Numerische Mathematik
https://doi.org/10.1007/s00211-022-01294-z

Numerische Mathematik

Error estimates for deep learning methods in fluid dynamics

A. Biswas[1] · J. Tian[2] · S. Ulusoy[3]

Received: 8 August 2020 / Revised: 20 September 2021 / Accepted: 18 March 2022

Abstract
In this study, we provide error estimates and stability analysis of deep learning techniques for certain partial differential equations including the incompressible Navier–Stokes equations. In particular, we obtain explicit error estimates (in suitable norms) for the solution computed by optimizing a loss function in a Deep Neural Network approximation of the solution, with a fixed complexity.

Mathematics Subject Classification 35Q35 · 35Q30 · 65M70

1 Introduction

Machine Learning, which has been at the forefront of the data science and artificial intelligence revolution in the last twenty years, has a wide range of applications in natural language processing, computer vision, speech and image recognition, among others [11, 13, 19]. Recently, its use has proliferated in computational sciences and physical modeling such as the modeling of turbulence [7, 18, 33–36]. Moreover, machine learning methods (*physics informed neural networks* [21, 22, 25, 27, 38] which are mesh-free) have gained wide applicability in obtaining numerical solutions of various types of partial differential equations (PDEs); see [2, 3, 12, 21, 23, 26–28,

✉ J. Tian
jtian@towson.edu

A. Biswas
abiswas@umbc.edu

S. Ulusoy
suleyman.ulusoy@aurak.ac.ae

1 Department of Mathematics and Statistics, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, USA

2 Department of Mathematics, Towson University, 7800 York Road, Towson, MD 21252, USA

3 Department of Mathematics and Natural Sciences, American University of Ras Al Khaimah, P.O. Box 10021, Ras Al Khaimah, UAE

Published online: 31 May 2022

30, 38] and the references therein. The need for these studies stems from the fact that when using traditional numerical methods in a high-dimensional PDE, the methods sometimes become infeasible. High-dimensional PDEs appear in a number of models for instance in the financial industry, in a variety of contexts such as in derivative pricing models, credit valuation adjustment models, or portfolio optimization models. Such high-dimensional fully nonlinear PDEs are exceedingly difficult to solve as the computational effort for standard approximation methods grows exponentially with the dimension. For example, in finite difference methods, as the dimension of the PDEs increases, the number of grids increases considerably and there is a need for reduced time step-size. This increases the computational cost and memory demands. Under these circumstances, implementing the deep learning algorithms can be helpful. In particular, the neural networks approach in partial differential equations (PDEs) offer implicit regularization and can overcome the curse of high dimensions [2, 3]. Additionally, this approach provides a natural framework for estimating unknown parameters [7, 27, 28, 33, 35].

Here our main focus is on numerical analysis of the neural networks techniques for solving the Navier–Stokes equations, after illustrating the fundamental issues involved in the elliptic case. The Navier–Stokes equations, either alone or coupled with governing equations of other physical quantities such as the temperature and/or the magnetic field, are the fundamental equations governing the motion of fluids. They appear in the study of diverse physical phenomena such as aerodynamics, geophysics, atmospheric physics, meteorology and plasma physics. For example, they are used in modeling the water flow in a pipe, air flow around a wing, ocean currents and weather. They are employed in the design of cars, aircrafts, and power stations, in the study of blood flow and many other applications.

There is by now an abundant literature proposing numerical schemes (and demonstrating their efficacy) employing DNN and machine learning tools for partial differential equations, including those arising in fluid dynamics; see for instance [7, 12, 21–23, 25–28, 30, 33, 38] and the references therein. However, concrete and complete mathematical analysis are relatively meager for such methods applied to PDEs, in particular, for the Navier–Stokes equations, although some results on convergence (as the complexity of the neural network tends to infinity) in the weak topology for some semilinear PDEs can be found in [30]. The goal of this paper is to provide a rigorous error analysis of deep learning methods employed in [12, 26–28], similar in spirit to the probabilistic error analysis for machine learning algorithms for the Black-Scholes equations in [3].

The computational algorithm employed in machine learning of PDEs (for instance in [12, 26–28]) involves representing the approximate solution by a *Deep Neural Network (DNN)*, in lieu of a spectral or finite element approximation, and then minimizing, over all such representations, an appropriate loss function, measuring the deviation of this representation from the PDE and the initial and boundary conditions. One important thing to note in this approach is the following. It is well-known that optimization of loss functions in a deep neural network is a non-convex optimization problem. Therefore, neither the existence nor the uniqueness of a global optimum is guaranteed. Typically, repeated application of stochastic gradient descent results in reaching a local minimum, which may or may not be global. Nevertheless, we side step this issue by taking advantage of the fact that in these applications, the minimum value of the loss function at the exact solution must be zero, and obtaining an explicit error estimate in terms of the attained value of the loss function. The estimate we obtain in turn guarantees that the approximate solution thus constructed converges, *in the strong topology,* to the true solution as the complexity of the networks tends to infinity.

Here, we briefly describe the results of this paper. The first part of the results are devoted to a second order elliptic equation posed on a bounded domain in $\mathbb{R}^2$. In the elliptic case, we choose a loss function that includes two parts; one part measures the fidelity to the equation, while the other measures the discrepancy with the boundary condition. Using results from approximation theory, we first show in Theorem 1 that we can find an approximate solution in a DNN and when the approximate solution and actual solution are close to each other, we can control the loss function (in terms of the complexity of the DNN). Theorems 2 and 3 provide in some sense the converse. More precisely, by establishing explicit error estimates, we show that by controlling the loss function, we can have a good approximation to the solution using a DNN which in turn justifies the DNN based numerical scheme. The proof employs interpolation inequalities and the lifting operator. Moreover, in Theorem 3, we show that by using a more stringent loss function, the error estimate can be improved. The improved error rate is optimal for the given regularity of the right hand side. To summarize, our results have shown the optimal error rate for deep learning method in elliptic equation.

Due to the (nonlinear and nonlocal) nature of the equation, the Navier–Stokes equations is much harder to study compared with the elliptic equation. The (well-known) difficulties include a lack of an evolution equation for the pressure which essentially acts like a Lagrangian multiplier that enforces the divergence free condition on the velocity [32]. Moreover, the fact that the functions represented by the DNN do not satisfy the boundary condition poses an additional challenge to the theoretical analysis. This is expected since traditional analysis of the inhomogeneous boundary condition case for the Navier–Stokes equations (being a nonlinear equation) is involved and employs corrector and/or extension functions [8, 10, 29].

In this study, unlike the elliptic case, the loss function for the Navier–Stokes equations consists of five terms, corresponding to the boundary condition, the initial condition, the divergence-free condition, the equation itself, and one is a penalty term corresponding to the regularity gain. With the choice of this loss function, we show that when applying the deep learning algorithm on the Navier–Stokes equations, with a small loss function, the approximate solution and actual solution are close to each other. The proof of this result is nontrivial. By the Hodge decomposition, the approximate solution is decomposed into two parts, namely u_N and v_N, necessitated by the fact that functions represented by the DNN are not necessarily zero on the boundary. Using the properties of the bilinear map and Gronwall's inequality, we first estimate the difference between u_N and actual solution in the L^2 norm in a certain time interval. Then, by again decomposing v_N into two parts, we obtain a control of v_N. Applying the Leray projection on the Navier–Stokes equation and using the estimates on u_N and v_N, we obtain the estimate on the difference between the approximate solution and accurate solution. On the other hand, we prove that by using the deep learning, we can find an approximate solution such that the loss function is small. Here, we study the loss function term by term. Lastly, we show that our scheme is approximately stable. Due to differing regularity properties of the terms involved, choosing appropriate norms for various constituent parts of the loss function is somewhat delicate for the Navier–Stokes case.

Theorems 4–6 are, in some sense, the analogs of Theorems 1–3 for the $2D$ Navier–Stokes equations. More precisely, Theorem 4 states the precise error estimate between a strong solution of the $2D$ Navier–Stokes equations and the approximate solution. The reverse direction of Theorem 4 is shown in Theorem 5. Finally, Theorem 6 shows that our scheme is approximately stable. We note moreover that the constants occurring in the error estimates are explicit: they are either physical in nature (i.e. depend on the Grashoff number (L^2-norm) of the driving force or are domain dependent constants (i.e., they are either constants appearing in Sobolev inequalities or are operator norms of lifting, extension or trace operators)). Finally, we note that although our results are proven in the context of the two-dimensional Navier–Stokes equations, our analysis applies equally well to the three dimensional case, up to the interval of existence of a *strong solution*, which in the two dimensional case, exists globally in time.

