int_{[0,T]\times\mathbb{R}^d} \|x\|^{p\kappa}\,\nu_d(dt,dx)\right]^{1/p}
\end{aligned}
$$

$$+ \frac{1}{\sqrt{2}}\delta\big(\kappa d^{\kappa}\big)^{3/2} e^{\kappa^2 T} 2^{\iota}\Big(1 + \big[\kappa d^{\kappa} T + \varpi_{d,\max\{2\kappa,1\}} T^{1/2}\big]^{\kappa}\Big)[\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}. \tag{3.43}$$

Next note that (3.28) ensures for all $d \in \mathbb{N}$ that

$$\begin{aligned} &1 + \big[\kappa d^{\kappa} T + \varpi_{d,\max\{2\kappa,1\}} T^{1/2}\big]^{\kappa} \\ &\leq 1 + \Big[\kappa d^{\kappa} T + d^{\kappa/2} \max\Big\{1, \sqrt{2\max\{1, 2\kappa-1\}\kappa}\Big\} T^{1/2}\Big]^{\kappa} \\ &\leq 1 + d^{(\kappa^2)}(\max\{T,1\})^{\kappa}\Big[\kappa + \max\Big\{1, \sqrt{2\max\{1,2\kappa-1\}\kappa}\Big\}\Big]^{\kappa}. \end{aligned} \tag{3.44}$$

Hence, we obtain for all $d \in \mathbb{N}$ that

$$\begin{aligned} &1 + \big[\kappa d^{\kappa} T + \varpi_{d,\max\{2\kappa,1\}} T^{1/2}\big]^{\kappa} \\ &\leq d^{(\kappa^2)}(\max\{T,1\})^{\kappa}\Big(1 + \Big[\kappa + \max\Big\{1, \sqrt{2\max\{1,2\kappa-1\}\kappa}\Big\}\Big]^{\kappa}\Big). \end{aligned} \tag{3.45}$$

This, (3.43), (3.20), and the fact that $\kappa < 2\max\{2\kappa, 3\}$ establish that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned} &\left[\int_{[0,T]\times\mathbb{R}^d} \big|\mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(\mathcal{Y}^{d,x}_t)\big] - \mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(Y^{N,d,1,x}_t)\big]\big|^p \,\nu_d(dt,dx)\right]^{1/p} \\ &\leq \frac{1}{\sqrt{2}}\delta\big(\kappa d^{\kappa}\big)^{3/2} e^{\kappa^2 T} 2^{\iota} \eta d^{\eta} \max\Big\{1, [\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\Big\} \\ &\quad + \frac{1}{\sqrt{2}}\delta\big(\kappa d^{\kappa}\big)^{3/2} e^{\kappa^2 T} 2^{\iota} d^{(\kappa^2)}(\max\{T,1\})^{\kappa}[\nu_d([0,T]\times\mathbb{R}^d)]^{1/p} \\ &\quad \times \Big(1 + \Big[\kappa + \max\Big\{1, \sqrt{2\max\{1,2\kappa-1\}\kappa}\Big\}\Big]^{\kappa}\Big). \end{aligned} \tag{3.46}$$

Combining this with (3.13) and the fact that $\delta = \sqrt{\frac{T}{N}}$ implies that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned} &\left[\int_{[0,T]\times\mathbb{R}^d} \big|\mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(\mathcal{Y}^{d,x}_t)\big] - \mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(Y^{N,d,1,x}_t)\big]\big|^p \,\nu_d(dt,dx)\right]^{1/p} \\ &\leq \delta d^{3\kappa/2+\eta} \frac{1}{\sqrt{2}} \kappa^{3/2} e^{\kappa^2 T} 2^{\iota} \eta \max\Big\{1, [\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\Big\} \\ &\quad + \delta d^{3\kappa/2+\kappa^2} \frac{1}{\sqrt{2}} \kappa^{3/2} e^{\kappa^2 T} 2^{\iota} [\nu_d([0,T]\times\mathbb{R}^d)]^{1/p} \\ &\quad \times (\max\{T,1\})^{\kappa}\Big(1 + \Big[\kappa + \max\Big\{1, \sqrt{2\max\{1,2\kappa-1\}\kappa}\Big\}\Big]^{\kappa}\Big) \\ &\leq \delta d^{3\kappa/2+\max\{\eta,\kappa^2\}} \frac{1}{\sqrt{2}} \kappa^{3/2} e^{\kappa^2 T} 2^{\iota} \max\Big\{1, [\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\Big\} \\ &\quad \times (\max\{T,1\})^{\kappa}\Big(\eta + 1 + \Big[\kappa + \max\Big\{1, \sqrt{2\max\{1,2\kappa-1\}\kappa}\Big\}\Big]^{\kappa}\Big) \\ &\leq N^{-1/2} d^{3\kappa/2+\max\{\eta,\kappa^2\}} \frac{1}{\sqrt{2}} \kappa^{3/2} e^{\kappa^2 T} 2^{\iota} \max\Big\{1, [\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\Big\} \\ &\quad \times (\max\{T,1\})^{\kappa+1/2}\Big(\eta + 1 + \Big[\kappa + \max\Big\{1, \sqrt{2\max\{1,2\kappa-1\}\kappa}\Big\}\Big]^{\kappa}\Big) \\ &= N^{-1/2} d^{3\kappa/2+\max\{\eta,\kappa^2\}} \mathcal{C}_2 \max\Big\{1, [\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\Big\}. \end{aligned} \tag{3.47}$$

Next note that the triangle inequality proves that for all $d \in \mathbb{N}$ it holds that

$$\left[\int_{[0,T]\times\mathbb{R}^d} \big|\mathbb{E}\big[f^0_d(X^{d,x}_t)\big] - \mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(Y^{N,d,1,x}_t)\big]\big|^p \,\nu_d(dt,dx)\right]^{1/p}$$

$$\leq \left[\int_{[0,T]\times\mathbb{R}^d} \big|\mathbb{E}\big[f_d^0(X_t^{d,x})\big] - \mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(\mathcal{Y}_t^{d,x})\big]\big|^p \, \nu_d(dt,dx)\right]^{1/p} + \left[\int_{[0,T]\times\mathbb{R}^d} \big|\mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(\mathcal{Y}_t^{d,x})\big] - \mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(Y_t^{N,d,1,x})\big]\big|^p \, \nu_d(dt,dx)\right]^{1/p}. \tag{3.48}$$

This, (3.37), and (3.47) ensure that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned} &\left[\int_{[0,T]\times\mathbb{R}^d} \big|\mathbb{E}\big[f_d^0(X_t^{d,x})\big] - \mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(Y_t^{N,d,1,x})\big]\big|^p \, \nu_d(dt,dx)\right]^{1/p} \\ &\leq N^{-1/2}\, \mathcal{C}_1\, d^{3\kappa+\kappa\iota+\max\{\eta,\kappa(\kappa+\iota)\}} \max\Big\{1, [\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\Big\} \\ &\quad + N^{-1/2} d^{3\kappa/2+\max\{\eta,\kappa^2\}} \mathcal{C}_2 \max\Big\{1, [\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\Big\}. \end{aligned} \tag{3.49}$$

The fact that $\iota \leq \kappa + 1$, the fact that $\frac{3\kappa}{2} + \max\{\eta, \kappa^2\} \leq \kappa(\kappa+4) + \max\{\eta, \kappa(2\kappa+1)\}$, and (3.14) hence demonstrate that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned} &\left[\int_{[0,T]\times\mathbb{R}^d} \big|\mathbb{E}\big[f_d^0(X_t^{d,x})\big] - \mathbb{E}\big[F^0_{d,\min\{\delta,1\}}(Y_t^{N,d,1,x})\big]\big|^p \, \nu_d(dt,dx)\right]^{1/p} \\ &\leq N^{-1/2} d^{\kappa(\kappa+4)+\max\{\eta,\kappa(2\kappa+1)\}} \, [\mathcal{C}_1 + \mathcal{C}_2] \max\Big\{1, [\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\Big\} \\ &\leq N^{-1/2} d^{\kappa(\kappa+4)+\max\{\eta,\kappa(2\kappa+1)\}} \mathcal{C} \max\Big\{1, [\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\Big\}. \end{aligned} \tag{3.50}$$

Next observe that (3.5) and (3.6) prove for all $d \in \mathbb{N}$, $x \in \mathbb{R}^d$ that

$$\begin{aligned} |F^0_{d,\min\{\delta,1\}}(x)| &\leq |F^0_{d,\min\{\delta,1\}}(x) - f_d^0(x)| + |f_d^0(x)| \\ &\leq \min\{\delta,1\}\kappa d^\kappa(1+\|x\|^\kappa) + \kappa d^\kappa(1+\|x\|^\kappa) \\ &\leq 2\kappa d^\kappa(1+\|x\|^\kappa). \end{aligned} \tag{3.51}$$

Moreover, note that (3.41) and the fact that $\forall\, y,z \in \mathbb{R}, \alpha \in (0,\infty)\colon |y+z|^\alpha \leq 2^{\max\{0,\alpha-1\}}(|y|^\alpha + |z|^\alpha)$ imply that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned} &\left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[\|Y_t^{N,d,1,x}\|^{p\kappa}\big] \, \nu_d(dt,dx)\right]^{1/p} \\ &\leq e^{\kappa^2 T} \left[\int_{[0,T]\times\mathbb{R}^d} \big[\|x\| + \kappa d^\kappa T + \varpi_{d,\max\{p\kappa,1\}} T^{1/2}\big]^{p\kappa} \, \nu_d(dt,dx)\right]^{1/p} \\ &\leq e^{\kappa^2 T} 2^{\max\{0,\kappa-1\}} \left[\int_{[0,T]\times\mathbb{R}^d} \big[\|x\|^\kappa + (\kappa d^\kappa T + \varpi_{d,\max\{p\kappa,1\}} T^{1/2})^\kappa\big]^p \, \nu_d(dt,dx)\right]^{1/p}. \end{aligned} \tag{3.52}$$

Combining this with the triangle inequality ensures that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned} &\left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[\|Y_t^{N,d,1,x}\|^{p\kappa}\big] \, \nu_d(dt,dx)\right]^{1/p} \\ &\leq e^{\kappa^2 T} 2^{\max\{0,\kappa-1\}} \left[\int_{[0,T]\times\mathbb{R}^d} \|x\|^{p\kappa} \, \nu_d(dt,dx)\right]^{1/p} \\ &\quad + e^{\kappa^2 T} 2^{\max\{0,\kappa-1\}} (\kappa d^\kappa T + \varpi_{d,\max\{p\kappa,1\}} T^{1/2})^\kappa [\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}. \end{aligned} \tag{3.53}$$

This, (3.20), and the fact that $\kappa < 2\max\{2\kappa, 3\}$ demonstrate that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned}&\left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[\|Y_t^{N,d,1,x}\|^{p\kappa}\big]\,\nu_d(dt,dx)\right]^{1/p}\\&\leq e^{\kappa^2 T} 2^{\max\{0,\kappa-1\}}\left(\eta d^\eta + (\kappa d^\kappa T + \varpi_{d,\max\{p\kappa,1\}} T^{1/2})^\kappa\right)\\&\quad\times \max\left\{1, [\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\right\}. \end{aligned} \tag{3.54}$$

In addition, observe that (3.28) and the fact that $\max\{p\kappa, 1\} \leq p\iota$ prove that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned}\varpi_{d,\max\{p\kappa,1\}} &\leq d^{\kappa/2} \max\left\{1, \sqrt{2\max\{1, \max\{p\kappa,1\}-1\}\kappa}\right\}\\&\leq d^{\kappa/2} \max\left\{1, \sqrt{2\max\{1, p\iota-1\}\kappa}\right\}\\&= d^{\kappa/2}\max\left\{1, \sqrt{2(p\iota-1)\kappa}\right\}. \end{aligned} \tag{3.55}$$

This and (3.54) ensure that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned}&\left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[\|Y_t^{N,d,1,x}\|^{p\kappa}\big]\,\nu_d(dt,dx)\right]^{1/p}\\&\leq e^{\kappa^2 T}2^{\max\{0,\kappa-1\}}\left(\eta d^\eta + \left[\kappa d^\kappa T + d^{\kappa/2}\max\left\{1,\sqrt{2(p\iota-1)\kappa}\right\}T^{1/2}\right]^\kappa\right)\\&\quad\times\max\left\{1,[\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\right\}\\&\leq e^{\kappa^2 T}2^{\max\{0,\kappa-1\}}d^{\max\{\eta,\kappa^2\}}\left(\eta + \left[\kappa T + \max\left\{1,\sqrt{2(p\iota-1)\kappa}\right\}T^{1/2}\right]^\kappa\right)\\&\quad\times\max\left\{1,[\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\right\}\\&= C d^{\max\{\eta,\kappa^2\}}\max\left\{1,[\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\right\}. \end{aligned} \tag{3.56}$$

This, the triangle inequality, and (3.51) assure that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned}&\left[\int_{[0,T]\times\mathbb{R}^d}\mathbb{E}\big[|F^0_{d,\min\{\delta,1\}}(Y_t^{N,d,1,x})|^p\big]\,\nu_d(dt,dx)\right]^{1/p}\\&\leq 2\kappa d^\kappa\left([\nu_d([0,T]\times\mathbb{R}^d)]^{1/p} + C d^{\max\{\eta,\kappa^2\}}\max\left\{1,[\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\right\}\right)\\&\leq 2\kappa d^{\kappa+\max\{\eta,\kappa^2\}}\big(1+C\big)\max\left\{1,[\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\right\}. \end{aligned} \tag{3.57}$$

Combining this, Fubini's theorem, (3.50), (3.14), and (3.24) proves that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned}&\left(\mathbb{E}\left[\int_{[0,T]\times\mathbb{R}^d}\left|u_d(t,x) - \frac{1}{M}\left[\sum_{m=1}^M F^0_{d,\min\{\delta,1\}}\big(Y_t^{N,d,m,x}\big)\right]\right|^p\nu_d(dt,dx)\right]\right)^{1/p}\\&= N^{-1/2}\mathcal{C}d^{\kappa(\kappa+4)+\max\{\eta,\kappa(2\kappa+1)\}}\max\left\{1,[\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\right\}\\&\quad+\frac{4\sqrt{p-1}}{M^{1/2}}2\kappa d^{\kappa+\max\{\eta,\kappa^2\}}(1+C)\max\left\{1,[\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\right\}\\&\leq \mathcal{C}\left[\frac{d^{\kappa(\kappa+4)+\max\{\eta,\kappa(2\kappa+1)\}}}{N^{1/2}} + \frac{d^{\kappa+\max\{\eta,\kappa^2\}}}{M^{1/2}}\right]\left[\max\left\{1,\nu_d([0,T]\times\mathbb{R}^d)\right\}\right]^{1/p}. \end{aligned} \tag{3.58}$$

This establishes item (ii). This completes the proof of Proposition 3.2. □

4. Deep Artificial Neural Network (ANN) Approximations for PDEs

In this section we establish in Theorem 4.1 in Subsection 4.6 below the main result of this article.

Theorem 4.1, in particular, proves that for every $T \in (0, \infty)$, $a \in \mathbb{R}$, $b \in (a, \infty)$ it holds that solutions of certain Kolmogorov PDEs can be approximated by deep ANNs on the space-time region $[0, T] \times [a, b]^d$ without the curse of dimensionality. In our proof of Theorem 4.1 we employ the auxiliary intermediate result in Proposition 4.3 in Subsection 4.5 below. Our proof of Proposition 4.3, in turn, uses the error estimates for Monte Carlo Euler approximations which we have presented in Proposition 3.2 in Section 3 above as well as the ANN approximation result for Monte Carlo Euler approximations in Corollary 4.1 in Subsection 4.4 below.

Our proof of Corollary 4.1 employs the auxiliary results in Proposition 4.2 and Lemma 4.2 in Subsection 4.4 below. Our proof of Proposition 4.2, in turn, uses the ANN approximation result for Monte Carlo Euler approximations in Proposition 4.1 in Subsection 4.4 below. Our proof of Proposition 4.1 is based on an application of [26, Proposition 3.10] and is very similar to the proof of [26, Theorem 3.12].

Our proof of Theorem 4.1 in Subsection 4.6 below also employs several well-known concepts and results from an appropriate calculus for ANNs from the scientific literature which we briefly recall in Subsections 4.1-4.3 below. In particular, Definition 4.1 is, e.g., [26, Definition 2.1], Definition 4.2 is, e.g., [26, Definition 2.2], Definition 4.3 is, e.g., [26, Definition 2.3], Definition 4.5 is, e.g., [26, Definition 2.5], Lemma 4.1 is, e.g., [26, Lemma 2.8], and Definition 4.6 is, e.g., [26, Definition 2.15].

4.1. ANNs

Definition 4.1 (ANNs). *We denote by* $\mathbf{N}$ *the set given by*

$$\mathbf{N} = \cup_{L \in \mathbb{N}} \cup_{l_0, l_1, \ldots, l_L \in \mathbb{N}} \left(\times_{k=1}^{L} (\mathbb{R}^{l_k \times l_{k-1}} \times \mathbb{R}^{l_k}) \right) \tag{4.1}$$

and we denote by $\mathcal{P}, \mathcal{L}, \mathcal{I}, \mathcal{O} \colon \mathbf{N} \to \mathbb{N}$ *and* $\mathcal{D} \colon \mathbf{N} \to \cup_{L=2}^{\infty} \mathbb{N}^L$ *the functions which satisfy for all* $L \in \mathbb{N}$, $l_0, l_1, \ldots, l_L \in \mathbb{N}$, $\Phi \in (\times_{k=1}^{L} (\mathbb{R}^{l_k \times l_{k-1}} \times \mathbb{R}^{l_k}))$ *that*

$$\mathcal{P}(\Phi) = \sum_{k=1}^{L} l_k (l_{k-1} + 1), \quad \mathcal{L}(\Phi) = L, \quad \mathcal{I}(\Phi) = l_0, \quad \mathcal{O}(\Phi) = l_L, \tag{4.2}$$

and $\mathcal{D}(\Phi) = (l_0, l_1, \ldots, l_L)$.

4.2. Realizations of ANNs

Definition 4.2 (Multidimensional Versions). *Let* $d \in \mathbb{N}$ *and let* $\psi \colon \mathbb{R} \to \mathbb{R}$ *be a function. Then we denote by* $\mathfrak{M}_{\psi,d} \colon \mathbb{R}^d \to \mathbb{R}^d$ *the function which satisfies for all* $x = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$ *that*

$$\mathfrak{M}_{\psi,d}(x) = (\psi(x_1), \psi(x_2), \ldots, \psi(x_d)). \tag{4.3}$$

Definition 4.3 (Realizations Associated to ANNs). *Let* $a \in C(\mathbb{R}, \mathbb{R})$*. Then we denote by* $\mathcal{R}_a \colon \mathbf{N} \to (\cup_{k,l \in \mathbb{N}} C(\mathbb{R}^k, \mathbb{R}^l))$ *the function which satisfies for all* $L \in \mathbb{N}$, $l_0, l_1, \ldots, l_L \in \mathbb{N}$, $\Phi = ((W_1, B_1), (W_2, B_2), \ldots, (W_L, B_L)) \in (\times_{k=1}^{L} (\mathbb{R}^{l_k \times l_{k-1}} \times \mathbb{R}^{l_k}))$, $x_0 \in \mathbb{R}^{l_0}, x_1 \in \mathbb{R}^{l_1}, \ldots, x_L \in \mathbb{R}^{l_L}$ *with* $\forall\, k \in \mathbb{N} \cap (0, L) \colon x_k = \mathfrak{M}_{a, l_k}(W_k x_{k-1} + B_k)$ *that*

$$\mathcal{R}_a(\Phi) \in C(\mathbb{R}^{l_0}, \mathbb{R}^{l_L}) \quad \text{and} \quad (\mathcal{R}_a(\Phi))(x_0) = W_L x_{L-1} + B_L \tag{4.4}$$

(cf. Definitions 4.1 *and* 4.2 *and Fig.* 4.1*).*

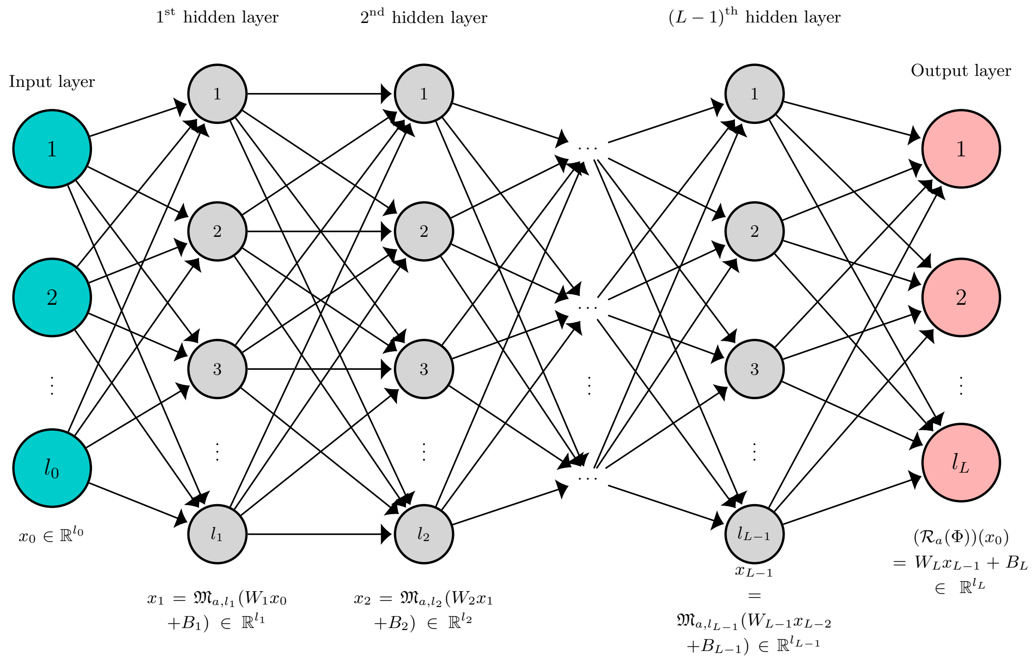

Fig. 4.1. Graphical illustration for the realization function and the architecture of an ANN $\Phi = ((W_1, B_1), \ldots, (W_L, B_L)) \in (\times_{k=1}^{L}(\mathbb{R}^{l_k \times l_{k-1}} \times \mathbb{R}^{l_k})) \subseteq \mathbf{N}$ (see Definition 4.1) where $L \in \mathbb{N}$ describes the number of affine linear transformations, where $l_0, l_1, \ldots, l_L \in \mathbb{N}$ describe the dimensions of the layers of the ANN, and where the function $a\colon \mathbb{R} \to \mathbb{R}$ represents the activation function (see Definition 4.3).

