The 8th International Conference on Computational Harmonic Analysis 2022 (ICCHA2022)

Effective deep neural network architectures for learning high-dimensional Banach-valued functions from limited data

Nick Dexter
Simon Fraser University
nicholas_dexter@sfu.ca

Sebastian Moraga
Simon Fraser University
sebastian_moraga_scheuermann@sfu.ca

Simone Brugiapaglia
Concordia University
simone.brugiapaglia@concordia.ca

Ben Adcock
Simon Fraser University
ben_adcock@sfu.ca

***Abstract*—In the past few decades the problem of reconstructing high-dimensional functions taking values in abstract spaces from limited samples has received increasing attention, largely due to its relevance to uncertainty quantification (UQ) for computational science and engineering. These UQ problems are often posed in terms of parameterized partial differential equations (PDE) whose solutions take values in Hilbert or Banach spaces. Impressive results have been achieved on such problems with deep learning (DL), i.e. machine learning with deep neural networks (DNN). This work focuses on approximating high-dimensional smooth functions taking values in reflexive and typically infinite-dimensional Banach spaces. Our novel approach to this problem is fully algorithmic, combining DL, compressed sensing (CS), orthogonal polynomials, and finite element discretization. We present a full theoretical analysis for DNN approximation with explicit guarantees on the error and sample complexity, and a clear accounting of all sources of error. We also provide numerical experiments showing that DNNs can produce accurate approximations on challenging Banach-valued benchmark problems.**

I. Introduction

A large number of recent works have proposed the use of DNNs in scientific computing and UQ over the past several years [1], [6], [8], [12], [13]. Many of these works demonstrate empirically that DNNs mitigate to a large degree the central challenge of high dimensionality in these problems, i.e., the *curse of dimensionality*. Parameterized PDE models for UQ can often be expressed in terms of differential operators in the spatial or temporal variables. The sampling-based approach for such problems considers the recovery of a function

$$f : \mathcal{U} \to \mathcal{V}, \quad \boldsymbol{y} \mapsto f(\boldsymbol{y}) \tag{I.1}$$

with $\mathcal{U} \subseteq \mathbb{R}^d$ with $d \in \mathbb{N}$ or $d = \infty$, and $\mathcal{V}$ the solution space of the PDE (a Hilbert or Banach space), from noisy samples $\{(\boldsymbol{y}_i, d_i)\}_{i=1}^m$ with $d_i = f(\boldsymbol{y}_i) + n_i \in \mathcal{V}_h$, a discretization of the infinite-dimensional space $\mathcal{V}$. Here the discretization process may introduce modelling errors, numerical error, or random noise to the measurements, which we account for in the convergence analysis.

II. Main Results

Our main result, stated below, extends results from [7], [10], [11] and quantifies all sources of error involved in the approximation of Banach-valued parameterized PDE solutions with DNNs.

Theorem II.1. *Let $f : \mathcal{U} \to \mathcal{V}$ be holomorphic w.r.t. $\boldsymbol{y}$ in a Bernstein polyellipse containing $\mathcal{U} \subseteq \mathbb{R}^d$, $d = \infty$. Then there exists (i) a class of ReLU or $\tanh$ DNNs, (ii) a loss function (regularized ℓ^2-loss), and (iii) a choice of regularization parameter λ depending on m only, for which every approximate minimizer f_Φ achieves an error scaling like $E_{\mathsf{app}} + m^{1/4}(E_{\mathsf{samp}} + E_{\mathsf{disc}} + E_{\mathsf{opt}})$, where E_{app}, E_{samp}, E_{disc}, and E_{opt} are the approximation, sampling, discretization, and objective function errors, respectively. Moreover the DNN f_Φ has both subexponential width and polylogarithmic depth in m.*

This result asserts convergence of DNNs with approximation error $E_{\mathsf{approx}} = C(m/c_0 L)^{\frac{1}{2}(\frac{1}{2}-\frac{1}{p})}$ and subexponential size in the number of training samples m, even in the infinite-dimensional setting. Comparing with results for Hilbert-valued approximation, we note an increase in sample complexity over that setting as a consequence of working in Banach spaces. Finally, we remark that the sample complexity and size of the network can be further reduced in the setting where *a priori* information of the anisotropy of f is given.