The rest of the paper is organized as follows. Section 2 provides the preliminaries for both Neural Network settings and approximation properties which will be used in this study. Section 3 is devoted to the statement of our main results. In Sect. 4, we present the mathematical analysis of the neural network algorithm in the elliptic system. This also serves as a template for our analysis of the Navier–Stokes equations. In Sect. 5, we present our main results in two dimensional Navier–Stokes equations. By using Hodge decomposition, we have shown that the approximate solution using the neural network algorithm is close to the actual solution of the two dimensional Navier–Stokes equations under certain conditions. Moreover, we have proved that our scheme is approximately stable. The existence of the approximate solution is shown by applying approximation properties of neural networks.

2 Preliminaries

2.1 Neural networks

In a DNN, we consider a mapping $f : x \mapsto y$, where x is the input variable and y is the output variable. The mapping function f is obtained by (function) composition of *layer functions*, comprising of an input layer, an output layer and multiple hidden layers, connected in *neural network*. The details are as follows.

In a DNN, each layer is a function of the form $\sigma(wx + b)$, $x \in \mathbb{R}^d, w = (w_1, \ldots, w_d)$, $b \in \mathbb{R}$. Here, σ is called the *activation function* and is usually taken to be either a sigmoid ($\sigma(x) = \frac{e^x}{e^x+1}$) , tanh or $\Re\ln$, where $\Re\ln(\xi) := \max(0, \xi)$. In applications to PDE, where we require adequate regularity of solutions, a popular choice is the tanh function where $\tanh(\xi) = \dfrac{e^{\xi} - e^{-\xi}}{e^{\xi} + e^{-\xi}}$.

Consider the collection of functions of the form

$$\sum \alpha_j f_1 \circ f_2 \circ f_3 \circ \cdots \circ f_{l_j}(x), \tag{2.1}$$

where f_i is a function of the form $\sigma(wx+b)$ described above. In (2.1), $\max l_j$ is called the depth of the network. Henceforth, we will denote by $\mathcal{F}_N$ the class of functions in (2.1), where N represents the network complexity (e.g. N could be the sum of the ranks of the weight matrices w and the number of layers in the DNN).

For the sake of completeness, we give a schematic representation of a neural network. Here, we adapt the standard dense neural networks which can be expressed as a series of compositions:

$$\begin{aligned} y_2(x) &= \sigma(W_1 x + b_1), \\ y_3(y_2) &= \sigma(W_2 y_2 + b_2), \\ &\cdot \\ &\cdot \\ &\cdot \\ y_{n_l}(y_{n_l-1}) &= \sigma(W_{n_l-1} y_{n_l-1} + b_{n_l-1}), \\ y_{n_l+1}(y_{n_l}) &= \sigma(W_{n_l} y_{n_l} + b_{n_l}), \\ f_\theta &= y_{n_l+1}(y_{n_l}(\cdots(y_2(x)))), \end{aligned}$$

where θ ensembles all the weights and parameters.

$$\theta = \left\{W_1, W_2, \ldots, W_{n_l}, b_1, \ldots, b_{n_l}\right\}. \tag{2.2}$$

In practice, different neural network architectures are possible such as those involving recurrent cells [16], convolutional layers [19], sparse convolutional neural networks [17], pooling layers, residual connections [13].

In this study, we assume that our neural networks are equipped with uniformly bounded weights and the final bias term b_{n_l}. We do not need any boundedness assumption on the other bias terms b_i.

2.2 Function approximation

Approximation properties of different DNNs has been studied extensively since the work of Cybenko [4] and Hornik [15]; see [24, 37] and the references therein for more recent work. An important question in the approximation process is how many neural network layers are needed to guarantee the approximation accuracy? In [1], the author showed that by using the sigmoidal activation funciton, at most $O(\varepsilon^{-2})$ neurons are needed to achieve the order of approximation ε. In [4], Cybenko proved that continuous functions can be approximated with arbitrary precision by the DNNs with one internal layer and an arbitrary continuous sigmoidal function providing that no constraints are placed on the number of nodes or the size of the weights. Also, in [14], Hornik *et. al.* provided the conditions ensuring that DNNs with a single hidden layer and an appropriately smooth hidden layer activation function are capable of arbitrarily accurate approximation to an arbitrary function and its derivatives.

3 Main results

Let $\mathcal{F}_N$ be a DNN with complexity N, which is a finite dimensional function space on a bounded domain. Throughout the paper, we use c to denote the absolute constants and the domain dependent constants occurring in the Poincaré and Sobolev inequalities. Below is a list of our main results.

3.1 Elliptic case

Consider a bounded domain Ω of $\mathbb{R}^2$ and the following partial differential equation

$$\begin{cases} \mathcal{L}u = f, \\ u|_{\partial\Omega} = g, \end{cases} \tag{3.1}$$

where $\mathcal{L} : H^2(\Omega) \to L^2(\Omega)$ is a second order uniformly bounded elliptic operator. In this study, for simplicity, we consider only the case $g = 0$, although the general case is similar.

Recall that (3.1) is well-posed and a unique solution exists satisfying

$$M := \|u\|_{H^2(\Omega)} \leq c\left(\|f\|_{L^2(\Omega)} + \|g\|_{H^{\frac{3}{2}}(\partial\Omega)}\right). \tag{3.2}$$

Consequently, the minimization problem

$$\inf_{u \in \text{ appropriate Sobolev class}} \left\{\|(\mathcal{L}u)(x) - f(x)\|^2_{L^2(\Omega)} + \|u|_{\partial\Omega}\|^2_{L^2(\partial\Omega)}\right\} \tag{3.3}$$

has a unique solution, with the value of the infimum being 0, and the infimum is attained at the solution u of (3.1). More generally, the same conclusion holds if we consider a loss function of the type

$$L = \alpha^2\|\mathcal{L}u - f\|^2_{L^2(\Omega)} + \beta^2\|u|_{\partial\Omega}\|^2_{L^2(\partial\Omega)}.$$

Thus, in order to approximate u using a DNN, one considers the loss function

$$L = \alpha^2\|\mathcal{L}u_N - f\|^2_{L^2(\Omega)} + \beta^2\|u_N|_{\partial\Omega}\|^2_{L^2(\partial\Omega)}, \; u_N \in \mathcal{F}_N, \tag{3.4}$$

under the restriction that $\|u_N\|_{H^2(\Omega)} \leq \widetilde{M}$ (i.e. $\|u_N\|_{H^2(\Omega)}$ is bounded) for suitable $\widetilde{M}$ (e.g. $\widetilde{M} = 2M$ where M is as in (3.2)), with $\alpha, \beta > 0$. Since the chosen activation function $\sigma = \tanh$ is smooth, in practice, this is achieved by restricting the (finite dimensional) parameter set in the neural network to a compact subset.

In the neural network framework, the optimization is usually conducted in a discrete setting as follows [27]. More precisely, let $\mathcal{F}_N$ be a finite dimensional function space on a bounded domain Ω. Choose a collocation points $\{x_j\}_{j=1}^m \subset \Omega$ and $\{y_j\}_{j=1}^n \subset \partial\Omega$.

Find

$$\inf_{u\in\mathcal{F}_N,\|u_N\|_{H^2(\Omega)}\leq\widetilde{M}}\left\{\alpha^2\sum_{j=1}^{m}\left|(\mathcal{L}u)(x_j)-f(x_j)\right|^2+\beta^2\sum_{j=1}^{n}\left|u(y_j)-g(y_j)\right|^2\right\}. \tag{3.5}$$

Note that (3.5) may be regarded as a Monte Carlo approximation of the corresponding Lebesgue integrals. Consequently, for mathematical convenience, let us consider the following optimization problem, namely, find

$$\inf_{u\in\mathcal{F}_N,\|u_N\|_{H^2(\Omega)}\leq\widetilde{M}}\left\{\|(\mathcal{L}u)(x)-f(x)\|^2_{L^2(\Omega)}+\|u|_{\partial\Omega}-g\|^2_{L^2(\partial\Omega)}\right\}. \tag{3.6}$$

The infimum can be attained provided that we restrict the parameters in $\mathcal{F}_N$ in a compact set. However in this case, the infimum may not be unique.

Remark 1 We can also use an unrestricted optimization in (3.5) or (3.6). However, in this case, the condition on $\|u_N\|_{H^2(\Omega)}\leq\widetilde{M}$ can be replaced by suitably adding a penalty/regularization term in the loss function. This converts the restricted minimization problem to an unrestricted one and is illustrated in the Navier–Stokes case. This drawback is due to the fact in contrast to spectral or finite element methods, the boundary conditions are not encoded in a DNN, but rather are enforced “approximately".