Definition 4.4 (Rectifier Function). *We denote by $\mathfrak{r}\colon \mathbb{R} \to \mathbb{R}$ the function which satisfies for all $x \in \mathbb{R}$ that $\mathfrak{r}(x) = \max\{x, 0\}$.*

4.3. Compositions of ANNs

Definition 4.5 (Standard Compositions of ANNs). *We denote by $(\cdot) \bullet (\cdot)\colon \{(\Phi_1, \Phi_2) \in \mathbf{N} \times \mathbf{N}\colon \mathcal{I}(\Phi_1) = \mathcal{O}(\Phi_2)\} \to \mathbf{N}$ the function which satisfies for all $L, \mathfrak{L} \in \mathbb{N}$, $l_0, l_1, \ldots, l_L, \mathfrak{l}_0, \mathfrak{l}_1, \ldots, \mathfrak{l}_{\mathfrak{L}} \in \mathbb{N}$, $\Phi_1 = ((W_1, B_1), (W_2, B_2), \ldots, (W_L, B_L)) \in (\times_{k=1}^{L}(\mathbb{R}^{l_k \times l_{k-1}} \times \mathbb{R}^{l_k}))$, $\Phi_2 = ((\mathcal{W}_1, \mathfrak{B}_1),,(\mathcal{W}_2, \mathfrak{B}_2), \ldots, (\mathcal{W}_{\mathfrak{L}}, \mathfrak{B}_{\mathfrak{L}})) \in (\times_{k=1}^{\mathfrak{L}}(\mathbb{R}^{\mathfrak{l}_k \times \mathfrak{l}_{k-1}} \times \mathbb{R}^{\mathfrak{l}_k}))$ with $l_0 = \mathcal{I}(\Phi_1) = \mathcal{O}(\Phi_2) = \mathfrak{l}_{\mathfrak{L}}$ that*

$$\Phi_1 \bullet \Phi_2 = \begin{cases} \big((\mathcal{W}_1, \mathfrak{B}_1), (\mathcal{W}_2, \mathfrak{B}_2), \ldots, (\mathcal{W}_{\mathfrak{L}-1}, \mathfrak{B}_{\mathfrak{L}-1}), (W_1\mathcal{W}_{\mathfrak{L}}, W_1\mathfrak{B}_{\mathfrak{L}} + B_1), \\ \qquad (W_2, B_2), (W_3, B_3), \ldots, (W_L, B_L)\big), & \text{if } L > 1 < \mathfrak{L}, \\ \big((W_1\mathcal{W}_1, W_1\mathfrak{B}_1 + B_1), (W_2, B_2), (W_3, B_3), \ldots, (W_L, B_L)\big), & \text{if } L > 1 = \mathfrak{L}, \\ \big((\mathcal{W}_1, \mathfrak{B}_1), (\mathcal{W}_2, \mathfrak{B}_2), \ldots, (\mathcal{W}_{\mathfrak{L}-1}, \mathfrak{B}_{\mathfrak{L}-1}), (W_1\mathcal{W}_{\mathfrak{L}}, W_1\mathfrak{B}_{\mathfrak{L}} + B_1)\big), & \text{if } L = 1 < \mathfrak{L}, \\ (W_1\mathcal{W}_1, W_1\mathfrak{B}_1 + B_1), & \text{if } L = 1 = \mathfrak{L}, \end{cases} \tag{4.5}$$

(cf. Definition 4.1).

Lemma 4.1. *Let $\Phi_1, \Phi_2, \Phi_3 \in \mathbf{N}$ satisfy that $\mathcal{I}(\Phi_1) = \mathcal{O}(\Phi_2)$ and $\mathcal{I}(\Phi_2) = \mathcal{O}(\Phi_3)$ (cf. Definition 4.1). Then*

$$(\Phi_1 \bullet \Phi_2) \bullet \Phi_3 = \Phi_1 \bullet (\Phi_2 \bullet \Phi_3) \tag{4.6}$$

(cf. Definition 4.5).

Definition 4.6 (Compositions of ANNs Involving Artificial Identities). *Let $\Psi \in \mathbf{N}$. Then we denote by*

$$(\cdot) \odot_\Psi (\cdot) \colon \{(\Phi_1, \Phi_2) \in \mathbf{N} \times \mathbf{N} \colon \mathcal{I}(\Phi_1) = \mathcal{O}(\Psi) \quad \text{and} \quad \mathcal{O}(\Phi_2) = \mathcal{I}(\Psi)\} \to \mathbf{N} \tag{4.7}$$

the function which satisfies for all $\Phi_1, \Phi_2 \in \mathbf{N}$ with $\mathcal{I}(\Phi_1) = \mathcal{O}(\Psi)$ and $\mathcal{O}(\Phi_2) = \mathcal{I}(\Psi)$ that

$$\Phi_1 \odot_\Psi \Phi_2 = \Phi_1 \bullet (\Psi \bullet \Phi_2) = (\Phi_1 \bullet \Psi) \bullet \Phi_2 \tag{4.8}$$

(cf. Definitions 4.1 and 4.5 and Lemma 4.1).

4.4. Deep ANN approximations for Monte Carlo Euler approximations

Proposition 4.1. *Let $N, d \in \mathbb{N}$, $c, C \in [0, \infty)$, $T, \mathfrak{D} \in (0, \infty)$, $q \in (2, \infty)$, $\varepsilon \in (0, 1]$, $(\tau_n)_{n \in \{0,1,\ldots,N\}} \subseteq \mathbb{R}$ satisfy for all $n \in \{0, 1, \ldots, N\}$ that $\tau_n = \frac{nT}{N}$ and*

$$\mathfrak{D} = \left[\frac{720q}{(q-2)}\right] \left[\log_2(\varepsilon^{-1}) + q + 1\right] - 504, \tag{4.9}$$

let $\Phi \in \mathbf{N}$ satisfy for all $x \in \mathbb{R}^d$ that $\mathcal{I}(\Phi) = \mathcal{O}(\Phi) = d$ and $\|(\mathcal{R}_\mathfrak{r}(\Phi))(x)\| \leq C + c\|x\|$, let $Y = (Y_t^{x,y})_{(t,x,y) \in [0,T] \times \mathbb{R}^d \times (\mathbb{R}^d)^N} \colon [0, T] \times \mathbb{R}^d \times (\mathbb{R}^d)^N \to \mathbb{R}^d$ satisfy for all $n \in \{0, 1, \ldots, N-1\}$, $t \in [\tau_n, \tau_{n+1}]$, $x \in \mathbb{R}^d$, $y = (y_1, y_2, \ldots, y_N) \in (\mathbb{R}^d)^N$ that $Y_0^{x,y} = x$ and

$$Y_t^{x,y} = Y_{\tau_n}^{x,y} + \left(\frac{tN}{T} - n\right) \left[\frac{T}{N} (\mathcal{R}_\mathfrak{r}(\Phi))(Y_{\tau_n}^{x,y}) + y_{n+1}\right], \tag{4.10}$$

and let $g_n \colon \mathbb{R}^d \times (\mathbb{R}^d)^N \to [0, \infty)$, $n \in \{0, 1, \ldots, N\}$, satisfy for all $n \in \{0, 1, \ldots, N\}$, $x \in \mathbb{R}^d$, $y = (y_1, y_2, \ldots, y_N) \in (\mathbb{R}^d)^N$ that

$$g_n(x, y) = \left[\|x\| + C\tau_n + \max_{m \in \{0,1,\ldots,n\}} \left\| \sum_{k=1}^m y_k \right\| \right] \exp(c\tau_n) \tag{4.11}$$

(cf. Definitions 2.1,4.1,4.3, and 4.4). Then there exist $\Psi_y \in \mathbf{N}$, $y \in (\mathbb{R}^d)^N$, such that

(i) it holds for all $y \in (\mathbb{R}^d)^N$ that $\mathcal{R}_\mathfrak{r}(\Psi_y) \in C(\mathbb{R}^{d+1}, \mathbb{R}^d)$,

(ii) it holds for all $n \in \{0, 1, \ldots, N-1\}$, $t \in [\tau_n, \tau_{n+1}]$, $x \in \mathbb{R}^d$, $y \in (\mathbb{R}^d)^N$ that

$$\|Y_t^{x,y} - (\mathcal{R}_\mathfrak{r}(\Psi_y))(t, x)\| \leq \varepsilon\left[2\sqrt{d} + (g_n(x, y))^q + (g_{n+1}(x, y))^q\right], \tag{4.12}$$

(iii) it holds for all $n \in \{0, 1, \ldots, N-1\}$, $t \in [\tau_n, \tau_{n+1}]$, $x \in \mathbb{R}^d$, $y \in (\mathbb{R}^d)^N$ that

$$\|(\mathcal{R}_\mathfrak{r}(\Psi_y))(t, x)\| \leq 6\sqrt{d} + 2\left[(g_n(x, y))^2 + (g_{n+1}(x, y))^2\right], \tag{4.13}$$

(iv) it holds for all $y \in (\mathbb{R}^d)^N$ that

$$\mathcal{P}(\Psi_y) \leq \frac{9}{2} N^6 d^{16} \left[2(\mathcal{L}(\Phi) - 1) + \mathfrak{D} + (24 + 6\mathcal{L}(\Phi) + [4 + \mathcal{P}(\Phi)]^2)^2\right]^2, \tag{4.14}$$

(v) it holds for all $t \in [0, T]$, $x \in \mathbb{R}^d$ that $[(\mathbb{R}^d)^N \ni y \mapsto (\mathcal{R}_\mathfrak{r}(\Psi_y))(t, x) \in \mathbb{R}^d] \in C((\mathbb{R}^d)^N, \mathbb{R}^d)$,

(vi) it holds for all $n \in \{0, 1, \dots, N\}$, $t \in [0, \tau_n]$, $x \in \mathbb{R}^d$, $y = (y_1, y_2, \dots, y_N)$, $z = (z_1, z_2, \dots, z_N) \in (\mathbb{R}^d)^N$ *with* $\forall\, k \in \mathbb{N} \cap [0, n]\colon y_k = z_k$ *that*

$$(\mathcal{R}_{\mathfrak{r}}(\Psi_y))(t, x) = (\mathcal{R}_{\mathfrak{r}}(\Psi_z))(t, x). \tag{4.15}$$

Proof. Throughout this proof let $\Psi_y \in \mathbf{N}$, $y \in (\mathbb{R}^d)^N$, satisfy that

(I) it holds for all $y \in (\mathbb{R}^d)^N$ that $\mathcal{R}_{\mathfrak{r}}(\Psi_y) \in C(\mathbb{R}^{d+1}, \mathbb{R}^d)$,

(II) it holds for all $n \in \{0, 1, \dots, N-1\}$, $t \in [\tau_n, \tau_{n+1}]$, $x \in \mathbb{R}^d$, $y \in (\mathbb{R}^d)^N$ that

$$\|Y_t^{x,y} - (\mathcal{R}_{\mathfrak{r}}(\Psi_y))(t, x)\| \leq \varepsilon\big(2\sqrt{d} + \|Y_{\tau_n}^{x,y}\|^q + \|Y_{\tau_{n+1}}^{x,y}\|^q\big), \tag{4.16}$$

(III) it holds for all $n \in \{0, 1, \dots, N-1\}$, $t \in [\tau_n, \tau_{n+1}]$, $x \in \mathbb{R}^d$, $y \in (\mathbb{R}^d)^N$ that

$$\|(\mathcal{R}_{\mathfrak{r}}(\Psi_y))(t, x)\| \leq 6\sqrt{d} + 2\big(\|Y_{\tau_n}^{x,y}\|^2 + \|Y_{\tau_{n+1}}^{x,y}\|^2\big), \tag{4.17}$$

(IV) it holds for all $y \in (\mathbb{R}^d)^N$ that

$$\begin{aligned}\mathcal{P}(\Psi_y) \leq \frac{1}{2}\Big[&6d^2N^2(\mathcal{L}(\Phi) - 1)\\ &+ 3N\big[d^2\mathfrak{D} + (23 + 6N(\mathcal{L}(\Phi) - 1) + 7d^2 + N[4d^2 + \mathcal{P}(\Phi)]^2)^2\big]\Big]^2,\end{aligned} \tag{4.18}$$

(V) it holds for all $t \in [0, T]$, $x \in \mathbb{R}^d$ that $[(\mathbb{R}^d)^N \ni y \mapsto (\mathcal{R}_{\mathfrak{r}}(\Psi_y))(t, x) \in \mathbb{R}^d] \in C((\mathbb{R}^d)^N, \mathbb{R}^d)$,

(VI) it holds for all $n \in \{0, 1, \dots, N\}$, $t \in [0, \tau_n]$, $x \in \mathbb{R}^d$, $y = (y_1, y_2, \dots, y_N)$, $z = (z_1, z_2, \dots, z_N) \in (\mathbb{R}^d)^N$ with $\forall\, k \in \mathbb{N} \cap [0, n]\colon y_k = z_k$ that

$$(\mathcal{R}_{\mathfrak{r}}(\Psi_y))(t, x) = (\mathcal{R}_{\mathfrak{r}}(\Psi_z))(t, x) \tag{4.19}$$

(cf. Grohs *et al.* [26, Proposition 3.10] (applied with $N \leftarrow N$, $d \leftarrow d$, $a \leftarrow \mathfrak{r}$, $T \leftarrow T$, $t_0 \leftarrow \tau_0$, $t_1 \leftarrow \tau_1, \dots, t_N \leftarrow \tau_N$, $\mathfrak{D} \leftarrow \mathfrak{D}$, $\varepsilon \leftarrow \varepsilon$, $q \leftarrow q$, $Y \leftarrow Y$ in the notation of Grohs *et al.* [26, Proposition 3.10])). Note that (IV) ensures that for all $y \in (\mathbb{R}^d)^N$ it holds that

$$\begin{aligned}\mathcal{P}(\Psi_y) &\leq \frac{1}{2}\Big[6d^2N^2(\mathcal{L}(\Phi) - 1) + 3N\big[d^2\mathfrak{D} + (23 + 6N(\mathcal{L}(\Phi) - 1) + 7d^2 + Nd^4[4 + \mathcal{P}(\Phi)]^2)^2\big]\Big]^2\\ &\leq \frac{1}{2}\Big[6d^2N^2(\mathcal{L}(\Phi) - 1) + 3N\big[d^2\mathfrak{D} + N^2d^8(23 + 6(\mathcal{L}(\Phi) - 1) + 7 + [4 + \mathcal{P}(\Phi)]^2)^2\big]\Big]^2\\ &= \frac{1}{2}\Big[6d^2N^2(\mathcal{L}(\Phi) - 1) + 3N\big[d^2\mathfrak{D} + N^2d^8(24 + 6\mathcal{L}(\Phi) + [4 + \mathcal{P}(\Phi)]^2)^2\big]\Big]^2.\end{aligned} \tag{4.20}$$

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

$$\begin{aligned}\mathcal{P}(\Psi_y) &\leq \frac{1}{2}\Big[6d^2N^2(\mathcal{L}(\Phi) - 1) + 3N^3d^8\big[\mathfrak{D} + (24 + 6\mathcal{L}(\Phi) + [4 + \mathcal{P}(\Phi)]^2)^2\big]\Big]^2\\ &\leq \frac{9}{2}N^6d^{16}\big[2(\mathcal{L}(\Phi) - 1) + \mathfrak{D} + (24 + 6\mathcal{L}(\Phi) + [4 + \mathcal{P}(\Phi)]^2)^2\big]^2.\end{aligned} \tag{4.21}$$

In addition, observe that, e.g., Grohs *et al.* [26, Lemma 3.11] (applied with $N \leftarrow N$, $d \leftarrow d$, $c \leftarrow c$, $C \leftarrow C$, $A_1 \leftarrow \frac{T}{N}\operatorname{id}_{\mathbb{R}^d}$, $A_2 \leftarrow \frac{T}{N}\operatorname{id}_{\mathbb{R}^d}, \dots, A_N \leftarrow \frac{T}{N}\operatorname{id}_{\mathbb{R}^d}$, $\mu \leftarrow \mathcal{R}_{\mathfrak{r}}(\Phi)$, $Y_0 \leftarrow$

$(Y^{x,y}_{\tau_0})_{(x,y)\in\mathbb{R}^d\times(\mathbb{R}^d)^N}$, $Y_1 \leftarrow (Y^{x,y}_{\tau_1})_{(x,y)\in\mathbb{R}^d\times(\mathbb{R}^d)^N}, \ldots, Y_N \leftarrow (Y^{x,y}_{\tau_N})_{(x,y)\in\mathbb{R}^d\times(\mathbb{R}^d)^N}$ in the notation of Grohs *et al.* [26, Lemma 3.11]) and the hypothesis that $\forall\, n \in \{0,1,\ldots,N\}\colon \tau_n = \frac{nT}{N}$ demonstrate that for all $n \in \{0,1,\ldots,N\}$, $x \in \mathbb{R}^d$, $y = (y_1, y_2, \ldots, y_N) \in (\mathbb{R}^d)^N$ it holds that

$$\begin{aligned}\|Y^{x,y}_{\tau_n}\| &\leq \left[\|x\| + \frac{CnT}{N} + \max_{m\in\{0,1,\ldots,n\}}\left\|\sum_{k=1}^m y_k\right\|\right]\exp\left(\frac{cnT}{N}\right)\\ &= \left[\|x\| + C\tau_n + \max_{m\in\{0,1,\ldots,n\}}\left\|\sum_{k=1}^m y_k\right\|\right]\exp(c\tau_n) = g_n(x,y). \end{aligned} \tag{4.22}$$

Combining this with (II) and (III) ensures that for all $n \in \{0,1,\ldots,N-1\}$, $t \in [\tau_n, \tau_{n+1}]$, $x \in \mathbb{R}^d$, $y \in (\mathbb{R}^d)^N$ it holds that

$$\begin{aligned}\|Y^{x,y}_t - (\mathcal{R}_{\mathfrak{r}}(\Psi_y))(t,x)\| &\leq \varepsilon\big[2\sqrt{d} + \|Y^{x,y}_{\tau_n}\|^q + \|Y^{x,y}_{\tau_{n+1}}\|^q\big]\\ &\leq \varepsilon\big[2\sqrt{d} + (g_n(x,y))^q + (g_{n+1}(x,y))^q\big]\end{aligned} \tag{4.23}$$

and

$$\begin{aligned}\|(\mathcal{R}_{\mathfrak{r}}(\Psi_y))(t,x)\| &\leq 6\sqrt{d} + 2\big(\|Y^{x,y}_{\tau_n}\|^2 + \|Y^{x,y}_{\tau_{n+1}}\|^2\big)\\ &\leq 6\sqrt{d} + 2\big[(g_n(x,y))^2 + (g_{n+1}(x,y))^2\big].\end{aligned} \tag{4.24}$$

This, (I), (V), (VI), and (4.21) establish items (i)-(vi). The proof of Proposition 4.1 is thus completed. □

Proposition 4.2. *Let $M, N, d, \mathfrak{d} \in \mathbb{N}$, $\alpha, c, C, \mathfrak{C} \in [0,\infty)$, $T, \mathfrak{D} \in (0,\infty)$, $q \in (2,\infty)$, $\varepsilon \in (0,1]$, $\mathfrak{f}^1, \mathfrak{f}^0 \in \mathbf{N}$ satisfy that $\mathcal{I}(\mathfrak{f}^1) = \mathcal{O}(\mathfrak{f}^1) = \mathcal{I}(\mathfrak{f}^0) = d$, $\mathcal{O}(\mathfrak{f}^0) = \mathfrak{d}$, and*

$$\mathfrak{D} = \left\lceil\frac{720q}{q-2}\right\rceil\big[\log_2(\varepsilon^{-1}) + q + 1\big] - 504, \tag{4.25}$$

assume for all $x, y \in \mathbb{R}^d$ that $\|(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^1))(x)\| \leq C + c\|x\|$ and

$$\|(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0))(x) - (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0))(y)\| \leq \mathfrak{C}(1 + \|x\|^\alpha + \|y\|^\alpha)\|x-y\|, \tag{4.26}$$

let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $W^m = (W^m_n)_{n\in\{0,1,\ldots,N\}}\colon \{0,1,\ldots,N\}\times\Omega \to \mathbb{R}^d$, $m \in \{1,2,\ldots,M\}$, be stochastic processes which satisfy for all $m \in \{1,2,\ldots,M\}$ that $W^m_0 = 0$, let $Y^m = (Y^{m,x}_t(\omega))_{(t,x,\omega)\in[0,T]\times\mathbb{R}^d\times\Omega}\colon [0,T]\times\mathbb{R}^d\times\Omega \to \mathbb{R}^d$, $m \in \{1,2,\ldots,M\}$, satisfy for all $m \in \{1,2,\ldots,M\}$, $x \in \mathbb{R}^d$, $n \in \{0,1,\ldots,N-1\}$, $t \in [\frac{nT}{N}, \frac{(n+1)T}{N}]$ that $Y^{m,x}_0 = x$ and

$$Y^{m,x}_t = Y^{m,x}_{\frac{nT}{N}} + \left(\frac{tN}{T} - n\right)\left[\frac{T}{N}(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^1))(Y^{m,x}_{\frac{nT}{N}}) + W^m_{n+1} - W^m_n\right], \tag{4.27}$$

and let $h_{m,r}\colon \mathbb{R}^d\times\Omega \to [0,\infty)$, $m \in \{1,2,\ldots,M\}$, $r \in (0,\infty)$, satisfy for all $m \in \{1,2,\ldots,M\}$, $r \in (0,\infty)$, $x \in \mathbb{R}^d$ that

$$h_{m,r}(x) = 1 + \left[\|x\| + CT + \max_{n\in\{0,1,\ldots,N\}}\|W^m_n\|\right]^r \exp(rcT) \tag{4.28}$$

(cf. Definitions 2.1, 4.1, 4.3, and 4.4). Then there exists $(\Psi_\omega)_{\omega\in\Omega} \subseteq \mathbf{N}$ such that

(i) it holds for all $\omega \in \Omega$ that $\mathcal{R}_{\mathfrak{r}}(\Psi_\omega) \in C(\mathbb{R}^{d+1}, \mathbb{R}^{\mathfrak{d}})$,

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

$$\left\| (\mathcal{R}_{\mathfrak{r}}(\Psi_\omega))(t,x) - \frac{1}{M}\left[\sum_{m=1}^{M} (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0))(Y_t^{m,x}(\omega))\right] \right\| \leq \frac{2\varepsilon\mathfrak{C}\sqrt{d}}{M}\left[\sum_{m=1}^{M}\left[1 + 2d^{\alpha/2}6^\alpha |h_{m,2}(x,\omega)|^\alpha\right] h_{m,q}(x,\omega)\right], \tag{4.29}$$

(iii) it holds for all $\omega \in \Omega$ that

$$\mathcal{P}(\Psi_\omega) \leq 2M^2\mathcal{P}(\mathfrak{f}^0) + 9M^2N^6d^{16}\big[2\mathcal{L}(\mathfrak{f}^1) + \mathfrak{D} + (24 + 6\mathcal{L}(\mathfrak{f}^1) + [4 + \mathcal{P}(\mathfrak{f}^1)]^2)^2\big]^2, \tag{4.30}$$

(iv) it holds for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that $\Omega \ni \omega \mapsto (\mathcal{R}_{\mathfrak{r}}(\Psi_\omega))(t,x) \in \mathbb{R}^{\mathfrak{d}}$ is measurable.