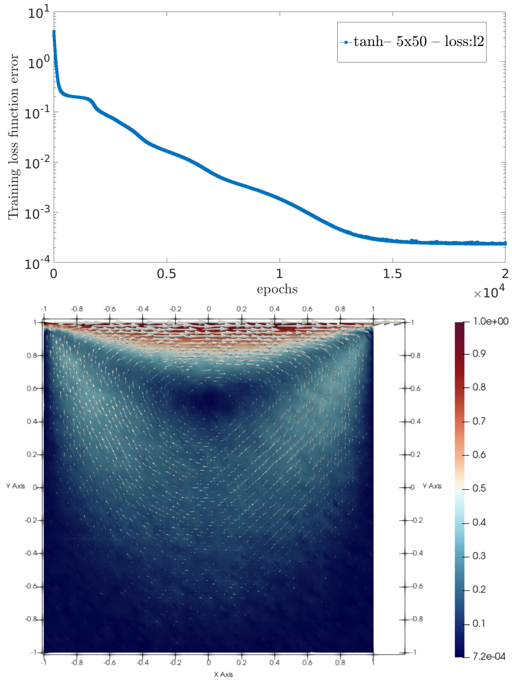

Fig. 1: For a $\tanh$ 5×50 DNN we show **(top)** MSE training error and **(bottom)** prediction of $u_{\Phi,h}(\boldsymbol{x}, \boldsymbol{y})$ at fixed $\boldsymbol{y}$ after 20,000 epochs of Adam [9].

III. Numerical Experiments

Our numerical results use the testing framework from [2] and focus on the case of recovering solutions to the parametric steady-state Navier-Stokes equations (NSE) with different input data on a challenging lid-driven cavity problem in two spatial dimensions. For more details see [5]. Figure 1 displays the training error in the mean-squared error (MSE) loss between the finite element coefficients computed by FEniCS [3] and those predicted by a DNN architecture inspired by [4]. This DNN is composed of multiple $\tanh$ 5×50 fully-connected subnetworks, each directly approximating one of the variables $(u, \boldsymbol{\tau}, p)$ of interest, corresponding to the solution, derivative, and pressure terms of the NSE. We also display the predicted solution $u_{\Phi,h}(\boldsymbol{x}, \boldsymbol{y})$ at a fixed $\boldsymbol{y}$ value, showing the DNN successfully captures the dynamics of this challenging benchmark problem.

IV. Conclusion

DNNs offer great potential in accurately resolving high-dimensional functions of interest in general abstract spaces, and their use in computational science and engineering is only increasing. Therefore it is paramount that we understand their practical benefits and limitations in various modeling contexts. In particular, it is critical that we understand key issues such as their approximation capabilities, sample complexity, and stability and robustness properties.

References

[1] B. Adcock, S. Brugiapaglia, N. Dexter, S. Moraga, *Deep neural networks are effective at learning high-dimensional Hilbert-valued functions from limited data.* In J. Bruna, J. S. Hesthaven, and L. Zdeborová, editors, *Proceedings of The 2nd Annual Conference on Mathematical and Scientific Machine Learning*, volume 145 of Proc. Mach. Learn. Res. (PMLR), pages 1–36. PMLR, 2021.

[2] B. Adcock and N. Dexter, *The gap between theory and practice in function approximation with deep neural networks*, SIAM Journal on Mathematics of Data Science, 3(2), pp. 624–655, 2021.

[3] M.S. Alnaes, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M.E. Rognes and G.N. Wells. *The FEniCS Project Version 1.5*, Archive of Numerical Software 3 (2015).

[4] S. Cai, Z. Wang, L. Lu, T.A. Zaki, G.E. Karniadakis. *DeepM&Mnet: Inferring the electroconvection multiphysics fields based on operator approximation by neural networks.* Journal of Computational Physics, (2021).

[5] E. Colmenares, G.N. Gatica, S. Moraga, *A Banach spaces-based analysis of a new fully-mixed finite element method for the Boussinesq problem*. ESAIM, Math. Model. Numer. Anal., (2020).

[6] N. Dal Santo, S. Deparis, and L. Pegolotti, *Data driven approximation of parametrized PDEs by reduced basis and neural networks*, Journal of Computational Physics, 416, 109550, 2020.

[7] J. Daws, and C. Webster, *Analysis of Deep Neural Networks with Quasi-optimal polynomial approximation rates*, arXiv preprint arXiv:1912.02302, 2019.

[8] M. Geist, P. Petersen, M. Raslan, R. Schneider, and G. Kutyniok, *Numerical Solution of the Parametric Diffusion Equation by Deep Neural Networks*, J. Sci. Comput. 88(22), 2021.

[9] D.P. Kingma, and J. Ba, *Adam: A method for stochastic optimization.* arXiv:1412.6980, 2014.

[10] J.A.A. Opschoor, C. Schwab, and J. Zech, *Exponential ReLU DNN expression of holomorphic maps in high dimension*, Constr. Approx., 2021.

[11] T. Ryck, S. Lanthaler, and S. Mishra, *On the approximation of functions by tanh neural networks.* Seminar for Applied Mathematics, ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland.

[12] L. Scarabosio, *Deep neural network surrogates for nonsmooth quantities of interest in shape uncertainty quantification*, arXiv:2101.07023, 2021.

[13] S. Wang, H. Wang, and P. Perdikaris, *Learning the solution operator of parametric partial differential equations with physics-informed DeepOnets*, arXiv:2103.10974, 2021.