In all the boundary integrals above, the quantity $u|_{\partial\Omega}$ is interpreted as trace in case $u\in H^1(\Omega)$. However, since u_N is smooth, its trace coincides with its restriction on the boundary. Recall that the trace operator is defined as a bounded operator $\gamma\in\mathcal{L}(H^1(\Omega),L^2(\Gamma))$ such that γu is the restriction of u to Γ for every function $u\in H^1(\Omega)$ which is twice continuously differentiable in $\overline{\Omega}$.

First, we show that when the approximate solution and the actual solution are close to each other, we can control the loss function. This (as also the analogous theorem for the Navier–Stokes equations) is based on the following numerical analysis result from [37], relating the accuracy of the approximation by a DNN (i.e. the class $\mathcal{F}_N$) with its complexity (quantified by N) and the regularity of the function being approximated.

Suppose that $\sigma\in C^\infty(\mathbb{R})$, $\sigma^{(v)}(0)\neq 0$ for $\nu=0,1,\ldots,$ and $K\subset\mathbb{R}^d$ is any compact set. If $f\in C^k(K)$, then a function ϕ_N represented by a DNN $\mathcal{F}_N$, with complexity $N\in\mathbb{N}$ exists such that

$$\|D^\alpha f-D^\alpha\phi_N\|_{C(K)}=O\left(\frac{1}{N^{(k-|\alpha|)/d}}\omega\left(D^\beta f,\frac{1}{N^{1/d}}\right)\right) \tag{3.7}$$

holds for all multi-indexes α,β with $|\alpha|\leq k$, $|\beta|=k$, where

$$\omega(g,\delta)=\sup_{x,y\in K,|x-y|\leq\delta}|g(x)-g(y)|.$$

Theorem 1 *Let u be the solution of (3.1) and $\alpha, \beta, \epsilon > 0$. Then there exists $u_N \in \mathcal{F}_N$ with $\|u - u_N\|_{H^2(\Omega)} \leq \varepsilon$ such that*

$$\alpha^2 \|\mathfrak{L} u_N - f\|^2_{L^2(\Omega)} + \beta^2 \|u_N|_{\partial\Omega}\|^2_{L^2(\partial\Omega)} \leq (C_{\mathfrak{L}} \alpha^2 + C_{Tr} \beta^2)\varepsilon^2$$

and

$$\|u_N\|_{H^2(\Omega)} \leq \widetilde{M},$$

where $\widetilde{M}$ can be taken to be $2M = 2\|u\|_{H^2(\Omega)}$, with M as in (3.2). Here $C_{\mathfrak{L}}$ is the operator norm bound of $\mathfrak{L}$, and C_{Tr} is the constant from the trace operator.

On the other hand, we can show that by controlling the loss function, we can have a good approximation to the solution u of (3.1) by using a DNN. The requisite error estimate is given in the theorem below, where the notation used is as in Theorem 1.

Theorem 2 *Let u be a solution of (3.1) and $\alpha, \beta, \epsilon > 0$. Assume that $u_N \in \mathcal{F}_N$ is such that*

$$\alpha^2 \|\mathfrak{L} u_N - f\|^2_{L^2(\Omega)} + \beta^2 \|u_N|_{\partial\Omega}\|^2_{L^2(\partial\Omega)} \leq \varepsilon^2, \tag{3.8}$$

with $\|u_N\|_{H^2(\Omega)} \leq \widetilde{M}$. Then

$$\|u - u_N\|_{H^1(\Omega)} \leq \left(\frac{C_{\mathfrak{L}^{-1}}}{\alpha} + \frac{C_{\mathfrak{L}^{-1}} C_{\mathfrak{L}} \|l_\Omega\| M^{1/3}}{\beta^{2/3}} \right) \varepsilon^{2/3},$$

where, l_Ω is the lifting operator.

We show in the theorem below how the error estimate in the H^1-norm can be improved further by altering the loss function.

Theorem 3 *Let u be a solution of (3.1) and let $u_N \in \mathcal{F}_N$ be such that*

$$\alpha^2 \|\mathfrak{L} u_N - f\|^2_{L^2(\Omega)} + \beta^2 \|u_N|_{\partial\Omega}\|^2_{H^{1/2}(\partial\Omega)} \leq \varepsilon^2, \tag{3.9}$$

with $\|u_N\|_{H^2(\Omega)} \leq \widetilde{M}$. Then

$$\|u - u_N\|_{H^1(\Omega)} < \left(\frac{C_{\mathfrak{L}^{-1}}}{\alpha} + \frac{C_{\mathfrak{L}^{-1}} C_{\mathfrak{L}} \|l_\Omega\|}{\beta} \right) \varepsilon,$$

where $C_{\mathfrak{L}^{-1}}$, $C_{\mathfrak{L}}$ and l_Ω are the operator norms of the respective operators.

3.2 Incompressible Navier–Stokes equations

The incompressible Navier–Stokes equations (NSE) are given by

$$\begin{aligned} \partial_t u - \Delta u + u \cdot \nabla u + \nabla p &= f, \\ \nabla \cdot u &= 0, \\ u|_{\partial \Omega} &= 0, \\ u(x, 0) &= u_0(x), x \in \Omega. \end{aligned} \tag{3.10}$$

In (3.10), u denotes the velocity of the fluid and p the pressure. The nondimensional Grashof number G is defined as $G = \frac{1}{\nu^2 \lambda_1} \sup_{t \to \infty} \|f(t)\|_{L^2}$ [9], where λ_1 is the smallest eigenvalue of the Stokes operator. In our case (3.10), we take the fluid viscosity ν to be 1.

Similar to the elliptic case, we show that when applying the $\mathcal{F}_N$ on the Navier–Stokes equations, with a small loss function, the approximate solution and actual solution are close to each other.

Theorem 4 *Assume that u is a strong solution of the 2D NSE (3.10) and $\tilde{u}_N \in \mathcal{F}_N$ such that*

$$\begin{aligned} &\|\tilde{u}_N|_{\partial\Omega}\|^4_{L^4([0,T];H^{1/2}(\partial\Omega))} + \|\tilde{u}_N(x,0) - u_0(x)\|^2_{L^2(\Omega)} \\ &\quad + \|\partial_t \tilde{u}_N - \Delta \tilde{u}_N + \tilde{u}_N \cdot \nabla \tilde{u}_N + \nabla \tilde{p}_N - f\|^2_{L^2(\Omega \times [0,T])} \\ &\quad + \|\nabla \cdot \tilde{u}_N\|^4_{L^4([0,T];L^2(\Omega))} + \lambda \|\tilde{u}_N\|^4_{L^4([0,T];H^1(\Omega))} \leq \varepsilon^2. \end{aligned} \tag{3.11}$$

Then

$$\|u - \tilde{u}_N\|^4_{L^4([0,T];L^2(\Omega))} \leq (ce^{F(G,u_0)T} + C_{Tr})\varepsilon^2 + ce^{F(G,u_0)T} \frac{\varepsilon^4}{\lambda}, \tag{3.12}$$

where $F(G, u_0)$ is a function of the Grashof number G and the initial data u_0.

Remark 2 The quantity $F(G, u_0)$ can be bounded above by an adequate power of M where $M = \sup_{t \geq 0} \|u\|_{H^1}$ which is known to be finite for the two-dimensional Navier–Stokes equations [5, 32].

The reverse direction of Theorem 4 has also been proved.

Theorem 5 *Given any $\varepsilon > 0$, we can find $\tilde{u}_N \in \mathcal{F}_N$, such that*

$$\begin{aligned} &\|\tilde{u}_N|_{\partial\Omega}\|^4_{L^4([0,T];H^{1/2}(\partial\Omega))} + \|\tilde{u}_N(x,0) - u_0(x)\|^2_{L^2(\Omega)} \\ &\quad + \|\partial_t \tilde{u}_N - \Delta \tilde{u}_N + \tilde{u}_N \cdot \nabla \tilde{u}_N + \nabla \tilde{p}_N - f\|^2_{L^2(\Omega \times [0,T])} \\ &\quad + \|\nabla \cdot \tilde{u}_N\|^4_{L^4([0,T];L^2(\Omega))} + \lambda \|\tilde{u}_N\|^4_{L^4([0,T];H^1(\Omega))} \leq F(G, u_0) T \varepsilon^2. \end{aligned} \tag{3.13}$$

Furthermore, we prove that our scheme is *approximately* stable.