Proof. Throughout this proof let $\tau_0, \tau_1, \ldots, \tau_N \in \mathbb{R}$ satisfy for all $n \in \{0,1,\ldots,N\}$ that $\tau_n = \frac{nT}{N}$, let $g_m \colon \mathbb{R}^d \times \Omega \to [0,\infty)$, $m \in \{1,2,\ldots,M\}$, satisfy for all $m \in \{1,2,\ldots,M\}$, $x \in \mathbb{R}^d$ that

$$g_m(x) = \left[\|x\| + CT + \max_{n\in\{0,1,\ldots,N\}}\left\|\sum_{l=1}^{n}\big(W_l^m - W_{l-1}^m\big)\right\|\right]\exp(cT), \tag{4.31}$$

let $(\Psi_{\omega,m})_{\omega\in\Omega, m\in\{1,2,\ldots,M\}} \subseteq \mathbf{N}$ satisfy that

(I) it holds for all $m \in \{1,2,\ldots,M\}$, $\omega \in \Omega$ that $\mathcal{R}_{\mathfrak{r}}(\Psi_{\omega,m}) \in C(\mathbb{R}^{d+1},\mathbb{R}^d)$,

(II) it holds for all $m \in \{1,2,\ldots,M\}$, $t \in [0,T]$, $x \in \mathbb{R}^d$, $\omega \in \Omega$ that

$$\|Y_t^{m,x}(\omega) - (\mathcal{R}_{\mathfrak{r}}(\Psi_{\omega,m}))(t,x)\| \leq 2\varepsilon\sqrt{d}\,\big[1 + (g_m(x,\omega))^q\big], \tag{4.32}$$

(III) it holds for all $m \in \{1,2,\ldots,M\}$, $t \in [0,T]$, $x \in \mathbb{R}^d$, $\omega \in \Omega$ that

$$\|(\mathcal{R}_{\mathfrak{r}}(\Psi_{\omega,m}))(t,x)\| \leq 6\sqrt{d}\,\big[1 + (g_m(x,\omega))^2\big], \tag{4.33}$$

(IV) it holds for all $m \in \{1,2,\ldots,M\}$, $\omega \in \Omega$ that

$$\mathcal{P}(\Psi_{\omega,m}) \leq \frac{9}{2}N^6d^{16}\big[2\mathcal{L}(\mathfrak{f}^1) + \mathfrak{D} + (24 + 6\mathcal{L}(\mathfrak{f}^1) + [4 + \mathcal{P}(\mathfrak{f}^1)]^2)^2\big]^2, \tag{4.34}$$

(V) it holds for all $m \in \{1,2,\ldots,M\}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that $\Omega \ni \omega \mapsto (\mathcal{R}_{\mathfrak{r}}(\Psi_{\omega,m}))(t,x) \in \mathbb{R}^d$ is measurable

(cf. Proposition 4.1 (applied with $N \leftarrow N$, $d \leftarrow d$, $c \leftarrow c$, $C \leftarrow C$, $T \leftarrow T$, $\mathfrak{D} \leftarrow \mathfrak{D}$, $q \leftarrow q$, $\varepsilon \leftarrow \varepsilon$, $\tau_0 \leftarrow \tau_0$, $\tau_1 \leftarrow \tau_1, \ldots, \tau_N \leftarrow \tau_N$, $\Phi \leftarrow \mathfrak{f}^1$ in the notation of Proposition 4.1)), let $\mathbb{I} \in \mathbf{N}$ satisfy for all $x \in \mathbb{R}^d$ that $\mathcal{R}_{\mathfrak{r}}(\mathbb{I}) \in C(\mathbb{R}^d,\mathbb{R}^d)$, $\mathcal{D}(\mathbb{I}) = (d, 2d, d)$, and $(\mathcal{R}_{\mathfrak{r}}(\mathbb{I}))(x) = x$ (cf. [39, Lemma 5.5]), let $(\psi_{\omega,m})_{\omega\in\Omega, m\in\{1,2,\ldots,M\}} \subseteq \mathbf{N}$ satisfy for all $m \in \{1,2,\ldots,M\}$, $\omega \in \Omega$ that $\psi_{\omega,m} = \mathfrak{f}^0 \odot_{\mathbb{I}} \Psi_{\omega,m}$ (cf. Definition 4.6), and let $(\Phi_\omega)_{\omega\in\Omega} \subseteq \mathbf{N}$ satisfy that

(A) it holds for all $\omega \in \Omega$ that $\mathcal{R}_{\mathfrak{r}}(\Phi_\omega) \in C(\mathbb{R}^{\mathcal{I}(\psi_{\omega,1})}, \mathbb{R}^{\mathcal{O}(\psi_{\omega,1})})$,

(B) it holds for all $\omega \in \Omega$ that $\mathcal{P}(\Phi_\omega) \leq M^2\mathcal{P}(\psi_{\omega,1})$,

(C) it holds for all $t \in \mathbb{R}$, $x \in \mathbb{R}^d$, $\omega \in \Omega$ that

$$(\mathcal{R}_{\mathfrak{r}}(\Phi_\omega))(t,x) = \frac{1}{M} \sum_{m=1}^{M} (\mathcal{R}_{\mathfrak{r}}(\psi_{\omega,m}))(t,x) \tag{4.35}$$

(cf. Grohs *et al.* [26, Proposition 2.25]). Note that (B), (IV), Grohs *et al.* [26, Proposition 2.16 item (iii)], and the fact that $\mathcal{D}(\mathbb{I}) = (d, 2d, d)$ demonstrate that for all $\omega \in \Omega$ it holds that

$$\begin{aligned}\mathcal{P}(\Phi_\omega) &\le M^2 \mathcal{P}(\psi_{\omega,1}) \le 2M^2[\mathcal{P}(\mathfrak{f}^0) + \mathcal{P}(\Psi_{\omega,1})] \\ &\le 2M^2\mathcal{P}(\mathfrak{f}^0) + 9M^2N^6d^{16}\big[2\mathcal{L}(\mathfrak{f}^1) + \mathfrak{D} + (24 + 6\mathcal{L}(\mathfrak{f}^1) + [4 + \mathcal{P}(\mathfrak{f}^1)]^2)^2\big]^2.\end{aligned} \tag{4.36}$$

Moreover, observe that (A), (C), and Grohs *et al.* [26, Proposition 2.16 item (iv)] imply that for all $t \in \mathbb{R}$, $x \in \mathbb{R}^d$, $\omega \in \Omega$ it holds that $\mathcal{R}_{\mathfrak{r}}(\Phi_\omega) \in C(\mathbb{R}^{d+1}, \mathbb{R}^{\mathfrak{d}})$ and

$$(\mathcal{R}_{\mathfrak{r}}(\Phi_\omega))(t,x) = \frac{1}{M} \sum_{m=1}^{M} (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0))((\mathcal{R}_{\mathfrak{r}}(\Psi_{\omega,m}))(t,x)). \tag{4.37}$$

Next note that the fact that $\forall\, m \in \{1, 2, \dots, M\}\colon W_0^m = 0$ ensures that for all $n \in \{1, 2, \dots, N\}$, $m \in \{1, 2, \dots, M\}$ it holds that

$$\sum_{l=1}^{n} \big(W_l^m - W_{l-1}^m\big) = W_n^m - W_0^m = W_n^m. \tag{4.38}$$

Combining this with (4.31) proves that for all $m \in \{1, 2, \dots, M\}$, $x \in \mathbb{R}^d$ it holds that

$$g_m(x) = \left[\|x\| + CT + \max_{n \in \{0,1,\dots,N\}} \|W_n^m\|\right] \exp(cT). \tag{4.39}$$

In addition, observe that (4.27) and the fact that $\forall\, n \in \{0, 1, \dots, N\}\colon \tau_n = \frac{nT}{N}$ assure that for all $m \in \{1, 2, \dots, M\}$, $n \in \{0, 1, \dots, N-1\}$, $x \in \mathbb{R}^d$ it holds that

$$Y_{\tau_{n+1}}^{m,x} = Y_{\tau_n}^{m,x} + \frac{T}{N}(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^1))(Y_{\tau_n}^{m,x}) + W_{n+1}^m - W_n^m. \tag{4.40}$$

Induction and (4.38) hence show that for all $m \in \{1, 2, \dots, M\}$, $n \in \{0, 1, \dots, N-1\}$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}Y_{\tau_{n+1}}^{m,x} &= Y_{\tau_0}^{m,x} + \frac{T}{N}\left[\sum_{l=0}^{n} (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^1))(Y_{\tau_l}^{m,x})\right] + \left[\sum_{l=0}^{n} (W_{l+1}^m - W_l^m)\right] \\ &= x + \frac{T}{N}\left[\sum_{l=0}^{n} (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^1))(Y_{\tau_l}^{m,x})\right] + W_{n+1}^m.\end{aligned} \tag{4.41}$$

This and the assumption that $\forall\, x \in \mathbb{R}^d\colon \|(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^1))(x)\| \le C + c\|x\|$ establish that for all $m \in \{1, 2, \dots, M\}$, $n \in \{0, 1, \dots, N\}$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}\|Y_{\tau_n}^{m,x}\| &\le \|x\| + \frac{T}{N}\left[\sum_{l=0}^{n-1} \|(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^1))(Y_{\tau_l}^{m,x})\|\right] + \|W_n^m\| \\ &\le \|x\| + \frac{T}{N}\left[\sum_{l=0}^{n-1} (C + c\|Y_{\tau_l}^{m,x}\|)\right] + \|W_n^m\| \\ &\le \|x\| + CT + \left[\max_{k \in \{0,1,\dots,N\}} \|W_k^m\|\right] + \frac{cT}{N}\left[\sum_{l=0}^{n-1} \|Y_{\tau_l}^{m,x}\|\right].\end{aligned} \tag{4.42}$$

The time-discrete Gronwall inequality, e.g., in Hutzenthaler *et al.* [37, Lemma 2.1] (applied with $N \leftarrow N$, $\alpha \leftarrow (\|x\| + CT + \max_{k\in\{0,1,\dots,N\}} \|W_k^m(\omega)\|)$, $\beta_0 \leftarrow \frac{cT}{N}$, $\beta_1 \leftarrow \frac{cT}{N}, \dots, \beta_{N-1} \leftarrow \frac{cT}{N}$, $\epsilon_0 \leftarrow \|x\|$, $\epsilon_1 \leftarrow \|Y_{\tau_1}^{m,x}(\omega)\|, \dots, \epsilon_N \leftarrow \|Y_{\tau_N}^{m,x}(\omega)\|$ for $m \in \{1,2,\dots,M\}$, $x \in \mathbb{R}^d$, $\omega \in \Omega$ in the notation of Hutzenthaler *et al.* [37, Lemma 2.1]) and (4.39) hence demonstrate that for all $m \in \{1,2,\dots,M\}$, $n \in \{0,1,\dots,N\}$, $x \in \mathbb{R}^d$ it holds that

$$\|Y_{\tau_n}^{m,x}\| \leq \left[\|x\| + CT + \max_{k\in\{0,1,\dots,N\}} \|W_k^m\|\right] \exp(cT) = g_m(x). \tag{4.43}$$

In addition, note that (4.27) and the fact that $\forall\, n \in \{0,1,\dots,N-1\}, t \in [\tau_n, \tau_{n+1}] \colon \frac{t-\tau_n}{\tau_{n+1}-\tau_n} = \frac{tN}{T} - n$ ensure that for all $m \in \{1,2,\dots,M\}$, $n \in \{0,1,\dots,N-1\}$, $t \in [\tau_n, \tau_{n+1}]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
&Y_{\tau_{n+1}}^{m,x}\left[\frac{t-\tau_n}{\tau_{n+1}-\tau_n}\right] + Y_{\tau_n}^{m,x}\left[1 - \frac{t-\tau_n}{\tau_{n+1}-\tau_n}\right] \\
&= \left(Y_{\tau_n}^{m,x} + \left[\frac{\tau_{n+1}N}{T} - n\right]\left[\frac{T}{N}(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^1))(Y_{\tau_n}^{m,x}) + W_{n+1}^m - W_n^m\right]\right)\frac{t-\tau_n}{\tau_{n+1}-\tau_n} \\
&\quad + Y_{\tau_n}^{m,x}\left[1 - \frac{t-\tau_n}{\tau_{n+1}-\tau_n}\right] \\
&= \left(Y_{\tau_n}^{m,x} + \left[\frac{T}{N}(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^1))(Y_{\tau_n}^{m,x}) + W_{n+1}^m - W_n^m\right]\right)\frac{t-\tau_n}{\tau_{n+1}-\tau_n} + Y_{\tau_n}^{m,x}\left[1 - \frac{t-\tau_n}{\tau_{n+1}-\tau_n}\right] \\
&= Y_{\tau_n}^{m,x} + \left[\frac{T}{N}(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^1))(Y_{\tau_n}^{m,x}) + W_{n+1}^m - W_n^m\right]\frac{t-\tau_n}{\tau_{n+1}-\tau_n} \\
&= Y_t^{m,x}.
\end{aligned} \tag{4.44}$$

Combining this with (4.43) implies that for all $m \in \{1,2,\dots,M\}$, $n \in \{0,1,\dots,N-1\}$, $t \in [\tau_n, \tau_{n+1}]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
\|Y_t^{m,x}\| &\leq \|Y_{\tau_n}^{m,x}\|\left[1 - \frac{t-\tau_n}{\tau_{n+1}-\tau_n}\right] + \|Y_{\tau_{n+1}}^{m,x}\|\left[\frac{t-\tau_n}{\tau_{n+1}-\tau_n}\right] \\
&\leq \max\left\{\|Y_{\tau_n}^{m,x}\|, \|Y_{\tau_{n+1}}^{m,x}\|\right\} \leq g_m(x) \leq 1 + |g_m(x)|^2.
\end{aligned} \tag{4.45}$$

This, (4.26), (II), and (III) ensure that for all $m \in \{1,2,\dots,M\}$, $t \in [0,T]$, $x \in \mathbb{R}^d$, $\omega \in \Omega$ it holds that

$$\begin{aligned}
&\left\|(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0))((\mathcal{R}_{\mathfrak{r}}(\Psi_{\omega,m}))(t,x)) - (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0))(Y_t^{m,x}(\omega))\right\| \\
&\leq \mathfrak{C}\big(1 + \|(\mathcal{R}_{\mathfrak{r}}(\Psi_{\omega,m}))(t,x)\|^\alpha + \|Y_t^{m,x}(\omega)\|^\alpha\big)\|(\mathcal{R}_{\mathfrak{r}}(\Psi_{\omega,m}))(t,x) - Y_t^{m,x}(\omega)\| \\
&\leq \mathfrak{C}\left[1 + 6^\alpha d^{\alpha/2}(1 + |g_m(x,\omega)|^2)^\alpha + (1 + |g_m(x,\omega)|^2)^\alpha\right] 2\varepsilon\sqrt{d}\,(1 + |g_m(x,\omega)|^q)\,.
\end{aligned} \tag{4.46}$$

Combining this, (4.28), and (4.39) demonstrates that for all $m \in \{1,2,\dots,M\}$, $t \in [0,T]$, $x \in \mathbb{R}^d$, $\omega \in \Omega$ it holds that

$$\begin{aligned}
&\left\|(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0))((\mathcal{R}_{\mathfrak{r}}(\Psi_{\omega,m}))(t,x)) - (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0))(Y_t^{m,x}(\omega))\right\| \\
&\leq 2\varepsilon\mathfrak{C}\sqrt{d}\left[1 + 2d^{\alpha/2}6^\alpha(1 + |g_m(x,\omega)|^2)^\alpha\right](1 + |g_m(x,\omega)|^q) \\
&= 2\varepsilon\mathfrak{C}\sqrt{d}\left[1 + 2d^{\alpha/2}6^\alpha|h_{m,2}(x,\omega)|^\alpha\right] h_{m,q}(x,\omega).
\end{aligned} \tag{4.47}$$

This and (4.37) show that for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $\omega \in \Omega$ it holds that

$$\begin{aligned}
&\left\| (\mathcal{R}_{\mathfrak{r}}(\Phi_\omega))(t,x) - \frac{1}{M}\left[\sum_{m=1}^{M} (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0))(Y_t^{m,x}(\omega))\right]\right\| \\
&= \left\| \frac{1}{M}\left[\sum_{m=1}^{M} (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0))((\mathcal{R}_{\mathfrak{r}}(\Psi_{\omega,m}))(t,x))\right] - \frac{1}{M}\left[\sum_{m=1}^{M} (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0))(Y_t^{m,x}(\omega))\right]\right\| \\
&\leq \frac{1}{M}\left[\sum_{m=1}^{M} \left\|(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0))((\mathcal{R}_{\mathfrak{r}}(\Psi_{\omega,m}))(t,x)) - (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0))(Y_t^{m,x}(\omega))\right\|\right] \\
&\leq \frac{2\varepsilon\mathfrak{C}\sqrt{d}}{M}\left[\sum_{m=1}^{M}\left[1 + 2d^{\alpha/2}6^\alpha |h_{m,2}(x,\omega)|^\alpha\right] h_{m,q}(x,\omega)\right].
\end{aligned} \tag{4.48}$$

Moreover, observe that (V), the fact that $\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0)$ is continuous, and (4.37) ensure that for all $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that $\Omega \ni \omega \mapsto (\mathcal{R}_{\mathfrak{r}}(\Phi_\omega))(t,x) \in \mathbb{R}^{\mathfrak{d}}$ is measurable. Combining this with (4.36), (4.48), and the fact that $\forall\, \omega \in \Omega\colon \mathcal{R}_{\mathfrak{r}}(\Phi_\omega) \in C(\mathbb{R}^{d+1}, \mathbb{R}^{\mathfrak{d}})$ establishes items (i)-(iv). This completes the proof of Proposition 4.2. ☐

Lemma 4.2. *Let $N, d \in \mathbb{N}$, $T \in (0,\infty)$, $\alpha, c, C, \mathcal{C} \in [0,\infty)$, $p \in (1,\infty)$, let $\nu\colon \mathcal{B}([0,T]\times\mathbb{R}^d) \to [0,\infty)$ be a finite measure which satisfies that*

$$\mathcal{C} = \max\left\{\left[\int_{[0,T]\times\mathbb{R}^d} \|x\|^{\max\{4p\alpha,6p\}}\,\nu(dt,dx)\right]^{1/\max\{4p\alpha,6p\}}, 1\right\}, \tag{4.49}$$

let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_n)_{n\in\{0,1,\dots,N\}})$ be a filtered probability space, let $\mathcal{M} = (\mathcal{M}_n)_{n\in\{0,1,\dots,N\}}\colon \{0,1,\dots,N\}\times\Omega \to [0,\infty)$ be an $(\mathbb{F}_n)_{n\in\{0,1,\dots,N\}}$-submartingale which satisfies for all $n \in \{0,1,\dots,N\}$ that $\mathbb{E}[|\mathcal{M}_n|^{\max\{4p\alpha,6p\}}] < \infty$, and let $h_r\colon \mathbb{R}^d\times\Omega \to [0,\infty)$, $r \in (0,\infty)$, satisfy for all $r \in (0,\infty)$, $x \in \mathbb{R}^d$ that

$$h_r(x) = 1 + e^{rc}\left[\|x\| + C + \max_{n\in\{0,1,\dots,N\}} \mathcal{M}_n\right]^r \tag{4.50}$$

(cf. Definition 2.1). Then

(i) it holds for all $q, r \in (0,\infty)$ with $1 < qr \leq \max\{4\alpha p, 6p\}$ that

$$\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[|h_r(x)|^q\big]\,\nu(dt,dx)\right]^{1/q} \\
&\leq 2e^{rc}\max\left\{2^{(1/q)-1}, 1\right\}\left[\mathcal{C} + C + \frac{qr}{qr-1}\left|\mathbb{E}\big[|\mathcal{M}_N|^{qr}\big]\right|^{1/qr}\right]^r \\
&\quad\times \max\left\{1, [\nu([0,T]\times\mathbb{R}^d)]^{1/q}\right\},
\end{aligned} \tag{4.51}$$

(ii) it holds that

$$\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\Big[\big|[h_2(x)]^\alpha h_3(x)\big|^p\Big]\,\nu(dt,dx)\right]^{1/p} \\
&\leq \left[\mathcal{C} + C + \frac{\max\{4p\alpha,6p\}}{\max\{4p\alpha,6p\}-1}\left|\mathbb{E}\big[|\mathcal{M}_N|^{\max\{4p\alpha,6p\}}\big]\right|^{1/\max\{4p\alpha,6p\}}\right]^{2\alpha+3} \\
&\quad\times 2^{\alpha+1}e^{(2\alpha+3)c}\max\left\{1, [\nu([0,T]\times\mathbb{R}^d)]^{1/p}\right\}.
\end{aligned} \tag{4.52}$$

Proof. Throughout this proof let $\mathscr{C}_q \in \mathbb{R}$, $q \in (0, \max\{4p\alpha, 6p\}]$, satisfy for all $q \in (0, \max\{4p\alpha, 6p\}]$ that

$$\mathscr{C}_q = \max\left\{ \left[\int_{[0,T]\times\mathbb{R}^d} \|x\|^q \, \nu(dt, dx) \right]^{1/q}, 1 \right\} \tag{4.53}$$

and let $g \colon \mathbb{R}^d \times \Omega \to [0, \infty)$ satisfy for all $x \in \mathbb{R}^d$ that

$$g(x) = \left[\|x\| + C + \max_{n\in\{0,1,\ldots,N\}} \mathcal{M}_n \right]. \tag{4.54}$$

Observe that Doob's inequality (cf., e.g., Klenke [43, Theorem 11.2]), Hölder's inequality, the hypothesis that $\mathcal{M}$ is a submartingale, the hypothesis that $\forall\, n \in \{0, 1, \ldots, N\} \colon \mathcal{M}_n \geq 0$, and the hypothesis that $\forall\, n \in \{0, 1, \ldots, N\} \colon \mathbb{E}[|\mathcal{M}_n|^{\max\{4p\alpha,6p\}}] < \infty$ demonstrate that for all $q \in (1, \max\{4p\alpha, 6p\}]$ it holds that

$$\begin{aligned} \left|\mathbb{E}\Big[\max_{n\in\{0,1,\ldots,N\}} |\mathcal{M}_n|^q\Big]\right|^{1/q} &\leq \frac{q}{q-1} \big|\mathbb{E}\big[|\mathcal{M}_N|^q\big]\big|^{1/q} \\ &\leq \frac{q}{q-1} \big|\mathbb{E}\big[|\mathcal{M}_N|^{\max\{4p\alpha,6p\}}\big]\big|^{1/\max\{4p\alpha,6p\}} < \infty. \end{aligned} \tag{4.55}$$

Moreover, note that the triangle inequality and (4.54) prove that for all $q, r \in (0, \infty)$ with $1 < qr \leq \max\{4\alpha p, 6p\}$ it holds that

$$\begin{aligned} &\left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[|g(x)|^{qr}\big] \, \nu(dt, dx) \right]^{1/qr} \\ &\leq \left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[\|x\|^{qr}\big] \, \nu(dt, dx) \right]^{1/qr} \\ &\quad + \left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\left[\left| C + \max_{n\in\{0,1,\ldots,N\}} \mathcal{M}_n \right|^{qr} \right] \nu(dt, dx) \right]^{1/qr}. \end{aligned} \tag{4.56}$$

The triangle inequality, (4.53), and (4.55) hence show that for all $q, r \in (0, \infty)$ with $1 < qr \leq \max\{4\alpha p, 6p\}$ it holds that

$$\begin{aligned} &\left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[|g(x)|^{qr}\big] \, \nu(dt, dx) \right]^{1/qr} \\ &\leq \left[\int_{[0,T]\times\mathbb{R}^d} \|x\|^{qr} \, \nu(dt, dx) \right]^{1/qr} + \left[C + \left|\mathbb{E}\Big[\max_{n\in\{0,1,\ldots,N\}} |\mathcal{M}_n|^{qr}\Big]\right|^{1/qr} \right] [\nu([0,T]\times\mathbb{R}^d)]^{1/qr} \\ &\leq \mathscr{C}_{qr} + \left[C + \frac{qr}{qr-1} \big|\mathbb{E}\big[|\mathcal{M}_N|^{qr}\big]\big|^{1/qr} \right] [\nu([0,T]\times\mathbb{R}^d)]^{1/qr}. \end{aligned} \tag{4.57}$$

Combining this with (4.50), (4.54), and the fact that $\forall\, a, b \in [0, \infty),\ q \in (0, \infty) \colon (a + b)^q \leq 2^{\max\{q-1,0\}}(a^q + b^q)$ ensures that for all $q, r \in (0, \infty)$ with $1 < qr \leq \max\{4\alpha p, 6p\}$ it holds that

$$\begin{aligned} &\left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[|h_r(x)|^q\big] \, \nu(dt, dx) \right]^{1/q} \\ &\leq \max\big\{2^{(1/q)-1}, 1\big\} \left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{1}_{[0,T]\times\mathbb{R}^d}(t, x) \, \nu(dt, dx) \right]^{1/q} \end{aligned}$$

$$
\begin{aligned}
&+ e^{rc} \max\left\{2^{(1/q)-1}, 1\right\} \left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[|g(x)|^{rq}\big]\, \nu(dt,dx)\right]^{1/q} \\
&\leq \max\left\{2^{(1/q)-1}, 1\right\} [\nu([0,T]\times\mathbb{R}^d)]^{1/q} \\
&\quad + e^{rc} \max\left\{2^{(1/q)-1}, 1\right\} \left[\mathscr{C}_{qr} + \left[C + \frac{qr}{qr-1}\big|\mathbb{E}\big[|\mathcal{M}_N|^{qr}\big]\big|^{1/qr}\right] [\nu([0,T]\times\mathbb{R}^d)]^{1/qr}\right]^r. \quad (4.58)
\end{aligned}
$$

In addition, observe that (4.49) and Hölder's inequality show that for all $q, r \in (0,\infty)$ with $1 < qr \leq \max\{4\alpha p, 6p\}$ it holds that

$$
\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} \|x\|^{qr}\, \nu(dt,dx)\right]^{1/qr} \\
&\leq \left[\int_{[0,T]\times\mathbb{R}^d} \|x\|^{\max\{4\alpha p,6p\}}\, \nu(dt,dx)\right]^{1/\max\{4\alpha p,6p\}} [\nu([0,T]\times\mathbb{R}^d)]^{[1/qr - 1/\max\{4\alpha p,6p\}]} \\
&\leq \mathcal{C} \max\left\{1, [\nu([0,T]\times\mathbb{R}^d)]^{1/qr}\right\}. \quad (4.59)
\end{aligned}
$$

This, the fact that $\mathcal{C} \geq 1$, and (4.53) prove that for all $q, r \in (0,\infty)$ with $1 < qr \leq \max\{4\alpha p, 6p\}$ it holds that

$$
\mathscr{C}_{qr} \leq \mathcal{C} \max\left\{1, [\nu([0,T]\times\mathbb{R}^d)]^{1/qr}\right\}. \quad (4.60)
$$

