Journal of Computational Mathematics
Vol.43, No.4, 2025, 918–975.
http://www.global-sci.org/jcm
doi:10.4208/cm.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 $d \in \mathbb{N} = \{1, 2, \dots\}$ 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* $\mathbf{N} = \cup_{L\in\mathbb{N}}\cup_{l_0,l_1,\dots,l_L\in\mathbb{N}}(\times_{k=1}^{L}(\mathbb{R}^{l_k\times l_{k-1}}\times\mathbb{R}^{l_k}))$, *let* $A\colon (\cup_{d\in\mathbb{N}}\mathbb{R}^d) \to (\cup_{d\in\mathbb{N}}\mathbb{R}^d)$ *satisfy for all* $d \in \mathbb{N}$, $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ *that*

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

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* $L \in \mathbb{N}$, $l_0, l_1, \dots, l_L \in \mathbb{N}$, $\Phi = ((W_1, B_1), \dots, (W_L, B_L)) \in (\times_{k=1}^{L}(\mathbb{R}^{l_k\times l_{k-1}} \times \mathbb{R}^{l_k}))$, $x_0 \in \mathbb{R}^{l_0}, x_1 \in \mathbb{R}^{l_1}, \dots, x_L \in \mathbb{R}^{l_L}$ *with*