Theorem 6 *Assume $\tilde{u}_{N_1} \in \mathfrak{F}_{N_1}$ is the approximate solution of*

$$\begin{aligned}\frac{\partial}{\partial t}u_1 - \Delta u_1 + u_1 \cdot \nabla u_1 + \nabla p_1 &= f_1,\\ \nabla \cdot u_1 &= 0,\\ u_1|_{\partial\Omega} &= 0,\\ u_1(x,0) &= u_{0,1}(x).\end{aligned} \tag{3.14}$$

Assume $\tilde{u}_{N_2} \in \mathfrak{F}_{N_2}$ is the approximate solution of

$$\begin{aligned}\frac{\partial}{\partial t}u_2 - \Delta u_2 + u_2 \cdot \nabla u_2 + \nabla p_2 &= f_2,\\ \nabla \cdot u_2 &= 0,\\ u_2|_{\partial\Omega} &= 0,\\ u_2(x,0) &= u_{0,2}(x).\end{aligned} \tag{3.15}$$

Here, $\tilde{u}_{N_1}$ and $\tilde{u}_{N_2}$ satisfy (3.11) with corresponding f_1 and f_2. Then, we have

$$\|u_{N_1} - u_{N_2}\|_{L^4([0,T];L^2(\Omega))} \leq C_1 + C_2 + C_3,$$

where $C_1 = \left((ce^{F(G,u_{0,1})T} + C_{Tr})\varepsilon^2 + ce^{F(G,u_{0,1})T}\frac{\varepsilon^4}{\lambda}\right)^{1/4}$, $C_2 = \left((ce^{F(G,u_{0,2})T} + C_{Tr})\varepsilon^2 + ce^{F(G,u_{0,2})T}\frac{\varepsilon^4}{\lambda}\right)^{1/4}$, $C_3 = c(\|u_{0,1} - u_{0,2}\|_{L^2(\Omega)} + \|f_1 - f_2\|_{L^4([0,T];L^2(\Omega))})$.

4 Elliptic equations: proofs of main theorems

Proof of Theorem 1 We remark first that given any $\epsilon > 0$, from (3.7), there exists a DNN $\mathfrak{F}_N$ of complexity N and $u_N \in \mathfrak{F}_N$ such that $\|u - u_N\|_{H^2(\Omega)} \leq \epsilon$.

$$\begin{aligned}\|\mathcal{L}u_N - f\|^2_{L^2(\Omega)} &= \|\mathcal{L}u_N - \mathcal{L}u\|^2_{L^2(\Omega)}\\ &\leq C_{\mathcal{L}}\|u_N - u\|^2_{H^2(\Omega)}\\ &\leq C_{\mathcal{L}}\varepsilon^2,\end{aligned}$$

where $C_{\mathcal{L}}$ is the operator norm bound of $\mathcal{L}$. Therefore

$$\alpha^2\|\mathcal{L}u_N - f\|^2_{L^2(\Omega)} \leq C_{\mathcal{L}}\alpha^2\varepsilon^2. \tag{4.1}$$

We also have

$$\begin{aligned}\|u_N|_{\partial\Omega}\|^2_{L^2(\partial\Omega)} &= \|u_N|_{\partial\Omega} - u|_{\partial\Omega}\|^2_{L^2(\partial\Omega)}\\ &\leq C_{Tr}\|u_N - u\|^2_{H^2(\Omega)}\\ &\leq C_{Tr}\varepsilon^2,\end{aligned}$$

where C_{Tr} is the constant from the trace operator. Therefore

$$\beta^2 \|u_N|_{\partial\Omega}\|^2_{L^2(\partial\Omega)} \leq C_{Tr}\beta^2\varepsilon^2. \tag{4.2}$$

Combining (4.1) and (4.2), we have

$$\alpha^2 \|\mathcal{L}u_N - f\|^2_{L^2(\Omega)} + \beta^2 \|u_N|_{\partial\Omega}\|^2_{L^2(\partial\Omega)} \leq (C_{\mathcal{L}}\alpha^2 + C_{Tr}\beta^2)\varepsilon^2.$$

Finally,

$$\begin{aligned}\|u_N\|_{H^2(\Omega)} &\leq \|u_N - u\|_{H^2(\Omega)} + \|u\|_{H^2(\Omega)} \\ &\leq \varepsilon + M < 2M =: \widetilde{M}.\end{aligned}$$

Let us consider the converse of Theorem 1. Same as in the previous settings, u is the unique solution of (3.1) and $\mathcal{F}_N$ is a DNN. We have the following results.

Proof of Theorem 2 Since

$$\begin{aligned}\|u - u_N\|_{L^2(\Omega)} &\leq \|u - u_N\|_{H^2(\Omega)} \\ &\leq \|u\|_{H^2(\Omega)} + \|u_N\|_{H^2(\Omega)} \\ &\leq M + \widetilde{M} < cM,\end{aligned}$$

we have

$$\begin{aligned}\|(u - u_N)|_{\partial\Omega}\|_{H^{1/2}(\partial\Omega)} &\leq c\|(u - u_N)|_{\partial\Omega}\|^{2/3}_{L^2(\partial\Omega)}\|(u - u_N)|_{\partial\Omega}\|^{1/3}_{H^{3/2}(\partial\Omega)} \\ &\leq c\frac{\varepsilon^{2/3}}{\beta^{2/3}}\|u - u_N\|^{1/3}_{H^2(\Omega)} \\ &\leq cM^{1/3}\frac{\varepsilon^{2/3}}{\beta^{2/3}}.\end{aligned} \tag{4.3}$$

Denote $\mathcal{L}u_N = f_\varepsilon$ and $u_N|_{\partial\Omega} = Tr(u_N) = g_\varepsilon$. From (3.8), we have $\|f - f_\varepsilon\|_{L^2(\Omega)} \leq \frac{\varepsilon}{\alpha}$, from (4.3), we have $\|g_\varepsilon\|_{H^{1/2}(\partial\Omega)} < cM^{1/3}\frac{\varepsilon^{2/3}}{\beta^{2/3}}$.

Consider the lifting operator $l_\Omega : H^{1/2}(\partial\Omega) \to H^1(\Omega)$, which is linear and bounded such that $Trl_\Omega = I$. Here, Tr is the trace operator and I is the identity operator. Let $\tilde{u}_N = u_N - l_\Omega g_\varepsilon$. Then

$$\begin{aligned}\mathcal{L}\tilde{u}_N &= f_\varepsilon - \mathcal{L}l_\Omega g_\varepsilon, \\ \tilde{u}_N|_{\partial\Omega} &= 0.\end{aligned} \tag{4.4}$$

Note that since $\mathcal{L}$ is a second order elliptic operator and $l_\Omega g_\varepsilon \in H^1(\Omega)$, we have $\mathcal{L}l_\Omega g_\varepsilon \in H^{-1}(\Omega)$. From Lax-Milgram [6], we have

$$\|u - \tilde{u}_N\|_{H^1(\Omega)} \leq C_{\mathcal{L}^{-1}}\|(f - f_\varepsilon) + \mathcal{L}l_\Omega g_\varepsilon\|_{H^{-1}(\Omega)}. \tag{4.5}$$

Therefore, we have

$$\begin{aligned}
&\|u - u_N\|_{H^1(\Omega)} \\
&\quad = \|u - (u_N - l_\Omega g_\varepsilon) - l_\Omega g_\varepsilon\|_{H^1(\Omega)} \\
&\quad \le \|u - \tilde{u}_N\|_{H^1(\Omega)} + \|l_\Omega g_\varepsilon\|_{H^1(\Omega)} \\
&\quad \le C_{\mathcal{L}^{-1}} \|(f - f_\varepsilon) + \mathcal{L} l_\Omega g_\varepsilon)\|_{H^{-1}(\Omega)} + \|l_\Omega\| \|g_\varepsilon\|_{H^{1/2}(\partial\Omega)}. \qquad (4.6)
\end{aligned}$$

Since

$$\|f - f_\varepsilon\|_{H^{-1}(\Omega)} \le c\|f - f_\varepsilon\|_{L^2(\Omega)}$$

and

$$\|\mathcal{L} l_\Omega g_\varepsilon\|_{H^{-1}(\Omega)} \le C_{\mathcal{L}} \|l_\Omega g_\varepsilon\|_{H^1(\Omega)} \le C_{\mathcal{L}} \|l_\Omega\| \|g_\varepsilon\|_{H^{1/2}(\partial\Omega)}.$$

Thus

$$\begin{aligned}
\|u - u_N\|_{H^1(\Omega)} &\le C_{\mathcal{L}^{-1}} \|f - f_\varepsilon\|_{L^2(\Omega)} + C_{\mathcal{L}^{-1}} C_{\mathcal{L}} \|l_\Omega\| \|g_\varepsilon\|_{H^{1/2}(\partial\Omega)} \\
&\le C_{\mathcal{L}^{-1}} \frac{\varepsilon}{\alpha} + C_{\mathcal{L}^{-1}} C_{\mathcal{L}} \|l_\Omega\| M^{1/3} \frac{\varepsilon^{2/3}}{\beta^{2/3}} \\
&\le \left(\frac{C_{\mathcal{L}^{-1}}}{\alpha} + \frac{C_{\mathcal{L}^{-1}} C_{\mathcal{L}} \|l_\Omega\| M^{1/3}}{\beta^{2/3}} \right) \varepsilon^{2/3}. \qquad (4.7)
\end{aligned}$$

We will show below that by considering the loss function to be

$$\alpha^2 \|\mathcal{L} u_N - f\|^2_{L^2(\Omega)} + \beta^2 \|u_N|_{\partial\Omega}\|^2_{H^{1/2}(\partial\Omega)},$$

we have an improved result on $\|u - u_N\|_{H^1(\Omega)}$ which is stated in Theorem 3.