Combining this with (4.58) implies that for all $q, r \in (0,\infty)$ with $1 < qr \leq \max\{4\alpha p, 6p\}$ it holds that

$$
\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[|h_r(x)|^{q}\big]\, \nu(dt,dx)\right]^{1/q} \\
&\leq \max\left\{2^{(1/q)-1}, 1\right\} [\nu([0,T]\times\mathbb{R}^d)]^{1/q} + e^{rc} \max\left\{2^{(1/q)-1}, 1\right\} \\
&\quad \times \left[\mathcal{C} + C + \frac{qr}{qr-1}\big|\mathbb{E}\big[|\mathcal{M}_N|^{qr}\big]\big|^{1/qr}\right]^r \max\left\{1, [\nu([0,T]\times\mathbb{R}^d)]^{1/q}\right\}. \quad (4.61)
\end{aligned}
$$

Therefore, we obtain that for all $q, r \in (0,\infty)$ with $1 < qr \leq \max\{4\alpha p, 6p\}$ it holds that

$$
\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[|h_r(x)|^{q}\big]\, \nu(dt,dx)\right]^{1/q} \\
&\leq 2e^{rc} \max\left\{2^{(1/q)-1}, 1\right\} \left[\mathcal{C} + C + \frac{qr}{qr-1}\big|\mathbb{E}\big[|\mathcal{M}_N|^{qr}\big]\big|^{1/qr}\right]^r \\
&\quad \times \max\left\{1, [\nu([0,T]\times\mathbb{R}^d)]^{1/q}\right\}. \quad (4.62)
\end{aligned}
$$

This establishes item (i). Next observe that Hölder's inequality assures that

$$
\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\Big[\big|[h_2(x)]^{\alpha} h_3(x)\big|^{p}\Big]\, \nu(dt,dx)\right]^{1/p} \\
&\leq \left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[|h_2(x)|^{2p\alpha}\big]\, \nu(dt,dx)\right]^{1/2p} \left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[|h_3(x)|^{2p}\big]\, \nu(dt,dx)\right]^{1/2p}. \quad (4.63)
\end{aligned}
$$

Moreover, note that Hölder's inequality demonstrates that

$$\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[|h_2(x)|^{2p\alpha}\big]\,\nu(dt,dx)\right]^{1/2p\alpha} \\
&\leq \left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[|h_2(x)|^{\max\{2p\alpha,3p\}}\big]\,\nu(dt,dx)\right]^{1/\max\{2p\alpha,3p\}} \\
&\quad \times [\nu([0,T]\times\mathbb{R}^d)]^{[1/2p\alpha - 1/\max\{2p\alpha,3p\}]}
\end{aligned} \tag{4.64}$$

and

$$\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[|h_3(x)|^{2p}\big]\,\nu(dt,dx)\right]^{1/2p} \\
&\leq \left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[|h_3(x)|^{\max\{(4/3)p\alpha,2p\}}\big]\,\nu(dt,dx)\right]^{1/\max\{(4/3)p\alpha,2p\}} \\
&\quad \times [\nu([0,T]\times\mathbb{R}^d)]^{[1/2p - 1/\max\{(4/3)p\alpha,2p\}]}.
\end{aligned} \tag{4.65}$$

Combining this with (4.62) implies that

$$\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[|h_2(x)|^{2p\alpha}\big]\,\nu(dt,dx)\right]^{1/2p\alpha} \\
&\leq 2e^{2c}\left[\mathcal{C} + C + \frac{\max\{4p\alpha,6p\}}{\max\{4p\alpha,6p\}-1}\big|\mathbb{E}\big[|\mathcal{M}_N|^{\max\{4p\alpha,6p\}}\big]\big|^{1/\max\{4p\alpha,6p\}}\right]^2 \\
&\quad \times \max\big\{1, [\nu([0,T]\times\mathbb{R}^d)]^{1/\max\{2p\alpha,3p\}}\big\}\ [\nu([0,T]\times\mathbb{R}^d)]^{[1/2p\alpha - 1/\max\{2p\alpha,3p\}]} \\
&\leq 2e^{2c}\left[\mathcal{C} + C + \frac{\max\{4p\alpha,6p\}}{\max\{4p\alpha,6p\}-1}\big|\mathbb{E}\big[|\mathcal{M}_N|^{\max\{4p\alpha,6p\}}\big]\big|^{1/\max\{4p\alpha,6p\}}\right]^2 \\
&\quad \times \max\big\{1, [\nu([0,T]\times\mathbb{R}^d)]^{1/2p\alpha}\big\}
\end{aligned} \tag{4.66}$$

and

$$\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[|h_3(x)|^{2p}\big]\,\nu(dt,dx)\right]^{1/2p} \\
&\leq 2e^{3c}\left[\mathcal{C} + C + \frac{\max\{4p\alpha,6p\}}{\max\{4p\alpha,6p\}-1}\big|\mathbb{E}\big[|\mathcal{M}_N|^{\max\{4p\alpha,6p\}}\big]\big|^{1/\max\{4p\alpha,6p\}}\right]^3 \\
&\quad \times \max\big\{1, [\nu([0,T]\times\mathbb{R}^d)]^{1/\max\{(4/3)p\alpha,2p\}}\big\}[\nu([0,T]\times\mathbb{R}^d)]^{[1/2p - 1/\max\{(4/3)p\alpha,2p\}]} \\
&\leq 2e^{3c}\left[\mathcal{C} + C + \frac{\max\{4p\alpha,6p\}}{\max\{4p\alpha,6p\}-1}\big|\mathbb{E}\big[|\mathcal{M}_N|^{\max\{4p\alpha,6p\}}\big]\big|^{1/\max\{4p\alpha,6p\}}\right]^3 \\
&\quad \times \max\big\{1, [\nu([0,T]\times\mathbb{R}^d)]^{1/2p}\big\}.
\end{aligned} \tag{4.67}$$

This and (4.63) show that

$$\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\Big[\big|[h_2(x)]^\alpha h_3(x)\big|^p\Big]\,\nu(dt,dx)\right]^{1/p} \\
&\leq \left[\left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[|h_2(x)|^{2p\alpha}\big]\,\nu(dt,dx)\right]^{1/2p\alpha}\right]^\alpha \left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[|h_3(x)|^{2p}\big]\,\nu(dt,dx)\right]^{1/2p}
\end{aligned}$$

$$\leq 2^{\alpha+1}e^{(2\alpha+3)c}\left[\mathcal{C}+C+\frac{\max\{4p\alpha,6p\}}{\max\{4p\alpha,6p\}-1}\big|\mathbb{E}\big[|\mathcal{M}_N|^{\max\{4p\alpha,6p\}}\big]\big|^{1/\max\{4p\alpha,6p\}}\right]^{2\alpha+3}$$
$$\times\max\big\{1,[\nu([0,T]\times\mathbb{R}^d)]^{1/2p}\big\}\Big[\max\big\{1,[\nu([0,T]\times\mathbb{R}^d)]^{1/2p\alpha}\big\}\Big]^{\alpha}. \tag{4.68}$$

Hence, we obtain that

$$\left[\int_{[0,T]\times\mathbb{R}^d}\mathbb{E}\Big[\big|[h_2(x)]^{\alpha}h_3(x)\big|^p\Big]\,\nu(dt,dx)\right]^{1/p}$$
$$\leq 2^{\alpha+1}e^{(2\alpha+3)c}\left[\mathcal{C}+C+\frac{\max\{4p\alpha,6p\}}{\max\{4p\alpha,6p\}-1}\big|\mathbb{E}\big[|\mathcal{M}_N|^{\max\{4p\alpha,6p\}}\big]\big|^{1/\max\{4p\alpha,6p\}}\right]^{2\alpha+3}$$
$$\times\max\big\{1,[\nu([0,T]\times\mathbb{R}^d)]^{1/p}\big\}. \tag{4.69}$$

This establishes item (ii). This completes the proof of Lemma 4.2. □

Corollary 4.1. *Let* $M,N,d,\mathfrak{d},k\in\mathbb{N}$, $p\in[2,\infty)$, $\alpha,c,C,\mathcal{C},\mathfrak{C}\in[0,\infty)$, $T,\mathfrak{D}\in(0,\infty)$, $B\in\mathbb{R}^{d\times k}$, $\varepsilon\in(0,1]$, $\mathfrak{f}^1,\mathfrak{f}^0\in\mathbf{N}$ *satisfy that* $\mathcal{I}(\mathfrak{f}^1)=\mathcal{O}(\mathfrak{f}^1)=\mathcal{I}(\mathfrak{f}^0)=d$, $\mathcal{O}(\mathfrak{f}^0)=\mathfrak{d}$, *and*

$$\mathfrak{D}=2160[\log_2(\varepsilon^{-1})+4]-504, \tag{4.70}$$

let $\nu\colon\mathcal{B}([0,T]\times\mathbb{R}^d)\to[0,\infty)$ *be a finite measure which satisfies that*

$$\mathcal{C}=\max\left\{\left[\int_{[0,T]\times\mathbb{R}^d}\|x\|^{\max\{4p\alpha,6p\}}\,\nu(dt,dx)\right]^{1/\max\{4p\alpha,6p\}},1\right\}, \tag{4.71}$$

assume for all $x,y\in\mathbb{R}^d$ *that* $\|(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^1))(x)\|\leq C+c\|x\|$ *and*

$$\|(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0))(x)-(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0))(y)\|\leq\mathfrak{C}(1+\|x\|^{\alpha}+\|y\|^{\alpha})\|x-y\|, \tag{4.72}$$

let $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_t)_{t\in[0,T]})$ *be a filtered probability space which satisfies the usual conditions*[1)]*, let* $W^m\colon[0,T]\times\Omega\to\mathbb{R}^k$, $m\in\{1,2,\dots,M\}$, *be standard* $(\mathbb{F}_t)_{t\in[0,T]}$*-Brownian motions, and let* $Y^m=(Y_t^{m,x}(\omega))_{(t,x,\omega)\in[0,T]\times\mathbb{R}^d\times\Omega}\colon[0,T]\times\mathbb{R}^d\times\Omega\to\mathbb{R}^d$, $m\in\{1,2,\dots,M\}$, *satisfy for all* $m\in\{1,2,\dots,M\}$, $x\in\mathbb{R}^d$, $n\in\{0,1,\dots,N-1\}$, $t\in[\frac{nT}{N},\frac{(n+1)T}{N}]$ *that* $Y_0^{m,x}=x$ *and*

$$Y_t^{m,x}=Y_{\frac{nT}{N}}^{m,x}+\left(\frac{tN}{T}-n\right)\left[\frac{T}{N}(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^1))(Y_{\frac{nT}{N}}^{m,x})+B\big(W_{\frac{(n+1)T}{N}}^m-W_{\frac{nT}{N}}^m\big)\right] \tag{4.73}$$

(cf. Definitions 2.1, 4.1, 4.3, and 4.4). Then there exists $(\Psi_\omega)_{\omega\in\Omega}\subseteq\mathbf{N}$ *such that*

(i) it holds for all $\omega\in\Omega$ *that* $\mathcal{R}_{\mathfrak{r}}(\Psi_\omega)\in C(\mathbb{R}^{d+1},\mathbb{R}^{\mathfrak{d}})$,

(ii) it holds for all $t\in[0,T]$, $x\in\mathbb{R}^d$ *that* $\Omega\ni\omega\mapsto(\mathcal{R}_{\mathfrak{r}}(\Psi_\omega))(t,x)\in\mathbb{R}^{\mathfrak{d}}$ *is measurable,*

(iii) it holds that

$$\left[\int_{[0,T]\times\mathbb{R}^d}\int_{\Omega}\left\|(\mathcal{R}_{\mathfrak{r}}(\Psi_\omega))(t,x)-\frac{1}{M}\left[\sum_{m=1}^{M}(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0))(Y_t^{m,x}(\omega))\right]\right\|^p\mathbb{P}(d\omega)\,\nu(dt,dx)\right]^{1/p}$$
$$\leq\big[2p\max\{C,\mathcal{C}\}\max\{T,1\}\max\{\alpha,1\}\big]^{2\alpha+3}\Big[1+\sqrt{\operatorname{Trace}(B^*B)}\Big]^{2\alpha+3}$$
$$\times\,\varepsilon d^{(\alpha+1)/2}\mathfrak{C}\,e^{(2\alpha+3)cT}2^{2\alpha+4}3^{\alpha}\max\big\{1,[\nu([0,T]\times\mathbb{R}^d)]^{1/p}\big\}, \tag{4.74}$$

[1)] Note that we say that a filtered probability space $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_t)_{t\in[0,T]})$ satisfies the usual conditions if and only if it holds for all $t\in[0,T)$ that $\{A\in\mathcal{F}\colon\mathbb{P}(A)=0\}\subseteq\mathbb{F}_t=(\cap_{s\in(t,T]}\mathbb{F}_s)$; cf., e.g., Liu & Röckner [48, Definition 2.1.11].

(iv) *it holds for all* $\omega \in \Omega$ *that*

$$\mathcal{P}(\Psi_\omega) \leq 2M^2\mathcal{P}(\mathfrak{f}^0) + 9M^2N^6d^{16}\big[2\mathcal{L}(\mathfrak{f}^1) + \mathfrak{D} + (24 + 6\mathcal{L}(\mathfrak{f}^1) + [4 + \mathcal{P}(\mathfrak{f}^1)]^2)^2\big]^2. \quad (4.75)$$

Proof. Throughout this proof let $h_{m,r} \colon \mathbb{R}^d \times \Omega \to [0,\infty)$, $m \in \{1,2,\dots,M\}$, $r \in \mathbb{R}$, satisfy for all $m \in \{1,2,\dots,M\}$, $r \in \mathbb{R}$, $x \in \mathbb{R}^d$ that

$$h_{m,r}(x) = 1 + \Big[\|x\| + CT + \max_{n \in \{0,1,\dots,N\}} \big\|BW^m_{\frac{nT}{N}}\big\|\Big]^r \exp(rcT), \quad (4.76)$$

let $(\Psi_\omega)_{\omega\in\Omega} \subseteq \mathbf{N}$ satisfy that

(I) it holds for all $\omega \in \Omega$ that $\mathcal{R}_\mathfrak{r}(\Psi_\omega) \in C(\mathbb{R}^{d+1}, \mathbb{R}^\mathfrak{d})$,

(II) it holds for all $\omega \in \Omega$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$\begin{aligned}&\left\|(\mathcal{R}_\mathfrak{r}(\Psi_\omega))(t,x) - \frac{1}{M}\left[\sum_{m=1}^M (\mathcal{R}_\mathfrak{r}(\mathfrak{f}^0))(Y_t^{m,x}(\omega))\right]\right\| \\ &\leq \frac{2\varepsilon\mathfrak{C}\sqrt{d}}{M}\left[\sum_{m=1}^M \Big[1 + 2d^{\alpha/2}6^\alpha |h_{m,2}(x,\omega)|^\alpha\Big] h_{m,3}(x,\omega)\right],\end{aligned} \quad (4.77)$$

(III) it holds for all $\omega \in \Omega$ that

$$\mathcal{P}(\Psi_\omega) \leq 2M^2\mathcal{P}(\mathfrak{f}^0) + 9M^2N^6d^{16}\big[2\mathcal{L}(\mathfrak{f}^1) + \mathfrak{D} + (24 + 6\mathcal{L}(\mathfrak{f}^1) + [4 + \mathcal{P}(\mathfrak{f}^1)]^2)^2\big]^2, \quad (4.78)$$

(IV) it holds for all $t \in [0,T]$, $x \in \mathbb{R}^d$ that $\Omega \ni \omega \mapsto (\mathcal{R}_\mathfrak{r}(\Psi_\omega))(t,x) \in \mathbb{R}^\mathfrak{d}$ is measurable

(cf. Proposition 4.2), let $Z = (Z_t^{x,y})_{(t,x,y)\in[0,T]\times\mathbb{R}^d\times(\mathbb{R}^d)^N} \colon [0,T] \times \mathbb{R}^d \times (\mathbb{R}^d)^N \to \mathbb{R}^d$ satisfy for all $n \in \{0,1,\dots,N-1\}$, $t \in [\frac{nT}{N}, \frac{(n+1)T}{N}]$, $x \in \mathbb{R}^d$, $y = (y_1,y_2,\dots,y_N) \in (\mathbb{R}^d)^N$ that $Z_0^{x,y} = x$ and

$$Z_t^{x,y} = Z^{x,y}_{\frac{nT}{N}} + \left(\frac{tN}{T} - n\right)\left[\frac{T}{N}(\mathcal{R}_\mathfrak{r}(\mathfrak{f}^1))\big(Z^{x,y}_{\frac{nT}{N}}\big) + y_{n+1}\right], \quad (4.79)$$

and let $\mathcal{W}^m \colon [0,T] \times \mathbb{R}^d \times \Omega \to [0,T] \times \mathbb{R}^d \times (\mathbb{R}^d)^N$, $m \in \{1,2,\dots,M\}$, satisfy for all $m \in \{1,2,\dots,M\}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ that

$$\mathcal{W}^m(t,x) = \Big(t, x, B\big(W^m_{\frac{T}{N}} - W^m_0\big), B\big(W^m_{\frac{2T}{N}} - W^m_{\frac{T}{N}}\big), \dots, B\big(W^m_{\frac{NT}{N}} - W^m_{\frac{(N-1)T}{N}}\big)\Big). \quad (4.80)$$

Note that (I), (IV), and Beck *et al.* [2, Lemma 2.4] demonstrate that $[0,T]\times\mathbb{R}^d\times\Omega \ni (t,x,\omega) \mapsto (\mathcal{R}_\mathfrak{r}(\Psi_\omega))(t,x) \in \mathbb{R}^\mathfrak{d}$ is measurable. In addition, observe that (4.73), (4.79), and (4.80) ensure that for all $m \in \{1,2,\dots,M\}$, $t \in [0,T]$, $x \in \mathbb{R}^d$ it holds that

$$Y_t^{m,x} = (Z \circ \mathcal{W}^m)(t,x). \quad (4.81)$$

Next note that Grohs *et al.* [26, Lemma 3.8] (applied with $N \leftarrow N$, $d \leftarrow d$, $\mu \leftarrow \mathcal{R}_\mathfrak{r}(\mathfrak{f}^1)$, $T \leftarrow T$, $(\{-1,0,1,\dots,N+1\} \ni n \mapsto t_n \in \mathbb{R}) \leftarrow (\{-1,0,1,\dots,N+1\} \ni n \mapsto \frac{nT}{N} \in \mathbb{R})$, $Y \leftarrow Z$ in the notation of Grohs *et al.* [26, Lemma 3.8]) proves that $Z \in C([0,T] \times \mathbb{R}^d \times (\mathbb{R}^d)^N, \mathbb{R}^d)$. Combining this with (4.81) and the fact that for all $m \in \{1,2,\dots,M\}$ it holds that $\mathcal{W}^m$ is measurable shows that for all $m \in \{1, 2, \dots, M\}$ it holds that Y^m is measurable. The fact that $\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0) \in C(\mathbb{R}^d, \mathbb{R}^{\mathfrak{d}})$ hence ensures that

$$[0,T] \times \mathbb{R}^d \times \Omega \ni (t,x,\omega) \mapsto \frac{1}{M}\left[\sum_{m=1}^{M} (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0))(Y_t^{m,x}(\omega))\right] \in \mathbb{R}^{\mathfrak{d}} \tag{4.82}$$

is measurable. Combining this with (4.77) and the fact that $[0,T] \times \mathbb{R}^d \times \Omega \ni (t,x,\omega) \mapsto (\mathcal{R}_{\mathfrak{r}}(\Psi_\omega))(t,x) \in \mathbb{R}^{\mathfrak{d}}$ is measurable proves that

$$\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} \int_\Omega \left\| (\mathcal{R}_{\mathfrak{r}}(\Psi_\omega))(t,x) - \frac{1}{M}\left[\sum_{m=1}^{M} (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0))(Y_t^{m,x}(\omega))\right]\right\|^p \mathbb{P}(d\omega)\,\nu(dt,dx)\right]^{1/p} \\
&\leq \frac{2\varepsilon\mathfrak{C}\sqrt{d}}{M}\left[\int_{[0,T]\times\mathbb{R}^d} \int_\Omega \left|\sum_{m=1}^{M}\left[1 + 2d^{\alpha/2}6^\alpha |h_{m,2}(x,\omega)|^\alpha\right] h_{m,3}(x,\omega)\right|^p \mathbb{P}(d\omega)\,\nu(dt,dx)\right]^{1/p}.
\end{aligned} \tag{4.83}$$

The triangle inequality therefore implies that

$$\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} \int_\Omega \left\| (\mathcal{R}_{\mathfrak{r}}(\Psi_\omega))(t,x) - \frac{1}{M}\left[\sum_{m=1}^{M} (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0))(Y_t^{m,x}(\omega))\right]\right\|^p \mathbb{P}(d\omega)\,\nu(dt,dx)\right]^{1/p} \\
&\leq \frac{2\varepsilon\mathfrak{C}\sqrt{d}}{M}\sum_{m=1}^{M}\left[\int_{[0,T]\times\mathbb{R}^d} \int_\Omega |h_{m,3}(x,\omega)|^p\,\mathbb{P}(d\omega)\,\nu(dt,dx)\right]^{1/p} \\
&\quad + \frac{4d^{\alpha/2}6^\alpha\varepsilon\mathfrak{C}\sqrt{d}}{M}\sum_{m=1}^{M}\left[\int_{[0,T]\times\mathbb{R}^d} \int_\Omega |[h_{m,2}(x,\omega)]^\alpha h_{m,3}(x,\omega)|^p\,\mathbb{P}(d\omega)\,\nu(dt,dx)\right]^{1/p}.
\end{aligned} \tag{4.84}$$

Next note that (4.76), Lemma 4.2, and the fact that for all $m \in \{1, 2, \dots, M\}$ it holds that $(\|BW^m_{\frac{nT}{N}}\|)_{n\in\{0,1,\dots,N\}}$ is a nonnegative $(\mathbb{F}_{\frac{nT}{N}})_{n\in\{0,1,\dots,N\}}$-submartingale demonstrate that for all $m \in \{1, 2, \dots, M\}$ it holds that

$$\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[|h_{m,3}(x)|^p\big]\,\nu(dt,dx)\right]^{1/p} \\
&\leq 2e^{3cT}\left[\mathcal{C} + CT + \frac{3p}{3p-1}\big|\mathbb{E}\big[\|BW_T^m\|^{3p}\big]\big|^{1/3p}\right]^3 \max\big\{1, [\nu([0,T]\times\mathbb{R}^d)]^{1/p}\big\}
\end{aligned} \tag{4.85}$$

and

$$\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[|[h_{m,2}(x)]^\alpha h_{m,3}(x)|^p\big]\,\nu(dt,dx)\right]^{1/p} \\
&\leq \left[\mathcal{C} + CT + \frac{\max\{4p\alpha, 6p\}}{\max\{4p\alpha, 6p\}-1}\big|\mathbb{E}\big[\|BW_T^m\|^{\max\{4p\alpha,6p\}}\big]\big|^{1/\max\{4p\alpha,6p\}}\right]^{2\alpha+3} \\
&\quad \times 2^{\alpha+1}e^{(2\alpha+3)cT}\max\big\{1, [\nu([0,T]\times\mathbb{R}^d)]^{1/p}\big\}.
\end{aligned} \tag{4.86}$$

Moreover, observe that Lemma 2.1, the fact that for all $m \in \{1, 2, \dots, M\}$ it holds that BW_T^m is a Gaussian random variable, and the fact that for all $m \in \{1, 2, \dots, M\}$ it holds that $\operatorname{Cov}(BW_T^m) = TB^*B$ ensure that for all $m \in \{1, 2, \dots, M\}$, $q \in [2, \infty)$ it holds that

$$\begin{aligned} \frac{q}{q-1} \big| \mathbb{E}\big[\|BW_T^m\|^q \big] \big|^{1/q} &\leq \frac{q}{q-1} \sqrt{\max\{1, q-1\} T \operatorname{Trace}(B^*B)} \\ &= \frac{q}{q-1} \sqrt{(q-1) T \operatorname{Trace}(B^*B)} \\ &= \frac{q}{\sqrt{q-1}} \sqrt{T \operatorname{Trace}(B^*B)}. \end{aligned} \tag{4.87}$$

Combining this with (4.85) and (4.86) assures that

$$\begin{aligned} &\frac{1}{M} \sum_{m=1}^M \left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[|h_{m,3}(x)|^p \big] \, \nu(dt, dx) \right]^{1/p} \\ &\leq 2e^{3cT} \left[\mathcal{C} + CT + \frac{3p}{\sqrt{3p-1}} \sqrt{T \operatorname{Trace}(B^*B)} \right]^3 \max\big\{1, [\nu([0,T]\times\mathbb{R}^d)]^{1/p}\big\} \end{aligned} \tag{4.88}$$

and

$$\begin{aligned} &\frac{1}{M} \sum_{m=1}^M \left[\int_{[0,T]\times\mathbb{R}^d} \mathbb{E}\big[\big| [h_{m,2}(x)]^\alpha h_{m,3}(x) \big|^p \big] \, \nu(dt, dx) \right]^{1/p} \\ &\leq \left[\mathcal{C} + CT + \frac{\max\{4p\alpha, 6p\}}{\sqrt{\max\{4p\alpha, 6p\} - 1}} \sqrt{T \operatorname{Trace}(B^*B)} \right]^{2\alpha+3} \\ &\quad \times 2^{\alpha+1} e^{(2\alpha+3)cT} \max\big\{1, [\nu([0,T]\times\mathbb{R}^d)]^{1/p}\big\}. \end{aligned} \tag{4.89}$$

This and (4.84) demonstrate that

$$\begin{aligned} &\left[\int_{[0,T]\times\mathbb{R}^d} \int_\Omega \left\| (\mathcal{R}_\mathfrak{r}(\Psi_\omega))(t,x) - \frac{1}{M} \left[\sum_{m=1}^M (\mathcal{R}_\mathfrak{r}(\mathfrak{f}^0))(Y_t^{m,x}(\omega)) \right] \right\|^p \mathbb{P}(d\omega) \, \nu(dt, dx) \right]^{1/p} \\ &\leq 4\varepsilon \mathfrak{C} e^{3cT} \sqrt{d} \max\big\{1, [\nu([0,T]\times\mathbb{R}^d)]^{1/p}\big\} \left[\mathcal{C} + CT + \frac{3p}{\sqrt{3p-1}} \sqrt{T \operatorname{Trace}(B^*B)} \right]^3 \\ &\quad + \left[\mathcal{C} + CT + \frac{\max\{4p\alpha, 6p\}}{\sqrt{\max\{4p\alpha, 6p\} - 1}} \sqrt{T \operatorname{Trace}(B^*B)} \right]^{2\alpha+3} \\ &\quad \times 4\varepsilon \mathfrak{C} 6^\alpha d^{\alpha/2} 2^{\alpha+1} e^{(2\alpha+3)cT} \sqrt{d} \max\big\{1, [\nu([0,T]\times\mathbb{R}^d)]^{1/p}\big\}. \end{aligned} \tag{4.90}$$

The fact that $[2, \infty) \ni x \mapsto \frac{x}{\sqrt{x-1}} \in \mathbb{R}$ is non-decreasing hence implies that

$$\begin{aligned} &\left[\int_{[0,T]\times\mathbb{R}^d} \int_\Omega \left\| (\mathcal{R}_\mathfrak{r}(\Psi_\omega))(t,x) - \frac{1}{M} \left[\sum_{m=1}^M (\mathcal{R}_\mathfrak{r}(\mathfrak{f}^0))(Y_t^{m,x}(\omega)) \right] \right\|^p \mathbb{P}(d\omega) \, \nu(dt, dx) \right]^{1/p} \\ &\leq \left[\mathcal{C} + CT + \frac{\max\{4p\alpha, 6p\}}{\sqrt{\max\{4p\alpha, 6p\} - 1}} \sqrt{T \operatorname{Trace}(B^*B)} \right]^{2\alpha+3} \\ &\quad \times \left[4\varepsilon \mathfrak{C} e^{3cT} \sqrt{d} + 4\varepsilon \mathfrak{C} 6^\alpha d^{\alpha/2} 2^{\alpha+1} e^{(2\alpha+3)cT} \sqrt{d} \right] \max\big\{1, [\nu([0,T]\times\mathbb{R}^d)]^{1/p}\big\} \\ &\leq |\max\{C, \mathcal{C}\}|^{2\alpha+3} |\max\{T, 1\}|^{2\alpha+3} \left[2 + \frac{\max\{4p\alpha, 6p\}}{\sqrt{\max\{4p\alpha, 6p\} - 1}} \sqrt{\operatorname{Trace}(B^*B)} \right]^{2\alpha+3} \end{aligned}$$

$$\times\, \varepsilon d^{(\alpha+1)/2} \mathfrak{C}\, e^{(2\alpha+3)cT}\big[4 + 6^{\alpha} 2^{\alpha+1} 4\big] \max\big\{1, [\nu([0,T]\times\mathbb{R}^d)]^{1/p}\big\}. \tag{4.91}$$

The fact that $\sqrt{\max\{4p\alpha, 6p\}-1} \geq \sqrt{6p-1} \geq \sqrt{11} \geq 3$ therefore ensures that

$$\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d}\int_\Omega \left\|(\mathcal{R}_{\mathfrak{r}}(\Psi_\omega))(t,x) - \frac{1}{M}\left[\sum_{m=1}^{M}(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0))(Y_t^{m,x}(\omega))\right]\right\|^p \mathbb{P}(d\omega)\,\nu(dt,dx)\right]^{1/p}\\
&\leq |\max\{C,\mathcal{C}\}|^{2\alpha+3}|\max\{T,1\}|^{2\alpha+3}\left[2+\max\left\{\left(\frac{4}{3}\right)p\alpha, 2p\right\}\sqrt{\operatorname{Trace}(B^*B)}\right]^{2\alpha+3}\\
&\quad\times \varepsilon d^{(\alpha+1)/2}\mathfrak{C}\, e^{(2\alpha+3)cT}\big[4+2^{2\alpha+3}3^{\alpha}\big]\max\big\{1,[\nu([0,T]\times\mathbb{R}^d)]^{1/p}\big\}\\
&\leq |\max\{C,\mathcal{C}\}|^{2\alpha+3}|\max\{T,1\}|^{2\alpha+3}[2p\max\{\alpha,1\}]^{2\alpha+3}\left[1+\sqrt{\operatorname{Trace}(B^*B)}\right]^{2\alpha+3}\\
&\quad\times \varepsilon d^{(\alpha+1)/2}\mathfrak{C}\, e^{(2\alpha+3)cT}2^{2\alpha+4}3^{\alpha}\max\big\{1,[\nu([0,T]\times\mathbb{R}^d)]^{1/p}\big\}.
\end{aligned}\tag{4.92}$$

Combining this with (I), (III), and (IV) establishes items (i)-(iv). This completes the proof of Corollary 4.1. □

4.5. Approximation error estimates for deep ANNs

Proposition 4.3. *Let $T,\kappa\in(0,\infty)$, $\eta\in[1,\infty)$, $p\in[2,\infty)$, let $A_d=(a_{d,i,j})_{(i,j)\in\{1,2,\dots,d\}^2}\in\mathbb{R}^{d\times d}$, $d\in\mathbb{N}$, be symmetric positive semidefinite matrices, let $\nu_d\colon\mathcal{B}([0,T]\times\mathbb{R}^d)\to[0,\infty)$, $d\in\mathbb{N}$, be finite measures which satisfy for all $d\in\mathbb{N}$ that*

$$\left[\int_{[0,T]\times\mathbb{R}^d}\|x\|^{2p\max\{2\kappa,3\}}\,\nu_d(dt,dx)\right]^{1/p}\leq \eta d^{\eta}, \tag{4.93}$$

let $f_d^m\colon\mathbb{R}^d\to\mathbb{R}^{md-m+1}$, $m\in\{0,1\}$, $d\in\mathbb{N}$, be functions, let $(\mathfrak{f}^m_{d,\varepsilon})_{(m,d,\varepsilon)\in\{0,1\}\times\mathbb{N}\times(0,1]}\subseteq\mathbf{N}$, assume for all $m\in\{0,1\}$, $d\in\mathbb{N}$, $\varepsilon\in(0,1]$, $x,y\in\mathbb{R}^d$ that

$$\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0_{d,\varepsilon})\in C(\mathbb{R}^d,\mathbb{R}),\quad \mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^1_{d,\varepsilon})\in C(\mathbb{R}^d,\mathbb{R}^d),\quad \mathcal{P}(\mathfrak{f}^m_{d,\varepsilon})\leq\kappa d^{\kappa}\varepsilon^{-\kappa}, \tag{4.94}$$
$$\big\|f_d^1(x)-f_d^1(y)\big\|\leq\kappa\|x-y\|,\quad \|(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^1_{d,\varepsilon}))(x)\|\leq\kappa(d^{\kappa}+\|x\|), \tag{4.95}$$
$$\big|(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0_{d,\varepsilon}))(x)-(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0_{d,\varepsilon}))(y)\big|\leq\kappa d^{\kappa}(1+\|x\|^{\kappa}+\|y\|^{\kappa})\|x-y\|, \tag{4.96}$$
$$\big\|f_d^m(x)-(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^m_{d,\varepsilon}))(x)\big\|\leq\varepsilon\kappa d^{\kappa}(1+\|x\|^{\kappa}), \tag{4.97}$$
$$\big|f_d^0(x)\big|+\operatorname{Trace}(A_d)\leq\kappa d^{\kappa}(1+\|x\|^{\kappa}), \tag{4.98}$$

and for every $d\in\mathbb{N}$ let $u_d\in\{v\in C([0,T]\times\mathbb{R}^d,\mathbb{R})\colon \inf_{q\in(0,\infty)}\sup_{(t,y)\in[0,T]\times\mathbb{R}^d}\frac{|v(t,y)|}{1+\|y\|^q}<\infty\}$ be a viscosity solution of

$$\left(\frac{\partial}{\partial t}u_d\right)(t,x)=\left(\frac{\partial}{\partial x}u_d\right)(t,x)\,f_d^1(x)+\sum_{i,j=1}^{d}a_{d,i,j}\left(\frac{\partial^2}{\partial x_i\partial x_j}u_d\right)(t,x) \tag{4.99}$$

with $u_d(0,x)=f_d^0(x)$ for $(t,x)\in(0,T)\times\mathbb{R}^d$ (cf. Definitions 2.1, 4.1, 4.3, and 4.4). Then there exist $\mathcal{C}\in\mathbb{R}$ and $(\mathfrak{u}_{d,N,M,\delta})_{(d,N,M,\delta)\in\mathbb{N}^3\times(0,1]}\subseteq\mathbf{N}$ such that

(i) it holds for all $d,N,M\in\mathbb{N}$, $\delta\in(0,1]$ that $\mathcal{R}_{\mathfrak{r}}(\mathfrak{u}_{d,N,M,\delta})\in C(\mathbb{R}^{d+1},\mathbb{R})$,

(ii) it holds for all $d, N, M \in \mathbb{N}$, $\delta \in (0,1]$ that

$$\begin{aligned}&\left[\int_{[0,T]\times\mathbb{R}^d} |u_d(y) - (\mathcal{R}_{\mathfrak{r}}(\mathfrak{u}_{d,N,M,\delta}))(y)|^p \,\nu_d(dy)\right]^{1/p}\\ &\leq \mathcal{C}\left[\max\left\{1, \nu_d([0,T]\times\mathbb{R}^d)\right\}\right]^{1/p}\\ &\quad\times\left[\frac{d^{\kappa(\kappa+4)+\max\{\eta,\kappa(2\kappa+1)\}}}{N^{1/2}} + \frac{d^{\kappa+\max\{\eta,\kappa^2\}}}{M^{1/2}} + \delta d^{(2\kappa+3)\max\{\eta,\kappa\}+\kappa^2+(7\kappa+1)/2}\right],\end{aligned} \tag{4.100}$$

(iii) it holds for all $d, N, M \in \mathbb{N}$, $\delta \in (0,1]$ that

$$\mathcal{P}(\mathfrak{u}_{d,N,M,\delta}) \leq \mathcal{C} M^2 N^{6+4\kappa}[\log_2(\delta^{-1}) + 1]^2 d^{16+8\kappa}. \tag{4.101}$$

Proof. Throughout this proof let $\mathfrak{D}_\delta \in \mathbb{R}$, $\delta \in (0,1]$, satisfy for all $\delta \in (0,1]$ that

$$\mathfrak{D}_\delta = 2160[\log_2(\delta^{-1}) + 4] - 504, \tag{4.102}$$

let $C_d \in \mathbb{R}$, $d \in \mathbb{N}$, satisfy for all $d \in \mathbb{N}$ that

$$C_d = \max\left\{\left[\int_{[0,T]\times\mathbb{R}^d} \|x\|^{\max\{4p\kappa,6p\}} \,\nu_d(dt,dx)\right]^{1/\max\{4p\kappa,6p\}}, 1\right\}, \tag{4.103}$$

let $\mathcal{C}_1 \in (0,\infty)$ satisfy that

$$\mathcal{C}_1 = \left[2p\max\{\eta,\kappa\}\max\{T,1\}\max\{\kappa,1\}\right]^{2\kappa+3}[1+(2\kappa)^{1/2}]^{2\kappa+3}\kappa e^{(2\kappa+3)\kappa T} 2^{2\kappa+4}\, 3^\kappa, \tag{4.104}$$

let $\mathcal{C}_2 \in (0,\infty)$ satisfy that

$$\mathcal{C}_2 = 2^{57}[\max\{\kappa,1\}]^8[\max\{T^{-\kappa/2},1\}]^8, \tag{4.105}$$

let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_t)_{t\in[0,T]})$ be a filtered probability space which satisfies the usual conditions, let $W^{d,m}\colon [0,T]\times\Omega \to \mathbb{R}^d$, $d, m \in \mathbb{N}$, be independent standard $(\mathbb{F}_t)_{t\in[0,T]}$-Brownian motions, let $Y^{N,d,m,x}\colon [0,T]\times\Omega \to \mathbb{R}^d$, $x \in \mathbb{R}^d$, $N, d, m \in \mathbb{N}$, be stochastic processes which satisfy for all $N, d, m \in \mathbb{N}$, $x \in \mathbb{R}^d$, $n \in \{0, 1, \ldots, N-1\}$, $t \in [\frac{nT}{N}, \frac{(n+1)T}{N}]$ that $Y_0^{N,d,m,x} = x$ and

$$\begin{aligned}Y_t^{N,d,m,x} &= Y_{\frac{nT}{N}}^{N,d,m,x} + \left(\frac{tN}{T} - n\right)\\ &\quad\times\left[\frac{T}{N}\big(\mathcal{R}_{\mathfrak{r}}\big(\mathfrak{f}^1_{d,\min\{(T/N)^{1/2},1\}}\big)\big)\big(Y_{\frac{nT}{N}}^{N,d,m,x}\big) + \sqrt{2A_d}\big(W^{d,m}_{\frac{(n+1)T}{N}} - W^{d,m}_{\frac{nT}{N}}\big)\right],\end{aligned} \tag{4.106}$$

let $(\psi_{d,N,M,\delta,\omega})_{(d,N,M,\delta,\omega)\in\mathbb{N}^3\times(0,1]\times\Omega} \subseteq \mathbf{N}$ satisfy that

(I) it holds for all $d, N, M \in \mathbb{N}$, $\delta \in (0,1]$, $\omega \in \Omega$ that $\mathcal{R}_{\mathfrak{r}}(\psi_{d,N,M,\delta,\omega}) \in C(\mathbb{R}^{d+1},\mathbb{R})$,

(II) it holds for all $d, N, M \in \mathbb{N}, \delta \in (0,1], t \in [0,T], x \in \mathbb{R}^d$ that $\Omega \ni \omega \mapsto (\mathcal{R}_{\mathfrak{r}}(\psi_{d,N,M,\delta,\omega}))(t,x) \in \mathbb{R}$ is measurable,

(III) it holds for all $d, N, M \in \mathbb{N}$, $\delta \in (0,1]$ that

$$\left[\int_{[0,T]\times\mathbb{R}^d}\int_\Omega\Bigg|(\mathcal{R}_{\mathfrak{r}}(\psi_{d,N,M,\delta,\omega}))(t,x) - \frac{1}{M}\left[\sum_{m=1}^{M}\big(\mathcal{R}_{\mathfrak{r}}\big(\mathfrak{f}^0_{d,\min\{(T/N)^{1/2},1\}}\big)\big)\big(Y^{N,d,m,x}_t(\omega)\big)\right]\Bigg|^p \mathbb{P}(d\omega)\,\nu_d(dt,dx)\right]^{1/p}$$
$$\leq \big[2p\max\{C_d,\kappa d^\kappa\}\max\{T,1\}\max\{\kappa,1\}\big]^{2\kappa+3}\left[1+\sqrt{\operatorname{Trace}(2A_d)}\right]^{2\kappa+3} \times \delta d^{(\kappa+1)/2}\kappa d^\kappa e^{(2\kappa+3)\kappa T}2^{2\kappa+4}\,3^\kappa\max\big\{1,[\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\big\}, \tag{4.107}$$

(IV) it holds for all $d, N, M \in \mathbb{N}$, $\delta \in (0,1]$, $\omega \in \Omega$ that

$$\begin{aligned}\mathcal{P}(\psi_{d,N,M,\delta,\omega}) &\leq 2M^2\mathcal{P}\big(\mathfrak{f}^0_{d,\min\{(T/N)^{1/2},1\}}\big) + 9M^2N^6d^{16}\Big[2\mathcal{L}\big(\mathfrak{f}^1_{d,\min\{(T/N)^{1/2},1\}}\big) + \mathfrak{D}_\delta \\ &\quad + \big(24 + 6\mathcal{L}\big(\mathfrak{f}^1_{d,\min\{(T/N)^{1/2},1\}}\big) + \big[4+\mathcal{P}\big(\mathfrak{f}^1_{d,\min\{(T/N)^{1/2},1\}}\big)\big]^2\big)^2\Big]^2\end{aligned} \tag{4.108}$$

(cf. Corollary 4.1 (applied with $M \leftarrow M$, $N \leftarrow N$, $d \leftarrow d$, $\mathfrak{d} \leftarrow 1$, $k \leftarrow d$, $p \leftarrow p$, $\alpha \leftarrow \kappa$, $c \leftarrow \kappa$, $C \leftarrow \kappa d^\kappa$, $\mathcal{C} \leftarrow C_d$, $\mathfrak{C} \leftarrow \kappa d^\kappa$, $T \leftarrow T$, $\mathfrak{D} \leftarrow \mathfrak{D}_\delta$, $B \leftarrow \sqrt{2A_d}$, $\varepsilon \leftarrow \delta$, $\mathfrak{f}^1 \leftarrow \mathfrak{f}^1_{d,\min\{(T/N)^{1/2},1\}}$, $\mathfrak{f}^0 \leftarrow \mathfrak{f}^0_{d,\min\{(T/N)^{1/2},1\}}$, $\nu \leftarrow \nu_d$, $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_t)_{t\in[0,T]}) \leftarrow (\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_t)_{t\in[0,T]})$, $W^m \leftarrow W^{d,m}$, $Y^m \leftarrow (Y^{N,d,m,x})_{x\in\mathbb{R}^d}$ for $d,N,M \in \mathbb{N}$, $m \in \{1,2,\ldots,M\}$, $\delta \in (0,1]$ in the notation of Corollary 4.1)), let $Z_{d,N,M,\delta}\colon \Omega \to [0,\infty]$, $d,N,M \in \mathbb{N}$, $\delta \in (0,1]$, satisfy for all $d,N,M \in \mathbb{N}$, $\delta \in (0,1]$, $\omega \in \Omega$ that

$$Z_{d,N,M,\delta}(\omega) = \int_{[0,T]\times\mathbb{R}^d}|u_d(y) - (\mathcal{R}_{\mathfrak{r}}(\psi_{d,N,M,\delta,\omega}))(y)|^p\,\nu_d(dy) \tag{4.109}$$

(cf. (I)), and let $\mathcal{C}_3 \in (0,\infty)$ satisfy that for all $d,N,M \in \mathbb{N}$ it holds that

$$\left(\mathbb{E}\left[\int_{[0,T]\times\mathbb{R}^d}\left|u_d(t,x) - \frac{1}{M}\left[\sum_{m=1}^{M}\big(\mathcal{R}_{\mathfrak{r}}\big(\mathfrak{f}^0_{d,\min\{(T/N)^{1/2},1\}}\big)\big)\big(Y^{N,d,m,x}_t\big)\right]\right|^p\nu_d(dt,dx)\right]\right)^{1/p}$$
$$\leq \mathcal{C}_3\left[\frac{d^{\kappa(\kappa+4)+\max\{\eta,\kappa(2\kappa+1)\}}}{N^{1/2}} + \frac{d^{\kappa+\max\{\eta,\kappa^2\}}}{M^{1/2}}\right]\big[\max\big\{1,\nu_d([0,T]\times\mathbb{R}^d)\big\}\big]^{1/p} \tag{4.110}$$

(cf. item (ii) in Proposition 3.2 (applied with $T \leftarrow T$, $\kappa \leftarrow \kappa$, $\eta \leftarrow \eta$, $p \leftarrow p$, $A_d \leftarrow A_d$, $\nu_d \leftarrow \nu_d$, $f^0_d \leftarrow f^0_d$, $f^1_d \leftarrow f^1_d$ $F^0_{d,\varepsilon} \leftarrow \mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0_{d,\varepsilon})$, $F^1_{d,\varepsilon} \leftarrow \mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^1_{d,\varepsilon})$, $(\Omega,\mathcal{F},\mathbb{P}) \leftarrow (\Omega,\mathcal{F},\mathbb{P})$, $W^{d,m} \leftarrow W^{d,m}$, $Y^{N,d,m,x} \leftarrow Y^{N,d,m,x}$ for $d,N,M \in \mathbb{N}$, $m \in \{1,2,\ldots,M\}$, $\varepsilon \in (0,1]$, $x \in \mathbb{R}^d$ in the notation of Proposition 3.2)). Observe that (4.93) and (4.103) demonstrate that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned}\max\{C_d,\kappa d^\kappa\} &= \max\left\{\left[\int_{[0,T]\times\mathbb{R}^d}\|x\|^{2p\max\{2\kappa,3\}}\,\nu_d(dt,dx)\right]^{1/2p\max\{2\kappa,3\}},1,\kappa d^\kappa\right\} \\ &\leq \max\left\{(\eta d^\eta)^{1/(2\max\{2\kappa,3\})},1,\kappa d^\kappa\right\} \leq \max\{\eta d^\eta,1,\kappa d^\kappa\} \\ &\leq \max\{\eta,1,\kappa\}\,d^{\max\{\eta,\kappa\}} = \max\{\eta,\kappa\}\,d^{\max\{\eta,\kappa\}}.\end{aligned} \tag{4.111}$$

Next note that (4.98) proves that

$$1+\sqrt{\mathrm{Trace}(2A_d)} \leq 1+(2\kappa d^\kappa)^{1/2} \leq d^{\kappa/2}[1+(2\kappa)^{1/2}]. \tag{4.112}$$

Combining this with (4.107) and (4.111) ensures that for all $d, N, M \in \mathbb{N}$, $\delta \in (0,1]$ it holds that

$$\begin{aligned}
&\Bigg[\int_{[0,T]\times\mathbb{R}^d}\int_\Omega \Bigg|(\mathcal{R}_\mathfrak{r}(\psi_{d,N,M,\delta,\omega}))(t,x) \\
&\qquad\qquad - \frac{1}{M}\Bigg[\sum_{m=1}^M \big(\mathcal{R}_\mathfrak{r}\big(\mathfrak{f}^0_{d,\min\{(T/N)^{1/2},1\}}\big)\big)\big(Y_t^{N,d,m,x}(\omega)\big)\Bigg]\Bigg|^p \,\mathbb{P}(d\omega)\,\nu_d(dt,dx)\Bigg]^{1/p} \\
&\leq \big[2p\max\{\eta,\kappa\}d^{\max\{\eta,\kappa\}}\max\{T,1\}\max\{\kappa,1\}\big]^{2\kappa+3}\Big[d^{\kappa/2}[1+(2\kappa)^{1/2}]\Big]^{2\kappa+3} \\
&\quad\times \delta d^{(\kappa+1)/2}\kappa d^\kappa e^{(2\kappa+3)\kappa T}2^{2\kappa+4}\,3^\kappa \max\big\{1,[\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\big\} \\
&\leq \delta\big[2p\max\{\eta,\kappa\}\max\{T,1\}\max\{\kappa,1\}\big]^{2\kappa+3}[1+(2\kappa)^{1/2}]^{2\kappa+3}\kappa e^{(2\kappa+3)\kappa T}2^{2\kappa+4}\,3^\kappa \\
&\quad\times d^{(2\kappa+3)\max\{\eta,\kappa\}+\kappa(\kappa+2)+\kappa+(\kappa+1)/2}\max\big\{1,[\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\big\}.
\end{aligned} \tag{4.113}$$