Proof of Theorem 3 Same as before, $\mathcal{L} u_N = f_\varepsilon$ and $u_N|_{\partial\Omega} = Tr(u_N) = g_\varepsilon$. From (3.9), we have $\|f - f_\varepsilon\|_{L^2(\Omega)} \le \frac{\varepsilon}{\alpha}$, and $\|g_\varepsilon\|_{H^{1/2}(\partial\Omega)} \le \frac{\varepsilon}{\beta}$. Let $\tilde{u}_N = u_N - l_\Omega g_\varepsilon$.

$$\begin{aligned}
\mathcal{L} \tilde{u}_N &= f_\varepsilon - \mathcal{L} l_\Omega g_\varepsilon, \\
\tilde{u}_N|_{\partial\Omega} &= 0.
\end{aligned} \qquad (4.8)$$

Similar to the Proof of Theorem 2, we have

$$\begin{aligned}
\|u - u_N\|_{H^1(\Omega)} &\le C_{\mathcal{L}^{-1}} \|(f - f_\varepsilon) + \mathcal{L} l_\Omega g_\varepsilon)\|_{H^{-1}(\Omega)} + \|l_\Omega\| \|g_\varepsilon\|_{H^{1/2}(\partial\Omega)} \\
&\le C_{\mathcal{L}^{-1}} \|f - f_\varepsilon\|_{L^2(\Omega)} + C_{\mathcal{L}^{-1}} C_{\mathcal{L}} \|l_\Omega\| \|g_\varepsilon\|_{H^{1/2}(\partial\Omega)} \\
&\le \left(\frac{C_{\mathcal{L}^{-1}}}{\alpha} + \frac{C_{\mathcal{L}^{-1}} C_{\mathcal{L}} \|l_\Omega\|}{\beta} \right) \varepsilon.
\end{aligned}$$

5 Navier–Stokes equations: proof of main theorems

5.1 Functional analytic framework

Let Ω be a bounded domain in $\mathbb{R}^2$ and

$$H := \left\{ u \in L^2(\Omega) | \; \nabla \cdot u = 0, \; \gamma(u) = 0 \right\},$$
$$V := \left\{ u \in H_0^1(\Omega) | \; \nabla \cdot u = 0 \right\}.$$

V' is the dual space of V.

Let $\mathbb{P}$ be the Leray Projection which is an orthogonal projection from L^2 onto the subset of L^2 consisting of those functions whose weak derivatives are divergence-free in the L^2 sense. A is the Stokes operator, defined as $A = -\mathbb{P}\Delta$. B is the bilinear form defined by $B(u, u) = \mathbb{P}\left[(u \cdot \nabla)u\right]$.

Applying the projection $\mathbb{P}$ on (3.10), the functional form of the NSE can be written as

$$\begin{aligned} &\frac{du}{dt} + Au + B(u, u) = \mathbb{P}f, \\ &u|_{\partial\Omega} = 0, \\ &u(x, 0) = u_0(x). \end{aligned} \tag{5.1}$$

We recall the definition of strong solutions from [31]:

Let $W = \left\{ u \in H^1_{loc}(\Omega) \text{ and } \nabla \cdot u = 0 \text{ in } \Omega \right\}$ and $u_0 \in W$, u is a strong solution of NSE if it solves the variational formulation of (3.10) as in [5, 31], and

$$u \in L^2(0, T; D(A)) \cap L^\infty(0, T; W),$$

for $T > 0$.

5.2 Hodge decomposition

The idea of Hodge decomposition is to decompose a vector $u \in L^2(\Omega)$ uniquely into a divergence-free part u_1 and an irrotational part u_2, which is orthogonal in $L^2(\Omega)$ to u_1:

$$u = u_1 + u_2, \;\; \nabla \cdot u_1 = 0, \;\text{ and } (u_1, u_2) = 0.$$

When we apply the Leray Projection $\mathbb{P}$ on u, we have

$$\mathbb{P}u = u_1.$$

More precisely, we have the following proposition, the proof of which can be found in [5].

Proposition 1 *Let Ω be open, bounded, connected with boundary of class C^2. Then $L^2(\Omega) = H \oplus H_1 \oplus H_2$, where H, H_1, H_2 are mutually orthogonal spaces and moreover*

$$H_1 = \{u \in L^2(\Omega) | \ u = \nabla p, \ p \in H^1(\Omega), \ \Delta p = 0\},$$

and

$$H_2 = \{u \in L^2(\Omega) | \ u = \nabla p, \ p \in H_0^1(\Omega)\}.$$

The decomposition above is obtained as follows. Let $v \in L^2(\Omega)$. Then,

$$v = u + u_1 + u_2, \ \ u \in H, \ \text{and} \ u_2 = \nabla p_2, \ \ \Delta p_2 = \nabla \cdot v \in H^{-1}(\Omega), \ \ p_2 \in H_0^1(\Omega).$$

Subsequently, u_1 is obtained by solving the Neumann problem

$$u_1 = \nabla p_1, \ \ \Delta p_1 = 0, \ \ \frac{\partial p_1}{\partial n_\Omega} = \gamma(v - u_2),$$

where n_Ω is the unit normal vector on the boundary of Ω and γ denotes the normal trace on the boundary (see [5, 32] for more details).

5.3 Proofs

Consider an approximate solution $\tilde{u}_N \in \mathcal{F}_N$, i.e. $\tilde{u}_N$ satisfies (3.11) and denote $\tilde{u}_N|_{\partial\Omega} = \tilde{g}, \ \nabla \cdot \tilde{u}_N = \tilde{h}$. Let $\tilde{f} := \partial_t \tilde{u}_N - \Delta \tilde{u}_N + \tilde{u}_N \cdot \nabla \tilde{u}_N + \nabla \tilde{p}_N - f$. Then

$$\begin{aligned} \partial_t \tilde{u}_N - \Delta \tilde{u}_N + \tilde{u}_N \cdot \nabla \tilde{u}_N + \nabla \tilde{p}_N &= f + \tilde{f}, \\ \nabla \cdot \tilde{u}_N &= \tilde{h}, \\ \tilde{u}_N|_{\partial\Omega} &= \tilde{g}. \end{aligned} \tag{5.2}$$

Applying the Hodge decomposition on $\tilde{u}_N$:

$$\tilde{u}_N = \mathbb{P}\tilde{u}_N + (\mathbb{I} - \mathbb{P})\tilde{u}_N =: u_N + v_N, \tag{5.3}$$

where $u_N = \mathbb{P}\tilde{u}_N$, $\nabla \cdot u_N = 0$, $u_N|_{\partial\Omega} = 0$, and $v_N = (\mathbb{I} - \mathbb{P})\tilde{u}_N$.

Before we prove our main theorems, we first introduce two Lemmas.

Lemma 1 *Consider u_N satisfying*

$$\begin{aligned} &\frac{du_N}{dt} + Au_N + B(u_N, u_N) = \mathbb{P}f + \varphi, \\ &u_N|_{\partial\Omega} = 0. \end{aligned} \tag{5.4}$$

where

$$\int_0^T \|\varphi\|_{V'}^2 \, dt \leq O\left(\varepsilon + \frac{\varepsilon^2}{\sqrt{\lambda}}\right), \tag{5.5}$$

and let u be a strong solution of (5.1), *with* $\|u_N(x,0) - u_0(x)\|_{L^2}^2 \leq \varepsilon^2$.
Then

$$\sup_{[0,T]} \|u(x,t) - u_N(x,t)\|_{L^2(\Omega)}^2 \leq ce^{F(G,u_0)T}\left(\varepsilon + \frac{\varepsilon^2}{\sqrt{\lambda}}\right),$$

where $F(G,u_0)$ *is a function of the Grashof number* G *and the initial data* u_0.