This and (4.104) imply that for all $d, N, M \in \mathbb{N}$, $\delta \in (0,1]$ it holds that

$$\begin{aligned}
&\Bigg[\int_{[0,T]\times\mathbb{R}^d}\int_\Omega \Bigg|(\mathcal{R}_\mathfrak{r}(\psi_{d,N,M,\delta,\omega}))(t,x) \\
&\qquad\qquad - \frac{1}{M}\Bigg[\sum_{m=1}^M \big(\mathcal{R}_\mathfrak{r}\big(\mathfrak{f}^0_{d,\min\{(T/N)^{1/2},1\}}\big)\big)\big(Y_t^{N,d,m,x}(\omega)\big)\Bigg]\Bigg|^p \,\mathbb{P}(d\omega)\,\nu_d(dt,dx)\Bigg]^{1/p} \\
&\leq \mathcal{C}_1\delta\, d^{(2\kappa+3)\max\{\eta,\kappa\}+\kappa(\kappa+2)+\kappa+(\kappa+1)/2}\max\big\{1,[\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\big\}.
\end{aligned} \tag{4.114}$$

Furthermore, observe that (4.94) shows that for all $d, N \in \mathbb{N}$, $m \in \{0,1\}$ it holds that

$$\begin{aligned}
\mathcal{L}\big(\mathfrak{f}^m_{d,\min\{(T/N)^{1/2},1\}}\big) &\leq \mathcal{P}\big(\mathfrak{f}^m_{d,\min\{(T/N)^{1/2},1\}}\big) \leq \kappa d^\kappa[\min\{(T/N)^{1/2},1\}]^{-\kappa} \\
&= \kappa d^\kappa N^{\kappa/2}[\min\{T^{1/2},N^{1/2}\}]^{-\kappa} \leq \kappa d^\kappa N^{\kappa/2}[\min\{T^{1/2},1\}]^{-\kappa} \\
&= \kappa d^\kappa N^{\kappa/2}\max\{T^{-\kappa/2},1\}.
\end{aligned} \tag{4.115}$$

Hence, we obtain that for all $d, N \in \mathbb{N}$, $\delta \in (0,1]$ it holds that

$$\begin{aligned}
&24+6\mathcal{L}\big(\mathfrak{f}^1_{d,\min\{(T/N)^{1/2},1\}}\big)+\big[4+\mathcal{P}\big(\mathfrak{f}^1_{d,\min\{(T/N)^{1/2},1\}}\big)\big]^2 \\
&\leq 24+6\mathcal{P}\big(\mathfrak{f}^1_{d,\min\{(T/N)^{1/2},1\}}\big)+25\big[\mathcal{P}\big(\mathfrak{f}^1_{d,\min\{(T/N)^{1/2},1\}}\big)\big]^2 \\
&\leq 24+31\big[\mathcal{P}\big(\mathfrak{f}^1_{d,\min\{(T/N)^{1/2},1\}}\big)\big]^2 \\
&\leq 55\big[\mathcal{P}\big(\mathfrak{f}^1_{d,\min\{(T/N)^{1/2},1\}}\big)\big]^2 \\
&\leq 2^6\kappa^2 d^{2\kappa}N^\kappa|\max\{T^{-\kappa/2},1\}|^2.
\end{aligned} \tag{4.116}$$

This, (IV), and (4.115) establish that for all $d, N, M \in \mathbb{N}$, $\delta \in (0,1]$, $\omega \in \Omega$ it holds that

$$\begin{aligned}
\mathcal{P}(\psi_{d,N,M,\delta,\omega}) &\leq 2M^2\kappa d^\kappa N^{\kappa/2}\max\{T^{-\kappa/2},1\} \\
&\quad + 9M^2N^6d^{16}\big[2\kappa d^\kappa N^{\kappa/2}\max\{T^{-\kappa/2},1\}+\mathfrak{D}_\delta+2^{12}\kappa^4d^{4\kappa}N^{2\kappa}|\max\{T^{-\kappa/2},1\}|^4\big]^2
\end{aligned}$$

$$
\begin{aligned}
&\le 2M^2\kappa d^\kappa N^{\kappa/2}\max\{T^{-\kappa/2},1\}\\
&\quad + 9M^2N^6d^{16}\big[(2^{12}+3)\mathfrak{D}_\delta|\max\{\kappa,1\}|^4 d^{4\kappa}N^{2\kappa}|\max\{T^{-\kappa/2},1\}|^4\big]^2\\
&\le 2M^2\kappa d^\kappa N^{\kappa/2}\max\{T^{-\kappa/2},1\}\\
&\quad + 2^{26}3^2M^2N^6d^{16}\big[\mathfrak{D}_\delta|\max\{\kappa,1\}|^4 d^{4\kappa}N^{2\kappa}|\max\{T^{-\kappa/2},1\}|^4\big]^2.
\end{aligned} \tag{4.117}
$$

Moreover, observe that (4.102) proves for all $\delta\in(0,1]$ that

$$
\mathfrak{D}_\delta = 2160\log_2(\delta^{-1}) + 8136 \le 8136[\log_2(\delta^{-1})+1] \le 2^{13}[\log_2(\delta^{-1})+1]. \tag{4.118}
$$

Combining this with (4.117) ensures that for all $d,N,M\in\mathbb{N}$, $\delta\in(0,1]$, $\omega\in\Omega$ it holds that

$$
\begin{aligned}
\mathcal{P}(\psi_{d,N,M,\delta,\omega}) &\le 2M^2\kappa d^\kappa N^{\kappa/2}\max\{T^{-\kappa/2},1\}\\
&\qquad + 2^{26}3^2M^2N^{6+4\kappa}d^{16+8\kappa}(\mathfrak{D}_\delta)^2[\max\{\kappa,1\}]^8[\max\{T^{-\kappa/2},1\}]^8\\
&\le 2^{27}3^2M^2N^{6+4\kappa}d^{16+8\kappa}(\mathfrak{D}_\delta)^2[\max\{\kappa,1\}]^8[\max\{T^{-\kappa/2},1\}]^8\\
&\le 2^{31}M^2N^{6+4\kappa}d^{16+8\kappa}(\mathfrak{D}_\delta)^2[\max\{\kappa,1\}]^8[\max\{T^{-\kappa/2},1\}]^8\\
&\le 2^{57}[\max\{\kappa,1\}]^8[\max\{T^{-\kappa/2},1\}]^8M^2N^{6+4\kappa}[\log_2(\delta^{-1})+1]^2d^{16+8\kappa}.
\end{aligned} \tag{4.119}
$$

This and (4.105) prove that for all $d,N,M\in\mathbb{N}$, $\delta\in(0,1]$, $\omega\in\Omega$ it holds that

$$
\mathcal{P}(\psi_{d,N,M,\delta,\omega}) \le \mathcal{C}_2M^2N^{6+4\kappa}[\log_2(\delta^{-1})+1]^2d^{16+8\kappa}. \tag{4.120}
$$

Next note that (I), (II), and, e.g., Beck *et al.* [2, Lemma 2.4] show that for all $d,N,M\in\mathbb{N}$, $\delta\in(0,1]$ it holds that $[0,T]\times\mathbb{R}^d\times\Omega\ni(t,x,\omega)\mapsto(\mathcal{R}_\mathfrak{r}(\psi_{d,N,M,\delta,\omega}))(t,x)\in\mathbb{R}$ is measurable. The triangle inequality and Fubini's theorem hence establish that for all $d,N,M\in\mathbb{N}$, $\delta\in(0,1]$ it holds that

$$
\begin{aligned}
&\left[\int_\Omega\int_{[0,T]\times\mathbb{R}^d}|u_d(y)-(\mathcal{R}_\mathfrak{r}(\psi_{d,N,M,\delta,\omega}))(y)|^p\,\nu_d(dy)\,\mathbb{P}(d\omega)\right]^{1/p}\\
&\le\left(\mathbb{E}\left[\int_{[0,T]\times\mathbb{R}^d}\left|u_d(t,x)-\frac{1}{M}\left[\sum_{m=1}^M\big(\mathcal{R}_\mathfrak{r}\big(\mathfrak{f}^0_{d,\min\{(T/N)^{1/2},1\}}\big)\big)\big(Y_t^{N,d,m,x}\big)\right]\right|^p\nu_d(dt,dx)\right]\right)^{1/p}\\
&\qquad+\left[\int_{[0,T]\times\mathbb{R}^d}\int_\Omega\left|(\mathcal{R}_\mathfrak{r}(\psi_{d,N,M,\delta,\omega}))(t,x)\right.\right.\\
&\qquad\qquad\left.\left.-\frac{1}{M}\left[\sum_{m=1}^M\big(\mathcal{R}_\mathfrak{r}\big(\mathfrak{f}^0_{d,\min\{(T/N)^{1/2},1\}}\big)\big)\big(Y_t^{N,d,m,x}(\omega)\big)\right]\right|^p\mathbb{P}(d\omega)\,\nu_d(dt,dx)\right]^{1/p}.
\end{aligned} \tag{4.121}
$$

Combining this with (4.110) and (4.114) ensures that for all $d,N,M\in\mathbb{N}$, $\delta\in(0,1]$ it holds that

$$
\begin{aligned}
&\left[\int_\Omega\int_{[0,T]\times\mathbb{R}^d}|u_d(y)-(\mathcal{R}_\mathfrak{r}(\psi_{d,N,M,\delta,\omega}))(y)|^p\,\nu_d(dy)\,\mathbb{P}(d\omega)\right]^{1/p}\\
&\le\mathcal{C}_3\left[\frac{d^{\kappa(\kappa+4)+\max\{\eta,\kappa(2\kappa+1)\}}}{N^{1/2}}+\frac{d^{\kappa+\max\{\eta,\kappa^2\}}}{M^{1/2}}\right]\big[\max\big\{1,\nu_d([0,T]\times\mathbb{R}^d)\big\}\big]^{1/p}\\
&\qquad+\mathcal{C}_1\delta\,d^{(2\kappa+3)\max\{\eta,\kappa\}+\kappa(\kappa+2)+\kappa+(\kappa+1)/2}\max\big\{1,[\nu_d([0,T]\times\mathbb{R}^d)]^{1/p}\big\}.
\end{aligned} \tag{4.122}
$$

Hence, we obtain that for all $d, N, M \in \mathbb{N}$, $\delta \in (0,1]$ it holds that

$$\begin{aligned}&\left[\int_\Omega \int_{[0,T]\times\mathbb{R}^d} |u_d(y) - (\mathcal{R}_{\mathfrak{r}}(\psi_{d,N,M,\delta,\omega}))(y)|^p \, \nu_d(dy)\, \mathbb{P}(d\omega)\right]^{1/p} \\ &\le \max\{\mathcal{C}_1, \mathcal{C}_3\} \left[\max\left\{1, \nu_d([0,T]\times\mathbb{R}^d)\right\}\right]^{1/p} \\ &\quad\times \left[\frac{d^{\kappa(\kappa+4)+\max\{\eta,\kappa(2\kappa+1)\}}}{N^{1/2}} + \frac{d^{\kappa+\max\{\eta,\kappa^2\}}}{M^{1/2}} + \delta d^{(2\kappa+3)\max\{\eta,\kappa\}+\kappa(\kappa+2)+\kappa+(\kappa+1)/2}\right].\end{aligned} \tag{4.123}$$

Next note that (I), (II), and, e.g., Beck *et al.* [2, Lemma 2.4] demonstrate that for all $d, N, M \in \mathbb{N}$, $\delta \in (0,1]$ it holds that $[0,T]\times\mathbb{R}^d\times\Omega \ni (t,x,\omega) \mapsto (\mathcal{R}_{\mathfrak{r}}(\psi_{d,N,M,\delta,\omega}))(t,x) \in \mathbb{R}$ is measurable. Combining this with Fubini's theorem, (4.109), (4.123), and the fact that $\forall\, d \in \mathbb{N}\colon u_d \in C([0,T]\times\mathbb{R}^d, \mathbb{R})$ proves that

A) it holds for all $d, N, M \in \mathbb{N}$, $\delta \in (0,1]$ that $Z_{d,N,M,\delta}$ is a random variable,

B) it holds for all $d, N, M \in \mathbb{N}$, $\delta \in (0,1]$ that

$$\begin{aligned}\mathbb{E}\big[|Z_{d,N,M,\delta}|\big] &\le [\max\{\mathcal{C}_1, \mathcal{C}_3\}]^p \max\left\{1, \nu_d([0,T]\times\mathbb{R}^d)\right\} \\ &\quad\times \left[\frac{d^{\kappa(\kappa+4)+\max\{\eta,\kappa(2\kappa+1)\}}}{N^{1/2}} + \frac{d^{\kappa+\max\{\eta,\kappa^2\}}}{M^{1/2}}\right. \\ &\qquad\left. + \delta d^{(2\kappa+3)\max\{\eta,\kappa\}+\kappa(\kappa+2)+\kappa+(\kappa+1)/2}\right]^p.\end{aligned} \tag{4.124}$$

This and, e.g., [39, Lemma 2.1] prove that there exist $\mathfrak{w}_{d,N,M,\delta} \in \Omega$, $d, N, M \in \mathbb{N}$, $\delta \in (0,1]$, which satisfy that for all $d, N, M \in \mathbb{N}$, $\delta \in (0,1]$ it holds that

$$\begin{aligned}&\int_{[0,T]\times\mathbb{R}^d} \left|u_d(y) - (\mathcal{R}_{\mathfrak{r}}(\psi_{d,N,M,\delta,\mathfrak{w}_{d,N,M,\delta}}))(y)\right|^p \nu_d(dy) \\ &= Z_{d,N,M,\delta}(\mathfrak{w}_{d,N,M,\delta}) \\ &\le [\max\{\mathcal{C}_1, \mathcal{C}_3\}]^p \max\left\{1, \nu_d([0,T]\times\mathbb{R}^d)\right\} \\ &\quad\times \left[\frac{d^{\kappa(\kappa+4)+\max\{\eta,\kappa(2\kappa+1)\}}}{N^{1/2}} + \frac{d^{\kappa+\max\{\eta,\kappa^2\}}}{M^{1/2}} + \delta d^{(2\kappa+3)\max\{\eta,\kappa\}+\kappa(\kappa+2)+\kappa+(\kappa+1)/2}\right]^p \\ &= [\max\{\mathcal{C}_1, \mathcal{C}_3\}]^p \max\left\{1, \nu_d([0,T]\times\mathbb{R}^d)\right\} \\ &\quad\times \left[\frac{d^{\kappa(\kappa+4)+\max\{\eta,\kappa(2\kappa+1)\}}}{N^{1/2}} + \frac{d^{\kappa+\max\{\eta,\kappa^2\}}}{M^{1/2}} + \delta d^{(2\kappa+3)\max\{\eta,\kappa\}+\kappa^2+(7\kappa+1)/2}\right]^p.\end{aligned} \tag{4.125}$$

Combining this, (I), and (4.120) establishes items (i)-(iii). This completes the proof of Proposition 4.3. □

4.6. Cost estimates for deep ANNs

Theorem 4.1. *Let $T, \kappa, \eta, c \in (0,\infty)$, $p \in [2,\infty)$ satisfy that*

$$c = 18 + 12\kappa + 4\max\{\eta, \kappa^2\} + 4\eta + [2\kappa(\kappa+4) + 2\max\{\eta, \kappa(2\kappa+1)\} + 2\eta](6+4\kappa), \tag{4.126}$$

for every $d \in \mathbb{N}$ let $A_d = (a_{d,i,j})_{(i,j)\in\{1,2,\dots,d\}^2} \in \mathbb{R}^{d\times d}$ be a symmetric positive semidefinite matrix, for every $d \in \mathbb{N}$ let $\nu_d \colon \mathcal{B}([0,T]\times\mathbb{R}^d) \to [0,\infty)$ be a finite measure which satisfies that

$$\left[\max\left\{1, \nu_d([0,T]\times\mathbb{R}^d), \int_{[0,T]\times\mathbb{R}^d} \|x\|^{2p\max\{2\kappa,3\}}\,\nu_d(dt,dx)\right\}\right]^{1/p} \leq \eta d^\eta, \tag{4.127}$$

for every $m \in \{0,1\}$, $d \in \mathbb{N}$ let $f_d^m \colon \mathbb{R}^d \to \mathbb{R}^{md-m+1}$ be a function, let $(\mathfrak{f}^m_{d,\varepsilon})_{(m,d,\varepsilon)\in\{0,1\}\times\mathbb{N}\times(0,1]} \subseteq \mathbf{N}$, assume for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$, $m \in \{0,1\}$, $x, y \in \mathbb{R}^d$ that

$$\mathcal{R}_\mathfrak{r}(\mathfrak{f}^0_{d,\varepsilon}) \in C(\mathbb{R}^d,\mathbb{R}), \quad \mathcal{R}_\mathfrak{r}(\mathfrak{f}^1_{d,\varepsilon}) \in C(\mathbb{R}^d,\mathbb{R}^d), \quad \mathcal{P}(\mathfrak{f}^m_{d,\varepsilon}) \leq \kappa d^\kappa \varepsilon^{-\kappa}, \tag{4.128}$$

$$\left\|f_d^1(x) - f_d^1(y)\right\| \leq \kappa\|x-y\|, \quad \left\|(\mathcal{R}_\mathfrak{r}(\mathfrak{f}^1_{d,\varepsilon}))(x)\right\| \leq \kappa(d^\kappa + \|x\|), \tag{4.129}$$

$$\left|(\mathcal{R}_\mathfrak{r}(\mathfrak{f}^0_{d,\varepsilon}))(x) - (\mathcal{R}_\mathfrak{r}(\mathfrak{f}^0_{d,\varepsilon}))(y)\right| \leq \kappa d^\kappa(1+\|x\|^\kappa+\|y\|^\kappa)\|x-y\|, \tag{4.130}$$

$$\left\|f_d^m(x) - (\mathcal{R}_\mathfrak{r}(\mathfrak{f}^m_{d,\varepsilon}))(x)\right\| \leq \varepsilon\kappa d^\kappa(1+\|x\|^\kappa), \tag{4.131}$$

$$\left|f_d^0(x)\right| + \operatorname{Trace}(A_d) \leq \kappa d^\kappa(1+\|x\|^\kappa), \tag{4.132}$$

and for every $d \in \mathbb{N}$ let $u_d \in \{v \in C([0,T]\times\mathbb{R}^d,\mathbb{R}) \colon \inf_{q\in(0,\infty)} \sup_{(t,y)\in[0,T]\times\mathbb{R}^d} \frac{|v(t,y)|}{1+\|y\|^q} < \infty\}$ be a viscosity solution of

$$\left(\frac{\partial}{\partial t}u_d\right)(t,x) = \left(\frac{\partial}{\partial x}u_d\right)(t,x)\, f_d^1(x) + \sum_{i,j=1}^d a_{d,i,j}\left(\frac{\partial^2}{\partial x_i \partial x_j}u_d\right)(t,x) \tag{4.133}$$

with $u_d(0,x) = f_d^0(x)$ for $(t,x) \in (0,T)\times\mathbb{R}^d$ (cf. Definitions 2.1, 4.1, 4.3, and 4.4). Then there exist $\mathfrak{C} \in \mathbb{R}$ and $(\mathfrak{u}_{d,\varepsilon})_{d\in\mathbb{N},\varepsilon\in(0,1]} \subseteq \mathbf{N}$ such that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ it holds that $\mathcal{R}_\mathfrak{r}(\mathfrak{u}_{d,\varepsilon}) \in C(\mathbb{R}^{d+1},\mathbb{R})$, $\mathcal{P}(\mathfrak{u}_{d,\varepsilon}) \leq \mathfrak{C}\varepsilon^{-(18+8\kappa)}d^c$, and

$$\left[\int_{[0,T]\times\mathbb{R}^d} |u_d(y) - (\mathcal{R}_\mathfrak{r}(\mathfrak{u}_{d,\varepsilon}))(y)|^p\,\nu_d(dy)\right]^{1/p} \leq \varepsilon. \tag{4.134}$$

Proof. Note that (4.127) implies that

$$\left[\max\left\{1, \nu_1([0,T]\times\mathbb{R}^d), \int_{[0,T]\times\mathbb{R}} \|x\|^{2p\max\{2\kappa,3\}}\,\nu_1(dt,dx)\right\}\right]^{1/p} \leq \eta. \tag{4.135}$$

This proves that $\eta \in [1,\infty)$. Proposition 4.3 hence ensures that there exist $\mathcal{C}_1 \in (0,\infty)$ and $(\Phi_{d,N,M,\delta})_{(d,N,M,\delta)\in\mathbb{N}^3\times(0,1]} \subseteq \mathbf{N}$ which satisfy that

(I) it holds for all $d, N, M \in \mathbb{N}$, $\delta \in (0,1]$ that $\mathcal{R}_\mathfrak{r}(\Phi_{d,N,M,\delta}) \in C(\mathbb{R}^{d+1},\mathbb{R})$,

(II) it holds for all $d, N, M \in \mathbb{N}$, $\delta \in (0,1]$ that

$$\begin{aligned}
&\left[\int_{[0,T]\times\mathbb{R}^d} |u_d(y) - (\mathcal{R}_\mathfrak{r}(\Phi_{d,N,M,\delta}))(y)|^p\,\nu_d(dy)\right]^{1/p} \\
&\leq \mathcal{C}_1\left[\max\left\{1, \nu_d([0,T]\times\mathbb{R}^d)\right\}\right]^{1/p} \\
&\quad\times\left[\frac{d^{\kappa(\kappa+4)+\max\{\eta,\kappa(2\kappa+1)\}}}{N^{1/2}} + \frac{d^{\kappa+\max\{\eta,\kappa^2\}}}{M^{1/2}} + \delta d^{(2\kappa+3)\max\{\eta,\kappa\}+\kappa^2+(7\kappa+1)/2}\right],
\end{aligned} \tag{4.136}$$

(III) it holds for all $d, N, M \in \mathbb{N}$, $\delta \in (0,1]$ that

$$\mathcal{P}(\Phi_{d,N,M,\delta}) \leq \mathcal{C}_1 M^2 N^{6+4\kappa} [\log_2(\delta^{-1}) + 1]^2 d^{16+8\kappa}. \tag{4.137}$$

Next let $\mathcal{C}_2, \mathfrak{C} \in (0,\infty)$ satisfy that

$$\mathcal{C}_2 = \max\{0, \log_2(3\mathcal{C}_1\eta)\} + \frac{1}{\ln(2)} + \frac{1}{\ln(2)} \left[2(\kappa+2)\max\{\eta,\kappa\} + \kappa^2 + \frac{(7\kappa+1)}{2}\right] \tag{4.138}$$

and

$$\mathfrak{C} = \mathcal{C}_1 2^{8+4\kappa} (3\mathcal{C}_1\eta)^{16+8\kappa} [\mathcal{C}_2 + 1]^2, \tag{4.139}$$

let $\mathcal{D}_{d,\varepsilon} \in (0,1]$, $\varepsilon \in (0,1]$, $d \in \mathbb{N}$, satisfy for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ that

$$\mathcal{D}_{d,\varepsilon} = \min\left\{1, (3\mathcal{C}_1\eta)^{-1} \varepsilon d^{-2(\kappa+2)\max\{\eta,\kappa\} - \kappa^2 - (7\kappa+1)/2}\right\}, \tag{4.140}$$

let $\mathfrak{M}_{d,\varepsilon} \in \mathbb{N}$, $\varepsilon \in (0,1]$, $d \in \mathbb{N}$, satisfy for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ that

$$\mathfrak{M}_{d,\varepsilon} = \min\left[\mathbb{N} \cap \left[(3\mathcal{C}_1\eta)^2 \varepsilon^{-2} d^{2\kappa+2\max\{\eta,\kappa^2\}+2\eta}, \infty\right)\right], \tag{4.141}$$

let $\mathfrak{N}_{d,\varepsilon} \in \mathbb{N}$, $\varepsilon \in (0,1]$, $d \in \mathbb{N}$, satisfy for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ that

$$\mathfrak{N}_{d,\varepsilon} = \min\left[\mathbb{N} \cap \left[(3\mathcal{C}_1\eta)^2 \varepsilon^{-2} d^{2\kappa(\kappa+4)+2\max\{\eta,\kappa(2\kappa+1)\}+2\eta}, \infty\right)\right], \tag{4.142}$$

and let $\mathfrak{u}_{d,\varepsilon} \in \mathbf{N}$, $\varepsilon \in (0,1]$, $d \in \mathbb{N}$, satisfy for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ that

$$\mathfrak{u}_{d,\varepsilon} = \Phi_{d,\mathfrak{N}_{d,\varepsilon},\mathfrak{M}_{d,\varepsilon},\mathcal{D}_{d,\varepsilon}}. \tag{4.143}$$

Observe that (4.127) and (4.136) ensure that for all $d, N, M \in \mathbb{N}$, $\delta \in (0,1]$ it holds that

$$\begin{aligned} &\left[\int_{[0,T]\times\mathbb{R}^d} |u_d(y) - (\mathcal{R}_\mathfrak{r}(\Phi_{d,N,M,\delta}))(y)|^p \, \nu_d(dy)\right]^{1/p} \\ &\leq \mathcal{C}_1\eta \left[\frac{d^{\kappa(\kappa+4)+\max\{\eta,\kappa(2\kappa+1)\}+\eta}}{N^{1/2}} + \frac{d^{\kappa+\max\{\eta,\kappa^2\}+\eta}}{M^{1/2}} + \delta d^{2(\kappa+2)\max\{\eta,\kappa\}+\kappa^2+(7\kappa+1)/2}\right]. \end{aligned} \tag{4.144}$$

Combining this, (4.143), (4.140), (4.141), and (4.142) implies that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ it holds that

$$\begin{aligned} &\left[\int_{[0,T]\times\mathbb{R}^d} |u_d(y) - (\mathcal{R}_\mathfrak{r}(\mathfrak{u}_{d,\varepsilon}))(y)|^p \, \nu_d(dy)\right]^{1/p} \\ &= \left[\int_{[0,T]\times\mathbb{R}^d} \left|u_d(y) - (\mathcal{R}_\mathfrak{r}(\Phi_{d,\mathfrak{N}_{d,\varepsilon},\mathfrak{M}_{d,\varepsilon},\mathcal{D}_{d,\varepsilon}}))(y)\right|^p \, \nu_d(dy)\right]^{1/p} \\ &\leq \mathcal{C}_1\eta \left[\frac{d^{\kappa(\kappa+4)+\max\{\eta,\kappa(2\kappa+1)\}+\eta}}{|\mathfrak{N}_{d,\varepsilon}|^{1/2}} + \frac{d^{\kappa+\max\{\eta,\kappa^2\}+\eta}}{|\mathfrak{M}_{d,\varepsilon}|^{1/2}} + \mathcal{D}_{d,\varepsilon} d^{2(\kappa+2)\max\{\eta,\kappa\}+\kappa^2+(7\kappa+1)/2}\right] \\ &\leq \mathcal{C}_1\eta \left[\frac{d^{\kappa(\kappa+4)+\max\{\eta,\kappa(2\kappa+1)\}+\eta}}{\left|(3\mathcal{C}_1\eta)^2\varepsilon^{-2}d^{2\kappa(\kappa+4)+2\max\{\eta,\kappa(2\kappa+1)\}+2\eta}\right|^{1/2}} + \frac{d^{\kappa+\max\{\eta,\kappa^2\}+\eta}}{\left|(3\mathcal{C}_1\eta)^2\varepsilon^{-2}d^{2\kappa+2\max\{\eta,\kappa^2\}+2\eta}\right|^{1/2}}\right. \\ &\qquad \left. + (3\mathcal{C}_1\eta)^{-1}\varepsilon d^{-2(\kappa+2)\max\{\eta,\kappa\}-\kappa^2-(7\kappa+1)/2} d^{2(\kappa+2)\max\{\eta,\kappa\}+\kappa^2+(7\kappa+1)/2}\right] \\ &= \mathcal{C}_1\eta \left(\frac{1}{3\mathcal{C}_1\eta\varepsilon^{-1}} + \frac{1}{3\mathcal{C}_1\eta\varepsilon^{-1}} + \frac{\varepsilon}{3\mathcal{C}_1\eta}\right) = \varepsilon. \end{aligned} \tag{4.145}$$

In addition, observe that (4.138), (4.140), and the fact that $\forall\, x \in [1,\infty)\colon \log_2(x) = \frac{\ln(x)}{\ln(2)} \leq \frac{x}{\ln(2)}$ demonstrate that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ it holds that

$$\begin{aligned}
&\log_2((\mathcal{D}_{d,\varepsilon})^{-1}) \\
&= \max\left\{0, \log_2\left(3\mathcal{C}_1\eta\,\varepsilon^{-1}\, d^{2(\kappa+2)\max\{\eta,\kappa\}+\kappa^2+(7\kappa+1)/2}\right)\right\} \\
&\leq \max\left\{0, \log_2(3\mathcal{C}_1\eta) + \log_2(\varepsilon^{-1}) + \left[2(\kappa+2)\max\{\eta,\kappa\} + \kappa^2 + \frac{(7\kappa+1)}{2}\right]\log_2(d)\right\} \\
&\leq \max\{0, \log_2(3\mathcal{C}_1\eta)\} + \frac{1}{\ln(2)}\,\varepsilon^{-1} + \frac{1}{\ln(2)}\left[2(\kappa+2)\max\{\eta,\kappa\} + \kappa^2 + \frac{(7\kappa+1)}{2}\right] d \\
&\leq \varepsilon^{-1} d\left(\max\{0, \log_2(3\mathcal{C}_1\eta)\} + \frac{1}{\ln(2)} + \frac{1}{\ln(2)}\left[2(\kappa+2)\max\{\eta,\kappa\} + \kappa^2 + \frac{(7\kappa+1)}{2}\right]\right) \\
&= \varepsilon^{-1} d\,\mathcal{C}_2.
\end{aligned} \tag{4.146}$$

Combining this with (III), (4.141), (4.142), and (4.143) proves that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ it holds that

$$\begin{aligned}
\mathcal{P}(\mathfrak{u}_{d,\varepsilon}) &= \mathcal{P}(\Phi_{d,\mathfrak{N}_{d,\varepsilon},\mathfrak{M}_{d,\varepsilon},\mathcal{D}_{d,\varepsilon}}) \\
&\leq \mathcal{C}_1(\mathfrak{M}_{d,\varepsilon})^2 d^{16+8\kappa}(\mathfrak{N}_{d,\varepsilon})^{6+4\kappa}[\log_2((\mathcal{D}_{d,\varepsilon})^{-1}) + 1]^2 \\
&\leq \mathcal{C}_1\left[(3\mathcal{C}_1\eta)^2\varepsilon^{-2}d^{2\kappa+2\max\{\eta,\kappa^2\}+2\eta} + 1\right]^2 d^{16+8\kappa}[\varepsilon^{-1}d\,\mathcal{C}_2 + 1]^2 \\
&\quad\times\left[(3\mathcal{C}_1\eta)^2\varepsilon^{-2}d^{2\kappa(\kappa+4)+2\max\{\eta,\kappa(2\kappa+1)\}+2\eta} + 1\right]^{6+4\kappa} \\
&\leq \mathcal{C}_1 2^{8+4\kappa}(3\mathcal{C}_1\eta)^4\varepsilon^{-4}d^{4\kappa+4\max\{\eta,\kappa^2\}+4\eta}d^{16+8\kappa}\varepsilon^{-2}d^2[\mathcal{C}_2+1]^2 \\
&\quad\times(3\mathcal{C}_1\eta)^{12+8\kappa}\varepsilon^{-(12+8\kappa)}d^{[2\kappa(\kappa+4)+2\max\{\eta,\kappa(2\kappa+1)\}+2\eta](6+4\kappa)}.
\end{aligned} \tag{4.147}$$

This, (4.126), and (4.139) ensure that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ it holds that

$$\begin{aligned}
\mathcal{P}(\mathfrak{u}_{d,\varepsilon}) &\leq \mathcal{C}_1 2^{8+4\kappa}(3\mathcal{C}_1\eta)^{16+8\kappa}\varepsilon^{-(18+8\kappa)}[\mathcal{C}_2+1]^2 \\
&\quad\times d^{2+4\kappa+4\max\{\eta,\kappa^2\}+4\eta+16+8\kappa+[2\kappa(\kappa+4)+2\max\{\eta,\kappa(2\kappa+1)\}+2\eta](6+4\kappa)} \\
&= \mathfrak{C}\varepsilon^{-(18+8\kappa)}d^{18+12\kappa+4\max\{\eta,\kappa^2\}+4\eta+[2\kappa(\kappa+4)+2\max\{\eta,\kappa(2\kappa+1)\}+2\eta](6+4\kappa)} \\
&= \mathfrak{C}\varepsilon^{-(18+8\kappa)}d^{c}.
\end{aligned} \tag{4.148}$$

Combining this, (I), and (4.145) establishes (4.134). This completes the proof of Theorem 4.1.□

Corollary 4.2. *Let $T,\kappa,\eta \in (0,\infty)$, $p \in [2,\infty)$, let $A_d = (a_{d,i,j})_{(i,j)\in\{1,2,\ldots,d\}^2} \in \mathbb{R}^{d\times d}$, $d \in \mathbb{N}$, be symmetric positive semidefinite matrices, let $\nu_d\colon \mathcal{B}([0,T]\times\mathbb{R}^d) \to [0,\infty)$, $d \in \mathbb{N}$, be finite measures which satisfy for all $d \in \mathbb{N}$ that*

$$\left[\max\left\{1, \nu_d([0,T]\times\mathbb{R}^d), \int_{[0,T]\times\mathbb{R}^d}\|x\|^{2p\max\{6\kappa,2\kappa+2,3\}}\,\nu_d(dt,dx)\right\}\right]^{1/p} \leq \eta d^{\eta}, \tag{4.149}$$

let $f_d^m\colon \mathbb{R}^d \to \mathbb{R}^{md-m+1}$, $m \in \{0,1\}$, $d \in \mathbb{N}$, be functions, let $(\mathfrak{f}^m_{d,\varepsilon})_{(m,d,\varepsilon)\in\{0,1\}\times\mathbb{N}\times(0,1]} \subseteq \mathbf{N}$, assume for all $m \in \{0,1\}$, $d \in \mathbb{N}$, $\varepsilon \in (0,1]$, $i \in \{1,2,\ldots,d\}$, $x,y \in \mathbb{R}^d$ that

$$\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0_{d,\varepsilon}) \in C(\mathbb{R}^d,\mathbb{R}), \quad \mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^1_{d,\varepsilon}) \in C(\mathbb{R}^d,\mathbb{R}^d), \quad \mathcal{P}(\mathfrak{f}^m_{d,\varepsilon}) \leq \kappa d^{\kappa}\varepsilon^{-\kappa}, \tag{4.150}$$

$$\left\|f^1_d(x) - f^1_d(y)\right\| \leq \kappa\|x-y\|, \quad \left\|(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^1_{d,\varepsilon}))(x)\right\| \leq \kappa(d^{\kappa} + \|x\|), \tag{4.151}$$

$$\varepsilon|f^0_d(x)| + \varepsilon|a_{d,i,i}| + \|f^m_d(x) - (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^m_{d,\varepsilon}))(x)\| \leq \varepsilon\kappa d^{\kappa}(1 + \|x\|^{\kappa}), \tag{4.152}$$

$$\left|(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0_{d,\varepsilon}))(x) - (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0_{d,\varepsilon}))(y)\right| \leq \kappa d^{\kappa}(1 + \|x\|^{\kappa} + \|y\|^{\kappa})\|x-y\|, \tag{4.153}$$

and for every $d \in \mathbb{N}$ let $u_d \in \{v \in C([0,T] \times \mathbb{R}^d, \mathbb{R})\colon \inf_{q\in(0,\infty)} \sup_{(t,y)\in[0,T]\times\mathbb{R}^d} \frac{|v(t,y)|}{1+\|y\|^q} < \infty\}$ be a viscosity solution of

$$\left(\frac{\partial}{\partial t}u_d\right)(t,x) = \left(\frac{\partial}{\partial x}u_d\right)(t,x)\, f_d^1(x) + \sum_{i,j=1}^{d} a_{d,i,j}\left(\frac{\partial^2}{\partial x_i \partial x_j}u_d\right)(t,x) \tag{4.154}$$

with $u_d(0,x) = f_d^0(x)$ for $(t,x) \in (0,T) \times \mathbb{R}^d$ (cf. Definitions 2.1, 4.1, 4.3, and 4.4). Then there exist $\mathfrak{C} \in \mathbb{R}$ and $(\mathfrak{u}_{d,\varepsilon})_{d\in\mathbb{N},\varepsilon\in(0,1]} \subseteq \mathbf{N}$ such that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ it holds that $\mathcal{R}_\mathfrak{r}(\mathfrak{u}_{d,\varepsilon}) \in C(\mathbb{R}^{d+1},\mathbb{R})$, $\mathcal{P}(\mathfrak{u}_{d,\varepsilon}) \le \mathfrak{C}\varepsilon^{-\mathfrak{C}}d^{\mathfrak{C}}$, and

$$\left[\int_{[0,T]\times\mathbb{R}^d} |u_d(y) - (\mathcal{R}_\mathfrak{r}(\mathfrak{u}_{d,\varepsilon}))(y)|^p\, \nu_d(dy)\right]^{1/p} \le \varepsilon. \tag{4.155}$$

Proof. Throughout this proof let $\iota = \max\{3\kappa, \kappa+1\}$. Observe that (4.151) and the fact that $\iota \ge \kappa$ prove that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$, $x, y \in \mathbb{R}^d$ it holds that

$$\left\|f_d^1(x) - f_d^1(y)\right\| \le \iota\|x-y\| \quad \text{and} \quad \left\|(\mathcal{R}_\mathfrak{r}(\mathfrak{f}_{d,\varepsilon}^1))(x)\right\| \le \iota(d^\iota + \|x\|). \tag{4.156}$$

Next note that (4.153) and the fact that $\iota \ge 3\kappa$ ensure that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$, $x, y \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
&\left|(\mathcal{R}_\mathfrak{r}(\mathfrak{f}_{d,\varepsilon}^0))(x) - (\mathcal{R}_\mathfrak{r}(\mathfrak{f}_{d,\varepsilon}^0))(y)\right| \\
&\le \kappa d^\iota(3 + \|x\|^\iota + \|y\|^\iota)\|x-y\| \\
&\le \iota d^\iota(1 + \|x\|^\iota + \|y\|^\iota)\|x-y\|.
\end{aligned} \tag{4.157}$$

In addition, observe that (4.152) and the fact that $\iota \ge 2\kappa$ imply that for all $m \in \{0,1\}$, $d \in \mathbb{N}$, $\varepsilon \in (0,1]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
&\left\|f_d^m(x) - (\mathcal{R}_\mathfrak{r}(\mathfrak{f}_{d,\varepsilon}^m))(x)\right\| \\
&\le \varepsilon\kappa d^\kappa(1 + \|x\|^\kappa) \le \varepsilon\kappa d^\iota(2 + \|x\|^\iota) \le \varepsilon\iota d^\iota(1 + \|x\|^\iota).
\end{aligned} \tag{4.158}$$

Moreover, note that (4.152) and the fact that $\iota \ge \max\{2\kappa, \kappa+1\}$ demonstrate that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
\left|f_d^0(x)\right| + \operatorname{Trace}(A_d) &\le \kappa d^{\kappa+1}(1 + \|x\|^\kappa) \le \kappa d^{\kappa+1}(2 + \|x\|^\iota) \\
&\le 2\kappa d^{\kappa+1}(1 + \|x\|^\iota) \le \iota d^\iota(1 + \|x\|^\iota).
\end{aligned} \tag{4.159}$$

Furthermore, observe that (4.150) implies that for all $m \in \{0,1\}$, $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ it holds that $\mathcal{P}(\mathfrak{f}_{d,\varepsilon}^m) \le \iota d^\iota \varepsilon^{-\iota}$. Combining this, (4.149), (4.156), (4.157), (4.158), (4.159), and Theorem 4.1 establishes that there exist $\mathcal{C} \in \mathbb{R}$ and $(\mathfrak{u}_{d,\varepsilon})_{d\in\mathbb{N},\varepsilon\in(0,1]} \subseteq \mathbf{N}$ which satisfy that

(I) it holds for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ that $\mathcal{R}_\mathfrak{r}(\mathfrak{u}_{d,\varepsilon}) \in C(\mathbb{R}^{d+1},\mathbb{R})$,

(II) it holds for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ that

$$\left[\int_{[0,T]\times\mathbb{R}^d} |u_d(y) - (\mathcal{R}_\mathfrak{r}(\mathfrak{u}_{d,\varepsilon}))(y)|^p\, \nu_d(dy)\right]^{1/p} \le \varepsilon, \tag{4.160}$$

(III) it holds for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ that

$$\mathcal{P}(\mathfrak{u}_{d,\varepsilon}) \le \mathcal{C}\varepsilon^{-(18+8\iota)}d^{18+12\iota+4\max\{\eta,\iota^2\}+4\eta+[2\iota(\iota+4)+2\max\{\eta,\iota(2\iota+1)\}+2\eta](6+4\iota)}. \tag{4.161}$$

This proves that for all $\mathfrak{C} \in [\max\{\mathcal{C}, 18 + 12\iota + 4\max\{\eta, \iota^2\} + 4\eta + [2\iota(\iota+4) + 2\max\{\eta, \iota(2\iota+1)\} + 2\eta](6+4\iota)\}, \infty)$, $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ it holds that $\mathcal{P}(\mathfrak{u}_{d,\varepsilon}) \leq \mathfrak{C}\varepsilon^{-\mathfrak{C}} d^{\mathfrak{C}}$. Combining this, (I), and (II) establishes (4.155). This completes the proof of Corollary 4.2. □

Corollary 4.3. *Let $T, \kappa \in (0,\infty)$, $\alpha \in \mathbb{R}$, $\beta \in (\alpha, \infty)$, let $A_d = (a_{d,i,j})_{(i,j)\in\{1,2,\ldots,d\}^2} \in \mathbb{R}^{d\times d}$, $d \in \mathbb{N}$, be symmetric positive semidefinite matrices, let $f_d^m \colon \mathbb{R}^d \to \mathbb{R}^{md-m+1}$, $m \in \{0,1\}$, $d \in \mathbb{N}$, be functions, let $(\mathfrak{f}_{d,\varepsilon}^m)_{(m,d,\varepsilon)\in\{0,1\}\times\mathbb{N}\times(0,1]} \subseteq \mathbf{N}$, assume for all $m \in \{0,1\}$, $d \in \mathbb{N}$, $\varepsilon \in (0,1]$, $i \in \{1,2,\ldots,d\}$, $x, y \in \mathbb{R}^d$ that*

$$\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}_{d,\varepsilon}^0) \in C(\mathbb{R}^d, \mathbb{R}), \quad \mathcal{R}_{\mathfrak{r}}(\mathfrak{f}_{d,\varepsilon}^1) \in C(\mathbb{R}^d, \mathbb{R}^d), \quad \mathcal{P}(\mathfrak{f}_{d,\varepsilon}^m) \leq \kappa d^\kappa \varepsilon^{-\kappa}, \tag{4.162}$$

$$\left\|f_d^1(x) - f_d^1(y)\right\| \leq \kappa\|x-y\|, \quad \left\|(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}_{d,\varepsilon}^1))(x)\right\| \leq \kappa(d^\kappa + \|x\|), \tag{4.163}$$

$$\varepsilon\left|f_d^0(x)\right| + \varepsilon|a_{d,i,i}| + \left\|f_d^m(x) - (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}_{d,\varepsilon}^m))(x)\right\| \leq \varepsilon\kappa d^\kappa(1 + \|x\|^\kappa), \tag{4.164}$$

$$\left|(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}_{d,\varepsilon}^0))(x) - (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}_{d,\varepsilon}^0))(y)\right| \leq \kappa d^\kappa(1 + \|x\|^\kappa + \|y\|^\kappa)\|x-y\|, \tag{4.165}$$

and for every $d \in \mathbb{N}$ let $u_d \in \{v \in C([0,T]\times\mathbb{R}^d, \mathbb{R}) \colon \inf_{q\in(0,\infty)} \sup_{(t,y)\in[0,T]\times\mathbb{R}^d} \frac{|v(t,y)|}{1+\|y\|^q} < \infty\}$ be a viscosity solution of

$$\left(\frac{\partial}{\partial t}u_d\right)(t,x) = \left(\frac{\partial}{\partial x}u_d\right)(t,x)\, f_d^1(x) + \sum_{i,j=1}^d a_{d,i,j}\left(\frac{\partial^2}{\partial x_i \partial x_j}u_d\right)(t,x) \tag{4.166}$$

with $u_d(0,x) = f_d^0(x)$ for $(t,x) \in (0,T)\times\mathbb{R}^d$ (cf. Definitions 2.1, 4.1, 4.3, and 4.4). Then for every $p \in (0,\infty)$ there exist $c \in \mathbb{R}$ and $(\mathfrak{u}_{d,\varepsilon})_{d\in\mathbb{N},\varepsilon\in(0,1]} \subseteq \mathbf{N}$ such that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ it holds that $\mathcal{R}_{\mathfrak{r}}(\mathfrak{u}_{d,\varepsilon}) \in C(\mathbb{R}^{d+1}, \mathbb{R})$, $\mathcal{P}(\mathfrak{u}_{d,\varepsilon}) \leq c\varepsilon^{-c}d^c$, and

$$\left[\int_{[0,T]\times[\alpha,\beta]^d} \frac{|u_d(y) - (\mathcal{R}_{\mathfrak{r}}(\mathfrak{u}_{d,\varepsilon}))(y)|^p}{|\beta-\alpha|^d}\,dy\right]^{1/p} \leq \varepsilon. \tag{4.167}$$

Proof. Throughout this proof let $p, q, \eta \in (0,\infty)$ satisfy that $q = \max\{p, 2\}$ and $\eta = \max\{6\kappa, 2\kappa+2, 3\} + [\max\{1,T\}]^{1/q}\max\{1, |\alpha|^{2\max\{6\kappa,2\kappa+2,3\}}, |\beta|^{2\max\{6\kappa,2\kappa+2,3\}}\}$, for every $d \in \mathbb{N}$ let $\mu_d \colon \mathcal{B}([0,T]\times\mathbb{R}^d) \to [0,\infty]$ be the Lebesgue-Borel measure on $[0,T]\times\mathbb{R}^d$, for every $d \in \mathbb{N}$ let $\nu_d \colon \mathcal{B}([0,T]\times\mathbb{R}^d) \to [0,\infty]$ be the measure which satisfies for all $d \in \mathbb{N}$, $B_1 \in \mathcal{B}([0,T])$, $B_2 \in \mathcal{B}(\mathbb{R}^d)$ that

$$\nu_d(B_1 \times B_2) = \frac{\mu_d(B_1 \times ([\alpha,\beta]^d \cap B_2))}{|\beta-\alpha|^d}, \tag{4.168}$$

let $\delta \colon (0,1] \to (0,1]$ satisfy for all $\varepsilon \in (0,1]$ that $\delta(\varepsilon) = \varepsilon[\max\{T,1\}]^{1/q-1/p}$, and let $r \colon (0,\infty) \to (0,\infty)$ satisfy for all $z \in (0,\infty)$ that $r(z) = z[\max\{T,1\}]^{z(1/p-1/q)}$. Observe that (4.168), Fubini's theorem, and, e.g., Grohs *et al.* [25, Lemma 3.15] prove that for all $d \in \mathbb{N}$ it holds that

$$\begin{aligned}
&\int_{[0,T]\times\mathbb{R}^d} \|x\|^{2q\max\{6\kappa,2\kappa+2,3\}}\,\nu_d(dt,dx) \\
&= \int_{[0,T]\times[\alpha,\beta]^d} \frac{\|x\|^{2q\max\{6\kappa,2\kappa+2,3\}}}{|\beta-\alpha|^d}\,\mu_d(dt,dx) \\
&= T\int_{[\alpha,\beta]^d} \frac{\|x\|^{2q\max\{6\kappa,2\kappa+2,3\}}}{|\beta-\alpha|^d}\,dx \\
&\leq Td^{q\max\{6\kappa,2\kappa+2,3\}}\max\left\{|\alpha|^{2q\max\{6\kappa,2\kappa+2,3\}}, |\beta|^{2q\max\{6\kappa,2\kappa+2,3\}}\right\}.
\end{aligned} \tag{4.169}$$