Proof Considering $w(t) = u(t) - u_N(t)$, from (5.1) and (5.4), we have

$$\begin{aligned} &\frac{dw}{dt} + Aw + B(u,u) - B(u_N,u_N) = -\varphi, \\ &w|_{\partial\Omega} = 0. \end{aligned} \tag{5.6}$$

Since

$$B(u,u) - B(u_N,u_N) = B(u,w) + B(w,u_N) = B(u,w) + B(w,u) - B(w,w),$$

we have

$$\frac{dw}{dt} + Aw + B(u,w) + B(w,u) - B(w,w) = -\varphi. \tag{5.7}$$

Taking inner product of (5.7) with w, we obtain

$$\frac{1}{2}\frac{d}{dt}\|w\|_{L^2}^2 + \|A^{1/2}w\|_{L^2}^2 + (B(w,u),w) = -(\varphi,w).$$

Therefore

$$\frac{1}{2}\frac{d}{dt}\|w\|_{L^2}^2 + \|A^{1/2}w\|_{L^2}^2 \leq |(B(w,u),w)| + |(\varphi,w)|.$$

Since

$$\begin{aligned} |(B(w,u),w)| &\leq c\|A^{1/2}u\|_{L^2}\|w\|_{L^2}\|A^{1/2}w\|_{L^2} \\ &\leq c\frac{\|A^{1/2}u\|_{L^2}^2\|w\|_{L^2}^2}{2} + \frac{\|A^{1/2}w\|_{L^2}^2}{2}, \end{aligned}$$

and

$$|(\varphi,w)| \leq \|\varphi\|_{V'}\|A^{1/2}w\|_{L^2} \leq \frac{\|\varphi\|_{V'}^2}{2} + \frac{\|A^{1/2}w\|_{L^2}^2}{2},$$

we have

$$\frac{1}{2}\frac{d}{dt}\|w\|_{L^2}^2 - c\frac{\|A^{1/2}u\|_{L^2}^2\|w\|_{L^2}^2}{2} \leq \frac{\|\varphi\|_{V'}^2}{2}.$$

Equivalently

$$\frac{d}{dt}\|w\|_{L^2}^2 - c\|A^{1/2}u\|_{L^2}^2\|w\|_{L^2}^2 \le \|\varphi\|_{V'}^2.$$

Applying Gronwall's inequality, we have

$$\|w(t)\|_{L^2}^2 \le e^{\int_0^t c\|A^{1/2}u\|_{L^2}^2 ds}\|w(0)\|_{L^2}^2 + e^{\int_0^t c\|A^{1/2}u\|_{L^2}^2 ds}\int_0^t e^{-\int_0^s c\|A^{1/2}u\|_{L^2}^2 d\tau}\|\varphi\|_{V'}^2 ds.$$

Since $\int_0^t c\|A^{1/2}u\|_{L^2}^2 ds \le F(G, u_0)t$ where, $F(G, u_0)$ is a function of the Grashof number G and the initial data u_0 [5, 32], we have

$$e^{\int_0^t c\|A^{1/2}u\|_{L^2}^2 ds} \le e^{F(G,u_0)t}.$$

Moreover

$$e^{-\int_0^s c\|A^{1/2}u\|_{L^2}^2 d\tau} \le 1,$$

and

$$\|w(0)\|_{L^2}^2 = \|u_N(x, 0) - u_0(x)\|_{L^2}^2 \le \varepsilon^2.$$

Therefore, we have

$$\|w(t)\|_{L^2}^2 \le e^{F(G,u_0)t}\|w(0)\|_{L^2}^2 + e^{F(G,u_0)t)}\int_0^t ||\varphi||_{V'}^2 ds \le ce^{F(G,u_0)t,}\left(\varepsilon + \frac{\varepsilon^2}{\sqrt{\lambda}}\right).$$

Therefore

$$\sup_{[0,T]}\|w(t)\|_{L^2}^2 \le ce^{F(G,u_0)T}\left(\varepsilon + \frac{\varepsilon^2}{\sqrt{\lambda}}\right).$$

Lemma 2 *Assume*

$$\|\tilde{u}_N|_{\partial\Omega}\|_{L^4([0,T];H^{1/2}(\partial\Omega))}^4 + \|\nabla\cdot\tilde{u}_N\|_{L^4([0,T];L^2(\Omega))}^4 \le \varepsilon^2, \tag{5.8}$$

and with the Hodge decomposition (5.3)*, we have*

$$\|v_N\|_{L^4([0,T];H^1(\Omega))} \le C_{Tr}\sqrt{\varepsilon}. \tag{5.9}$$

Proof Consider the Hodge decomposition

$$\tilde{u}_N = u_N + v_N = u_N \oplus v_1 \oplus v_2$$

with

$$v_N = v_1 \oplus v_2,$$

where $v_1 = \nabla p_1$ and $v_2 = \nabla p_2$, p_1 and p_2 are the solutions of the following two systems, respectively.

$$\begin{cases} \Delta p_1 = 0, \\ \dfrac{\partial p_1}{\partial n} = \gamma(\tilde{u}_N - v_2), \end{cases} \tag{5.10}$$

and

$$\begin{cases} \Delta p_2 = \nabla \cdot \tilde{u}_N, \\ Tr(p_2) = 0. \end{cases} \tag{5.11}$$

Here, γ is the trace operator and $Tr(p_2)$ means the value of p_2 on the boundary. According to Lions and Magenes [20], the above two systems have unique solutions (up to an additive constant). First, we solve for p_2 from (5.11). Accordingly, v_2 can be obtained. Then, we use v_2 to solve for p_1 from (5.10) and find v_1 afterwards.

From (5.8) and (5.11), we have

$$\begin{aligned} \|v_2\|_{L^4([0,T];H^1(\Omega))} &\le \|p_2\|_{L^4([0,T];H^2(\Omega))} \\ &\le c\|\nabla \cdot \tilde{u}_N\|_{L^4([0,T];L^2(\Omega))} \\ &\le c\varepsilon^{1/2}. \end{aligned}$$

From (5.8) and (5.10),

$$\begin{aligned} \|v_1\|_{L^4([0,T];H^1(\Omega))} &\le \|p_1\|_{L^4([0,T];H^2(\Omega))} \\ &\le c\|\gamma(\tilde{u}_N - v_2)\|_{L^4([0,T];H^{1/2}(\Omega))} \\ &\le c\|\tilde{u}_N|_{\partial\Omega}\|_{L^4([0,T];H^{1/2}(\partial\Omega))} + c\|\gamma(v_2)\|_{L^4([0,T];H^{1/2}(\Omega))} \\ &\le c\varepsilon^{1/2} + c\|Tr(v_2)\cdot n_\Omega\|_{L^4([0,T];H^{1/2}(\Omega))} \\ &\le c\varepsilon^{1/2} + C_{Tr}\|v_2\|_{L^4([0,T];H^1(\Omega))} \\ &\le C_{Tr}\sqrt{\varepsilon}. \end{aligned} \tag{5.12}$$

Therefore,

$$\|v_N\|_{L^4([0,T];H^1(\Omega))} = \|v_1\|_{L^4([0,T];H^1(\Omega))} + \|v_2\|_{L^4([0,T];H^1(\Omega))} \le C_{Tr}\sqrt{\varepsilon}.$$

Note that (5.9) also implies

$$\left(\int_0^T \|\nabla v_N\|^4_{L^2(\Omega)} dt\right)^{1/4} \le C_{Tr}\sqrt{\varepsilon},\ \left(\int_0^T \|v_N\|^4_{L^2(\Omega)} dt\right)^{1/4} \le C_{Tr}\sqrt{\varepsilon}.$$

Proof of Theorem 4 Applying the Leray projection operator $\mathbb{P}$ on (5.2), one obtains, under the assumption that $\mathbb{P}f = f$,

$$\partial_t \mathbb{P}\tilde{u}_N - \mathbb{P}\Delta\tilde{u}_N + \mathbb{P}\tilde{u}_N \cdot \nabla\tilde{u}_N = f + \mathbb{P}\tilde{f}.$$

Recall that $\mathbb{P}$ is the orthogonal projection. By Hodge decomposition

$$\tilde{u}_N = \mathbb{P}\tilde{u}_N + (\mathbb{I} - \mathbb{P})\tilde{u}_N =: u_N + v_N.$$

Then

$$\partial_t u_N + Au_N + \mathbb{P}u_N \cdot \nabla u_N + \mathbb{P}(\tilde{u}_N \cdot \nabla\tilde{u}_N - u_N \cdot \nabla u_N) = f + \mathbb{P}\tilde{f} + \mathbb{P}\Delta v_N.$$

Here, A is the Stokes' operator.

$$\begin{aligned} \mathbb{P}\left(\tilde{u}_N \cdot \nabla\tilde{u}_N - u_N \cdot \nabla u_N\right) &= \mathbb{P}\left((\tilde{u}_N - u_N) \cdot \nabla\tilde{u}_N + u_N \cdot \nabla(\tilde{u}_N - u_N)\right) \\ &= \mathbb{P}\left(v_N \cdot \nabla\tilde{u}_N + u_N \cdot \nabla v_N\right) \\ &=: \psi. \end{aligned} \tag{5.13}$$

Next, we will estimate $\int_0^T \|\psi\|_{V'}^2 \, dt$.