Therefore, we obtain for all $d \in \mathbb{N}$ that

$$\begin{aligned}&\left[\max\left\{1, \nu_d([0,T]\times\mathbb{R}^d), \int_{[0,T]\times\mathbb{R}^d} \|x\|^{2q\max\{6\kappa,2\kappa+2,3\}}\,\nu_d(dt,dx)\right\}\right]^{1/q} \\ &\leq \left[\max\left\{1, T, Td^{q\max\{6\kappa,2\kappa+2,3\}}\max\left\{|\alpha|^{2q\max\{6\kappa,2\kappa+2,3\}}, |\beta|^{2q\max\{6\kappa,2\kappa+2,3\}}\right\}\right\}\right]^{1/q} \\ &\leq d^{\max\{6\kappa,2\kappa+2,3\}}\left[\max\{1,T\}\right]^{1/q}\max\left\{1, |\alpha|^{2\max\{6\kappa,2\kappa+2,3\}}, |\beta|^{2\max\{6\kappa,2\kappa+2,3\}}\right\} \leq \eta d^{\eta}. \end{aligned} \tag{4.170}$$

Corollary 4.2 hence ensures that there exist $\mathfrak{C} \in (0,\infty)$ and $(\Phi_{d,\varepsilon})_{d\in\mathbb{N},\varepsilon\in(0,1]} \subseteq \mathbf{N}$ which satisfy for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ that $\mathcal{R}_{\mathfrak{r}}(\Phi_{d,\varepsilon}) \in C(\mathbb{R}^{d+1},\mathbb{R})$, $\mathcal{P}(\Phi_{d,\varepsilon}) \leq \mathfrak{C}\varepsilon^{-\mathfrak{C}}d^{\mathfrak{C}}$, and

$$\left[\int_{[0,T]\times\mathbb{R}^d} |u_d(y) - (\mathcal{R}_{\mathfrak{r}}(\Phi_{d,\varepsilon}))(y)|^q\,\nu_d(dy)\right]^{1/q} \leq \varepsilon. \tag{4.171}$$

Combining this with (4.168) and Hölder's inequality proves that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ it holds that

$$\begin{aligned}&\left[\int_{[0,T]\times[\alpha,\beta]^d} \frac{\left|u_d(y) - (\mathcal{R}_{\mathfrak{r}}(\Phi_{d,\delta(\varepsilon)}))(y)\right|^p}{|\beta-\alpha|^d}\,dy\right]^{1/p} \\ &\leq \left[\int_{[0,T]\times[\alpha,\beta]^d} \frac{\left|u_d(y) - (\mathcal{R}_{\mathfrak{r}}(\Phi_{d,\delta(\varepsilon)}))(y)\right|^q}{|\beta-\alpha|^d}\,dy\right]^{1/q} T^{1/p-1/q} \\ &= \left[\int_{[0,T]\times\mathbb{R}^d} \left|u_d(y) - (\mathcal{R}_{\mathfrak{r}}(\Phi_{d,\delta(\varepsilon)}))(y)\right|^q\,\nu_d(dy)\right]^{1/q} T^{1/p-1/q} \\ &\leq \delta(\varepsilon)\left[\max\{T,1\}\right]^{1/p-1/q} = \varepsilon[\max\{T,1\}]^{1/q-1/p}[\max\{T,1\}]^{1/p-1/q} = \varepsilon. \end{aligned} \tag{4.172}$$

In addition, observe that for all $z \in (0,\infty)$ it holds that

$$z \leq z[\max\{T,1\}]^{z(1/p-1/q)} = r(z). \tag{4.173}$$

The fact that $\forall\, d \in \mathbb{N}$, $\varepsilon \in (0,1]\colon \mathcal{R}_{\mathfrak{r}}(\Phi_{d,\varepsilon}) \in C(\mathbb{R}^{d+1},\mathbb{R})$ and the fact that $\forall\, d \in \mathbb{N}$, $\varepsilon \in (0,1]\colon \mathcal{P}(\Phi_{d,\varepsilon}) \leq \mathfrak{C}\varepsilon^{-\mathfrak{C}}d^{\mathfrak{C}}$ hence show that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ it holds that $\mathcal{R}_{\mathfrak{r}}(\Phi_{d,\delta(\varepsilon)}) \in C(\mathbb{R}^{d+1},\mathbb{R})$ and

$$\begin{aligned}\mathcal{P}(\Phi_{d,\delta(\varepsilon)}) &\leq \mathfrak{C}[\delta(\varepsilon)]^{-\mathfrak{C}}d^{\mathfrak{C}} = \mathfrak{C}[\max\{T,1\}]^{-\mathfrak{C}(1/q-1/p)}\varepsilon^{-\mathfrak{C}}d^{\mathfrak{C}} \\ &= \mathfrak{C}[\max\{T,1\}]^{\mathfrak{C}(1/p-1/q)}\varepsilon^{-\mathfrak{C}}d^{\mathfrak{C}} = r(\mathfrak{C})\varepsilon^{-\mathfrak{C}}d^{\mathfrak{C}} \leq r(\mathfrak{C})\varepsilon^{-r(\mathfrak{C})}d^{r(\mathfrak{C})}. \end{aligned} \tag{4.174}$$

This and (4.172) establish that there exist $c \in \mathbb{R}$ and $(\mathfrak{u}_{d,\varepsilon})_{d\in\mathbb{N},\varepsilon\in(0,1]} \subseteq \mathbf{N}$ such that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ it holds that $\mathcal{R}_{\mathfrak{r}}(\mathfrak{u}_{d,\varepsilon}) \in C(\mathbb{R}^{d+1},\mathbb{R})$, $\mathcal{P}(\mathfrak{u}_{d,\varepsilon}) \leq c\varepsilon^{-c}d^c$, and

$$\left[\int_{[0,T]\times[\alpha,\beta]^d} \frac{|u_d(y) - (\mathcal{R}_{\mathfrak{r}}(\mathfrak{u}_{d,\varepsilon}))(y)|^p}{|\beta-\alpha|^d}\,dy\right]^{1/p} \leq \varepsilon. \tag{4.175}$$

This completes the proof of Corollary 4.3. □

Corollary 4.4. *Let $f_d^m\colon \mathbb{R}^d \to \mathbb{R}^{md-m+1}$, $m \in \{0,1\}$, $d \in \mathbb{N}$, be functions, let $T,\kappa,p \in (0,\infty)$, $a \in \mathbb{R}$, $b \in (a,\infty)$, $(\mathfrak{f}^m_{d,\varepsilon})_{(m,d,\varepsilon)\in\{0,1\}\times\mathbb{N}\times(0,1]} \subseteq \mathbf{N}$, assume for all $m \in \{0,1\}$, $d \in \mathbb{N}$, $\varepsilon \in (0,1]$, $x,y \in \mathbb{R}^d$ that*

$$\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0_{d,\varepsilon}) \in C(\mathbb{R}^d, \mathbb{R}), \quad \mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^1_{d,\varepsilon}) \in C(\mathbb{R}^d, \mathbb{R}^d), \quad \mathcal{P}(\mathfrak{f}^m_{d,\varepsilon}) \leq \kappa d^\kappa \varepsilon^{-\kappa}, \tag{4.176}$$

$$\left\|f^1_d(x) - f^1_d(y)\right\| \leq \kappa\|x-y\|, \quad \left\|(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^1_{d,\varepsilon}))(x)\right\| \leq \kappa(d^\kappa + \|x\|), \tag{4.177}$$

$$\left|(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0_{d,\varepsilon}))(x) - (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0_{d,\varepsilon}))(y)\right| \leq \kappa d^\kappa(1 + \|x\|^\kappa + \|y\|^\kappa)\|x-y\|, \tag{4.178}$$

$$\varepsilon\left|f^0_d(x)\right| + \left\|f^m_d(x) - (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^m_{d,\varepsilon}))(x)\right\| \leq \varepsilon\kappa d^\kappa(1 + \|x\|^\kappa), \tag{4.179}$$

and for every $d \in \mathbb{N}$ let $u_d \in \{v \in C([0,T] \times \mathbb{R}^d, \mathbb{R}) \colon \inf_{q \in (0,\infty)} \sup_{(t,y) \in [0,T] \times \mathbb{R}^d} \frac{|v(t,y)|}{1+\|y\|^q} < \infty\}$ be a viscosity solution of

$$\left(\frac{\partial}{\partial t} u_d\right)(t,x) = (\Delta_x u_d)(t,x) + \left(\frac{\partial}{\partial x} u_d\right)(t,x)\, f^1_d(x) \tag{4.180}$$

with $u_d(0,x) = f^0_d(x)$ for $(t,x) \in (0,T) \times \mathbb{R}^d$ (cf. Definitions 2.1, 4.1, 4.3, and 4.4). Then there exist $c \in \mathbb{R}$ and $(\mathfrak{u}_{d,\varepsilon})_{(d,\varepsilon) \in \mathbb{N} \times (0,1]} \subseteq \mathbf{N}$ such that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ it holds that $\mathcal{R}_{\mathfrak{r}}(\mathfrak{u}_{d,\varepsilon}) \in C(\mathbb{R}^{d+1}, \mathbb{R})$, $\mathcal{P}(\mathfrak{u}_{d,\varepsilon}) \leq c\varepsilon^{-c} d^c$, and

$$\left[\int_{[0,T]\times[a,b]^d} \frac{|u_d(y) - (\mathcal{R}_{\mathfrak{r}}(\mathfrak{u}_{d,\varepsilon}))(y)|^p}{|b-a|^d}\, dy\right]^{1/p} \leq \varepsilon. \tag{4.181}$$

Proof. Throughout this proof let $\iota = \max\{3\kappa, 2(\kappa+1)\}$. Observe that (4.177) and the fact that $\iota \geq \kappa$ prove that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$, $x, y \in \mathbb{R}^d$ it holds that

$$\left\|f^1_d(x) - f^1_d(y)\right\| \leq \iota\|x-y\| \quad \text{and} \quad \left\|(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^1_{d,\varepsilon}))(x)\right\| \leq \iota(d^\iota + \|x\|). \tag{4.182}$$

Next note that (4.178) and the fact that $\iota \geq 3\kappa$ ensure that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$, $x, y \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
&\left|(\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0_{d,\varepsilon}))(x) - (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^0_{d,\varepsilon}))(y)\right| \\
&\leq \kappa d^\iota(3 + \|x\|^\iota + \|y\|^\iota)\|x-y\| \\
&\leq \iota d^\iota(1 + \|x\|^\iota + \|y\|^\iota)\|x-y\|.
\end{aligned} \tag{4.183}$$

In addition, observe that (4.179) and the fact that $\iota \geq 2(\kappa+1)$ show that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$, $m \in \{0,1\}$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned}
&\varepsilon\left|f^0_d(x)\right| + \varepsilon + \left\|f^m_d(x) - (\mathcal{R}_{\mathfrak{r}}(\mathfrak{f}^m_{d,\varepsilon}))(x)\right\| \\
&\leq \varepsilon\kappa d^\kappa(1 + \|x\|^\kappa) + \varepsilon \leq \varepsilon(\kappa+1)d^\kappa(1 + \|x\|^\kappa) \\
&\leq \varepsilon(\kappa+1)d^\kappa(2 + \|x\|^\iota) \leq \varepsilon\iota d^\iota(1 + \|x\|^\iota).
\end{aligned} \tag{4.184}$$

Combining this, (4.182), (4.183), and Corollary 4.3 implies that there exist $c \in \mathbb{R}$ and $(\mathfrak{u}_{d,\varepsilon})_{(d,\varepsilon) \in \mathbb{N} \times (0,1]} \subseteq \mathbf{N}$ such that for all $d \in \mathbb{N}$, $\varepsilon \in (0,1]$ it holds that $\mathcal{R}_{\mathfrak{r}}(\mathfrak{u}_{d,\varepsilon}) \in C(\mathbb{R}^{d+1}, \mathbb{R})$, $\mathcal{P}(\mathfrak{u}_{d,\varepsilon}) \leq c\varepsilon^{-c} d^c$, and

$$\left[\int_{[0,T]\times[a,b]^d} \frac{|u_d(y) - (\mathcal{R}_{\mathfrak{r}}(\mathfrak{u}_{d,\varepsilon}))(y)|^p}{|b-a|^d}\, dy\right]^{1/p} \leq \varepsilon. \tag{4.185}$$

This completes the proof of Corollary 4.4. □

Acknowledgements. Philipp Grohs is gratefully acknowledged for several useful comments. This work has been partially funded by the Swiss National Science Foundation (SNSF) through the research grant 200020_175699, by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC 1173, by the Karlsruhe House of Young Scientists (KHYS) through a research travel grant, by ETH Foundations of Data Science (ETH - FDS), and by the European Union (ERC, MONTECARLO, 101045811). The views and the opinions expressed in this work are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. In addition, the second author gratefully acknowledges the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC 2044-390685587, Mathematics Münster: Dynamics-Geometry-Structure.

References

[1] F. Bach, Breaking the curse of dimensionality with convex neural networks, *J. Mach. Learn. Res.*, **18** (2017), 1–53.

[2] C. Beck, S. Becker, P. Grohs, N. Jaafari, and A. Jentzen, Solving stochastic differential equations and Kolmogorov equations by means of deep learning, *arXiv:1806.00421*, (2018), 56 pp.

[3] C. Beck, S. Becker, P. Grohs, N. Jaafari, and A. Jentzen, Solving the Kolmogorov PDE by means of deep learning, *J. Sci. Comput.*, **88** (2021), Paper No. 73, 28 pp.

[4] C. Beck, M. Hutzenthaler, and A. Jentzen, On nonlinear Feynman-Kac formulas for viscosity solutions of semilinear parabolic partial differential equations, *Stoch. Dyn.*, **21** (2021), Paper No. 2150048, 68 pp.

[5] C. Beck, M. Hutzenthaler, A. Jentzen, and B. Kuckuck, An overview on deep learning-based approximation methods for partial differential equations, *Discrete Contin. Dyn. Syst. Ser. B*, **28** (2023), 3697–3746.

[6] S. Becker, A. Jentzen, M. S. Müller, and P. von Wurstemberger, Learning the random variables in Monte Carlo simulations with stochastic gradient descent: Machine learning for parametric PDEs and financial derivative pricing, *Math. Finance*, **34** (2024), 90–150.

[7] R. Bellman, *Dynamic Programming*, Princeton University Press, 1957.

[8] P. Beneventano, P. Cheridito, R. Graeber, A. Jentzen, and B. Kuckuck, Deep neural network approximation theory for high-dimensional functions, *arXiv:2112.14523*, (2021), 82 pp.

[9] J. Berg and K. Nyström, A unified deep artificial neural network approach to partial differential equations in complex geometries, *Neurocomputing*, **317** (2018), 28–41.

[10] J. Berner, M. Dablander, P. Grohs, *Numerically Solving Parametric Families of High-Dimensional Kolmogorov Partial Differential Equations via Deep Learning*, in: *Proceedings of the 34th International Conference on Neural Information Processing Systems*, **1394** (2020), 16615–16627.

[11] Q. Chan-Wai-Nam, J. Mikael, and X. Warin, Machine learning for semi linear PDEs, *J. Sci. Comput.*, **79** (2019), 1667–1712.

[12] P. Cheridito, A. Jentzen, and F. Rossmannek, Efficient approximation of high-dimensional functions with neural networks, *IEEE Trans. Neural Netw. Learn. Syst.*, **33** (2022), 3079–3093.

[13] M. W. M. G. Dissanayake and N. Phan-Thien, Neural-network-based approximations for solving partial differential equations, *Commun. Numer. Meth. Engng.*, **10** (1994), 195–201.

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

[15] W. E, J. Han, and A. Jentzen, Algorithms for solving high dimensional PDEs: From nonlinear Monte Carlo to machine learning, *Nonlinearity*, **35** (2022), 278–310.

[16] W. E, M. Hutzenthaler, A. Jentzen, and T. Kruse, 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** (2019), 1534–1571.

[17] W. E, M. Hutzenthaler, A. Jentzen, and T. Kruse, Multilevel Picard iterations for solving smooth semilinear parabolic heat equations, *Partial Differ. Equ. Appl.*, **2** (2021), Paper No. 80, 31 pp.

[18] W. E and Q. Wang, Exponential convergence of the deep neural network approximation for analytic functions, *Sci. China Math.*, **61** (2018), 1733–1740.

[19] W. E and B. Yu, The deep Ritz method: A deep learning-based numerical algorithm for solving variational problems, *Commun. Math. Stat.*, **6** (2018), 1–12.

[20] D. Elbrächter, P. Grohs, A. Jentzen, and C. Schwab, DNN expression rate analysis of high-dimensional PDEs: Application to option pricing, *Constr. Approx.*, **55** (2022), 3–71.

[21] A.-M. Farahmand, S. Nabi, and D. Nikovski, Deep reinforcement learning for partial differential equation control, *2017 American Control Conference (ACC)*, (2017), 3120–3127.

[22] M. Fujii, A. Takahashi, and M. Takahashi, Asymptotic expansion as prior knowledge in deep learning method for high dimensional BSDEs, *Asia-Pac. Financ. Markets*, **26** (2019), 391–408.

[23] L. Gonon, P. Grohs, A. Jentzen, D. Kofler, and D. Šiška, Uniform error estimates for artificial neural network approximations for heat equations, *IMA J. Numer. Anal.*, **42** (2022), 1991–2054.

[24] L. Goudenège, A. Molent, and A. Zanette, Machine learning for pricing American options in high-dimensional Markovian and non-Markovian models, *Quant. Finance*, **20** (2020), 573–591.

[25] P. Grohs, F. Hornung, A. Jentzen, and P. von Wurstemberger, A proof that artificial neural networks overcome the curse of dimensionality in the numerical approximation of Black-Scholes partial differential equations, *Mem. Amer. Math. Soc.*, **284** (2023), v+93.

[26] P. Grohs, F. Hornung, A. Jentzen, and P. Zimmermann, Space-time error estimates for deep neural network approximations for differential equations, *Adv. Comput. Math.*, **49** (2023), 78 pp.

[27] P. Grohs, A. Jentzen, and D. Salimova, Deep neural network approximations for solutions of PDEs based on Monte Carlo algorithms, *Partial Differ. Equ. Appl.*, **3** (2022), Paper No. 45, 41 pp.

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

[29] J. Han, A. Jentzen, and W. E, Solving high-dimensional partial differential equations using deep learning, *Proc. Natl. Acad. Sci. USA*, **115** (2018), 8505–8510.

[30] J. Han and J. Long, Convergence of the deep BSDE method for coupled FBSDEs, *Probab. Uncertain. Quant. Risk*, **5** (2020), Paper No. 5, 33 pp.

[31] P. Henry-Labordère, Counterparty risk valuation: A marked branching diffusion approach, *arXiv:1203.2369*, (2012), 17 pp.

[32] P. Henry-Labordère, Deep Primal-Dual Algorithm for BSDEs: Applications of Machine Learning to CVA and IM, SSRN, (2017), 16 pp., doi:10.2139/ssrn.3071506.

[33] P. Henry-Labordère, N. Oudjane, X. Tan, N. Touzi, and X. Warin, Branching diffusion representation of semilinear PDEs and Monte Carlo approximation, *Ann. Inst. Henri Poincaré Probab. Stat.*, **55** (2019), 184–210.

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

[35] C. Huré, H. Pham, and X. Warin, Deep backward schemes for high-dimensional nonlinear PDEs, *Math. Comp.*, **89** (2020), 1547–1579.

[36] M. Hutzenthaler, A. Jentzen, T. Kruse, T. A. Nguyen, and P. von Wurstemberger, Overcoming the curse of dimensionality in the numerical approximation of semilinear parabolic partial differential equations, *Proc. A.*, **476** (2020), 20190630, 25 pp.

[37] M. Hutzenthaler, A. Jentzen, and P. von Wurstemberger, Overcoming the curse of dimensionality in the approximative pricing of financial derivatives with default risks, *Electron. J. Probab.*, **25** (2020), Paper No. 101, 73 pp.

[38] A. Jacquier and M. Oumgari, Deep curve-dependent PDEs for affine rough volatility, *SIAM J. Financial Math.*, **14** (2023), 353–382.

[39] A. Jentzen, D. Salimova, and T. Welti, 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, *Commun. Math. Sci.*, **19** (2021), 1167–1205.

[40] L. Jianyu, L. Siwei, Q. Yingjian, and H. Yaping, Numerical solution of elliptic partial differential equation using radial basis function neural networks, *Neural Networks*, **16** (2003), 729–734.

[41] Y. Khoo, J. Lu, and L. Ying, Solving parametric PDE problems with artificial neural networks, *European J. Appl. Math.*, **32** (2021), 421–435.

[42] P. Kidger and T. Lyons, *Universal approximation with deep narrow networks*, in: *Proceedings of Machine Learning Research*, Vol. 125, (2020), 2306–2327.

[43] A. Klenke, *Probability Theory: A Comprehensive Course*, Second Edition, Universitext, Springer, 2014.

[44] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Stochastic Modelling and Applied Probability, Vol. 23, Springer-Verlag, 1992.

[45] G. Kutyniok, P. Petersen, M. Raslan, and R. Schneider, A theoretical analysis of deep neural networks and parametric PDEs, *Constr. Approx.*, **55** (2022), 73–125.

[46] I. E. Lagaris, A. Likas, and D. I. Fotiadis, Artificial neural networks for solving ordinary and partial differential equations, *IEEE Trans. Neural Netw.*, **9(5)** (1998), 987–1000.

[47] H. Li and L. Ying, A semigroup method for high dimensional elliptic PDEs and eigenvalue problems based on neural networks, *J. Comput. Phys.*, **453** (2022), Paper No. 110939, 15 pp.

[48] W. Liu and M. Röckner, *Stochastic Partial Differential Equations: An Introduction*, Universitext, Springer, Cham, 2015.

[49] Z. Long, Y. Lu, X. Ma, and B. Dong, *PDE-Net: Learning PDEs from Data*, in: *Proceedings of the 35th International Conference on Machine Learning*, (2018), 3208–3216.

[50] K. O. Lye, S. Mishra, and D. Ray, Deep learning observables in computational fluid dynamics, *J. Comput. Phys.*, **410** (2020), 109339, 26 pp.

[51] M. Magill, F. Qureshi, and H. W. de Haan, *Neural networks trained to solve differential equations learn general representations*, in: *Advances in Neural Information Processing Systems*, (2018), 4071–4081.

[52] E. Novak and K. Ritter, *The curse of dimension and a universal method for numerical integration*, in: *Multivariate approximation and splines (Mannheim, 1996)*, International Series of Numerical Mathematics, Vol. 125, Birkhäuser, Basel, (1997), 177–187.

[53] E. Novak and H. Woźniakowski, *Tractability of Multivariate Problems. Volume I: Linear Information*, EMS Tracts in Mathematics, Vol. 6, EMS, 2008.

[54] E. Novak and H. Woźniakowski, *Tractability of Multivariate Problems. Volume II: Standard Information for Functionals*, EMS Tracts in Mathematics, Vol. 12, EMS, 2010.

[55] N. Nüsken and L. Richter, Solving high-dimensional Hamilton-Jacobi-Bellman PDEs using neural networks: Perspectives from the theory of controlled diffusions and measures on path space, *Part. Diff. Equ. Appl.*, **2** (2021), Paper No. 48.

[56] H. Pham, X. Warin, and M. Germain, Neural networks-based backward scheme for fully nonlinear PDEs, *Partial Differ. Equ. Appl.*, **2** (2021), Paper No. 16, 24 pp.

[57] M. Raissi, Deep hidden physics models: Deep learning of nonlinear partial differential equations, *J. Mach. Learn. Res.*, **19** (2018), 932–955.

[58] C. Reisinger and Y. Zhang, Rectified deep neural networks overcome the curse of dimensionality for nonsmooth value functions in zero-sum games of nonlinear stiff systems, *Anal. Appl. (Singap.)*, **18** (2020), 951–999.

[59] J. Sirignano and K. Spiliopoulos, DGM: A deep learning algorithm for solving partial differential equations, *J. Comput. Phys.*, **375** (2018), 1339–1364.

[60] M. Zhou, J. Han, and J. Lu, Actor-critic method for high dimensional static Hamilton-Jacobi-Bellman partial differential equations based on neural networks, *SIAM J. Sci. Comput.*, **43** (2021), A4043–A4066.