Note that $\|\psi\|_{V'} = \sup_{w \in V, \|w\|_V \le 1} \langle \psi, w \rangle$, where

$$\langle \psi, w \rangle = \int_\Omega \mathbb{P}(v_N \cdot \nabla\tilde{u}_N + u_N \cdot \nabla v_N) \cdot w \, dx. \tag{5.14}$$

We estimate (5.14) term by term. Since w is divergence free, we have

$$\begin{aligned} \int_\Omega \mathbb{P}(v_N \cdot \nabla\tilde{u}_N) \cdot w \, dx &= \int_\Omega (v_N \cdot \nabla\tilde{u}_N) \cdot w \, dx \\ &\le \|\nabla\tilde{u}_N\|_{L^2(\Omega)} \|v_N\|_{L^4(\Omega)} \|w\|_{L^4(\Omega)}. \end{aligned}$$

By Sobolev inequality, $\|w\|_{L^4(\Omega)} \le c\|w\|_V \le c$ and thus

$$\begin{aligned} \|\mathbb{P}(v_N \cdot \nabla\tilde{u}_N)\|_{V'} &\le c\|\nabla\tilde{u}_N\|_{L^2(\Omega)} \|v_N\|_{L^4(\Omega)} \\ &\le c\|\nabla\tilde{u}_N\|_{L^2(\Omega)} \|v_N\|_{L^2(\Omega)}^{1/2} \|\nabla v_N\|_{L^2(\Omega)}^{1/2}, \end{aligned} \tag{5.15}$$

where in the last line, we used the Ladyzhenskaya's inequality [5].

Therefore

$$\begin{aligned} \int_0^T \|\mathbb{P}(v_N \cdot \nabla\tilde{u}_N)\|_{V'}^2 \, dt &\le c \int_0^T \|\nabla\tilde{u}_N\|_{L^2(\Omega)}^2 \|v_N\|_{L^2(\Omega)} \|\nabla v_N\|_{L^2(\Omega)} \, dt \\ &\le c \left(\int_0^T \|\nabla\tilde{u}_N\|_{L^2(\Omega)}^4 dt \right)^{1/2} \left(\int_0^T \|v_N\|_{L^2(\Omega)}^4 dt \right)^{1/4} \left(\int_0^T \|\nabla v_N\|_{L^2(\Omega)}^4 dt \right)^{1/4}. \end{aligned}$$

Since

$$\lambda\|\tilde{u}_N\|^4_{L^4([0,T];H^1(\Omega))} \leq \varepsilon^2,$$

we have

$$\int_0^T \|\nabla\tilde{u}_N\|^4_{L^2(\Omega)}dt \leq \frac{\varepsilon^2}{\lambda}.$$

Therefore

$$\left(\int_0^T \|\nabla\tilde{u}_N\|^4_{L^2(\Omega)}dt\right)^{1/2} \leq \frac{\varepsilon}{\sqrt{\lambda}}. \tag{5.16}$$

Applying Lemma 2, we have

$$\left(\int_0^T \|\nabla v_N\|^4_{L^2(\Omega)}dt\right)^{1/4} \leq C_{Tr}\sqrt{\varepsilon},\ \left(\int_0^T \|v_N\|^4_{L^2(\Omega)}dt\right)^{1/4} \leq C_{Tr}\sqrt{\varepsilon}.$$

Therefore

$$\int_0^T \|\mathbb{P}(v_N\cdot\nabla\tilde{u}_N)\|^2_{V'}\,dt \leq \frac{C_{Tr}\varepsilon^2}{\sqrt{\lambda}}. \tag{5.17}$$

Next, we estimate the second term of (5.14): $\int_\Omega \mathbb{P}(u_N\cdot\nabla v_N)\cdot w\,dx$.
Similarly, we have

$$\|\mathbb{P}(u_N\cdot\nabla v_N)\|_{V'} \leq \|\nabla v_N\|_{L^2(\Omega)}\|u_N\|_{L^4(\Omega)} \leq \|\nabla v_N\|_{L^2(\Omega)}\|u_N\|_{H^1(\Omega)}.$$

Therefore

$$\begin{aligned}\int_0^T \|\mathbb{P}(u_N\cdot\nabla v_N)\|^2_{V'}\,dt &\leq \int_0^T \|\nabla v_N\|^2_{L^2(\Omega)}\|u_N\|^2_{H^1(\Omega)}\,dt\\ &\leq \left(\int_0^T \|\nabla v_N\|^4_{L^2(\Omega)}dt\right)^{1/2}\left(\int_0^T \|u_N\|^4_{H^1(\Omega)}dt\right)^{1/2}.\end{aligned}$$

Since

$$\left(\int_0^T \|\nabla v_N\|^4_{L^2(\Omega)}dt\right)^{1/2} \leq C_{Tr}\varepsilon,$$

and

$$\left(\int_0^T \|u_N\|^4_{H^1(\Omega)}dt\right)^{1/2} \leq \left(\int_0^T \|\tilde{u}_N\|^4_{H^1(\Omega)}dt\right)^{1/2} \leq \frac{\varepsilon}{\sqrt{\lambda}}.$$

Therefore

$$\int_0^T \|\mathbb{P}(u_N\cdot\nabla v_N)\|^2_{V'}\,dt \leq \frac{C_{Tr}\varepsilon^2}{\sqrt{\lambda}}. \tag{5.18}$$

Combining (5.17) and (5.18), we have

$$\int_0^T \|\psi\|_{V'}^2\,dt \leq \frac{C_{Tr}\varepsilon^2}{\sqrt{\lambda}}. \tag{5.19}$$

Moreover, since

$$\|\mathbb{P}(\Delta v_N)\|_{V'} \leq \|\nabla v_N\|_{L^2(\Omega)}\|w\|_{H^1(\Omega)}, \tag{5.20}$$

we have

$$\begin{aligned}\int_0^T \|\mathbb{P}(\Delta v_N)\|_{V'}^2\,dt &\leq \int_0^T \|\nabla v_N\|_{L^2(\Omega)}^2\|w\|_{H^1(\Omega)}^2\,dt \\ &\leq \left(\int_0^T \|\nabla v_N\|_{L^2(\Omega)}^4 dt\right)^{1/2}\left(\int_0^T \|w\|_{H^1(\Omega)}^4 dt\right)^{1/2} \\ &\leq C_{Tr}\varepsilon. \end{aligned}\tag{5.21}$$

Denoting $\varphi := \mathbb{P}\tilde{f} + \mathbb{P}(\Delta v_N) - \psi$, we have

$$\frac{du_N}{dt} + Au_N + B(u_N, u_N) = \mathbb{P}f + \varphi.$$

Since

$$\int_0^T \|\mathbb{P}(\Delta v_N)\|_{V'}^2\,dt \leq C_{Tr}\varepsilon,\ \int_0^T \|\psi\|_{V'}^2\,dt \leq \frac{C_{Tr}\varepsilon^2}{\sqrt{\lambda}},$$
$$\int_0^T \|\mathbb{P}\tilde{f}\|_{V'}^2\,dt \leq \int_0^T \|\mathbb{P}\tilde{f}\|_{L^2}^2\,dt \leq \varepsilon^2,$$

we obtain

$$\int_0^T \|\varphi\|_{V'}^2\,dt \leq C_{Tr}\left(\varepsilon + \frac{\varepsilon^2}{\sqrt{\lambda}}\right).$$

Since $\|u_N(x,0) - u_0(x)\|_{L^2}^2 \leq \|\tilde{u}_N(x,0) - u_0(x)\|_{L^2}^2 \leq \varepsilon^2$. Applying Lemma 1, we have

$$\sup_{[0,T]} \|u(t) - u_N(t)\|_{L^2}^2 \leq ce^{F(G,u_0)T}\left(\varepsilon + \frac{\varepsilon^2}{\sqrt{\lambda}}\right).$$

Moreover, since

$$\left(\int_0^T \|v_N\|_{L^2(\Omega)}^4 dt\right)^{1/4} \leq C_{Tr}\sqrt{\varepsilon},$$

we have

$$\int_0^T \|u - \tilde{u}_N\|^4_{L^2(\Omega)} dt \le \int_0^T \|u - u_N\|^4_{L^2(\Omega)} dt + \int_0^T \|v_N\|^4_{L^2(\Omega)} dt$$
$$\le ce^{F(G,u_0)T}\left(\varepsilon^2 + \frac{\varepsilon^4}{\lambda}\right) + C_{Tr}\varepsilon^2$$
$$= (ce^{F(G,u_0)T} + C_{Tr})\varepsilon^2 + ce^{F(G,u_0)T}\frac{\varepsilon^4}{\lambda}.$$

Lemma 3 *Given* $\varepsilon > 0$*, assume that* (u, p) *satisfies (3.10), then, there exists* $(\tilde{u}_N, \tilde{p}_N) \in \mathfrak{F}_N$ *satisfying*

$$\sup_{t\in[0,T]} \|u - \tilde{u}_N\|_{L^2(\Omega)} \le \varepsilon,\ \|u - \tilde{u}_N\|_{H^{1,2}(\Omega\times[0,T])} \le \varepsilon,$$
$$\left(\int_0^T \|u - \tilde{u}_N\|^4_{W^{1,4}(\Omega)} dt\right)^{1/4} \le \varepsilon,\ \|p - \tilde{p}_N\|_{L^2([0,T];H^1(\Omega))} \le \varepsilon. \tag{5.22}$$

Proof From 3.7, as long as the solution (u, p) of the NSEs belongs to the spaces in (5.22), we can find $(\tilde{u}_N, \tilde{p}_N) \in \mathfrak{F}_N$ as smooth as we want and close to (u, p), which means (5.22) holds. From the classical results of the 2-D NSEs, we know that we can find the solution (u, p) that belongs to the spaces in (5.22).

Proof of Theorem 5 From Lemma 3, given $\varepsilon > 0$, assume that u is a strong solution of (3.10) and there is an N such that $\tilde{u}_N \in \mathfrak{F}_N$ satisfying

$$\sup_{t\in[0,T]} \|u - \tilde{u}_N\|_{L^2(\Omega)} \le \varepsilon,\ \|u - \tilde{u}_N\|_{H^{1,2}(\Omega\times[0,T])} \le \varepsilon,$$
$$\left(\int_0^T \|u - \tilde{u}_N\|^4_{W^{1,4}(\Omega)} dt\right)^{1/4} \le \varepsilon, \|p - \tilde{p}_N\|_{L^2([0,T];H^1(\Omega))} \le \varepsilon.$$

Now, we estimate the left hand side of (3.13) term by term:

Since $\sup_{t\in[0,T]} \|u - \tilde{u}_N\|_{L^2(\Omega)} \le \varepsilon$, we have

$$\|u(x,0) - \tilde{u}_N(x,0)\|^2_{L^2(\Omega)} \le \varepsilon^2. \tag{5.23}$$

Since $\nabla \cdot u = 0$, we have

$$\|\nabla \cdot \tilde{u}_N\|^4_{L^4([0,T];L^2(\Omega))} = \|\nabla \cdot \tilde{u}_N - \nabla \cdot u\|^4_{L^4([0,T];L^2(\Omega))}$$
$$\le c\|\tilde{u}_N - u\|^4_{L^4([0,T];H^1(\Omega))}$$
$$\le c\varepsilon^4. \tag{5.24}$$

Observe that

$$\lambda\|\tilde{u}_N\|^4_{L^4([0,T];H^1(\Omega))} \le c\lambda\|\tilde{u}_N - u\|^4_{L^4([0,T];H^1(\Omega))} + c\lambda\|u\|^4_{L^4([0,T];H^1(\Omega))}$$

$$\leq c\lambda\varepsilon^4 + \lambda F(G, u_0)T.$$

where $F(G, u_0)$ is a suitable function of G and u_0 (here we have used the fact that $\sup_{t\geq 0}\|u\|_{H^1} < \infty$ and can be expressend as a suitable function of u_0 and G [5, 32]). Picking λ small enough, we have

$$\lambda\|\tilde{u}_N\|^4_{L^4([0,T];H^1(\Omega))} \leq F(G, u_0)T\varepsilon^4. \tag{5.25}$$

Consider

$$\begin{aligned}
&\|\partial_t\tilde{u}_N - \Delta\tilde{u}_N + \tilde{u}_N\cdot\nabla\tilde{u}_N + \nabla\tilde{p}_N - f\|^2_{L^2(\Omega\times[0,T])} \\
&\quad= \|\partial_t\tilde{u}_N - \Delta\tilde{u}_N + \tilde{u}_N\cdot\nabla\tilde{u}_N + \nabla\tilde{p}_N - (\partial_t u - \Delta u + u\cdot\nabla u + \nabla p)\|^2_{L^2(\Omega\times[0,T])} \\
&\quad\leq c\|\partial_t\tilde{u}_N - \partial_t u\|^2_{L^2(\Omega\times[0,T])} + c\|\Delta u - \Delta\tilde{u}_N\|^2_{L^2(\Omega\times[0,T])} \\
&\qquad + c\|\tilde{u}_N\cdot\nabla\tilde{u}_N - u\cdot\nabla u\|^2_{L^2(\Omega\times[0,T])} + c\|\nabla\tilde{p}_N - \nabla p\|^2_{L^2(\Omega\times[0,T])}.
\end{aligned} \tag{5.26}$$

We have

$$\begin{aligned}
\|\partial_t\tilde{u}_N - \partial_t u\|^2_{L^2(\Omega\times[0,T])} &\leq \varepsilon^2, \\
\|\Delta u - \Delta\tilde{u}_N\|^2_{L^2(\Omega\times[0,T])} &\leq \varepsilon^2,
\end{aligned}$$

and

$$\|\nabla\tilde{p}_N - \nabla p\|^2_{L^2(\Omega\times[0,T])} \leq \varepsilon^2.$$

Moreover

$$\begin{aligned}
&\|\tilde{u}_N\cdot\nabla\tilde{u}_N - u\cdot\nabla u\|^2_{L^2(\Omega\times[0,T])} \\
&\quad= \|\tilde{u}_N\cdot\nabla\tilde{u}_N - u\cdot\nabla\tilde{u}_N + u\cdot\nabla\tilde{u}_N - u\cdot\nabla u\|^2_{L^2(\Omega\times[0,T])} \\
&\quad= \|(\tilde{u}_N - u)\cdot\nabla\tilde{u}_N + u\cdot(\nabla\tilde{u}_N - \nabla u)\|^2_{L^2(\Omega\times[0,T])} \\
&\quad\leq \|\tilde{u}_N - u\|^2_{L^4(\Omega\times[0,T])}\|\nabla\tilde{u}_N\|^2_{L^4(\Omega\times[0,T])} \\
&\qquad + \|u\|^2_{L^4(\Omega\times[0,T])}\|\nabla\tilde{u}_N - \nabla u\|^2_{L^4(\Omega\times[0,T])}.
\end{aligned} \tag{5.27}$$

We have

$$\begin{aligned}
\|\nabla\tilde{u}_N\|_{L^4(\Omega\times[0,T])} &\leq \|\nabla\tilde{u}_N - \nabla u\|_{L^4(\Omega\times[0,T])} + \|\nabla u\|_{L^4(\Omega\times[0,T])} \\
&\leq c\varepsilon + F(G, u_0)T \lesssim F(G, u_0)T.
\end{aligned}$$

Moreover,

$$\begin{aligned}
&\|\tilde{u}_N - u\|_{L^4(\Omega\times[0,T])} \leq c\varepsilon,\ \ \|u\|_{L^4(\Omega\times[0,T])} \leq F(G, u_0)T \text{ and} \\
&\|\nabla\tilde{u}_N - \nabla u\|_{L^4(\Omega\times[0,T])} \leq c\varepsilon.
\end{aligned}$$

Thus

$$\|\tilde{u}_N \cdot \nabla \tilde{u}_N - u \cdot \nabla u\|^2_{L^2(\Omega\times[0,T])} \leq F(G, u_0)T\varepsilon^2.$$

Therefore

$$\|\partial_t \tilde{u}_N - \Delta \tilde{u}_N + \tilde{u}_N \cdot \nabla \tilde{u}_N + \nabla \tilde{p}_N - f\|^2_{L^2(\Omega\times[0,T])} \leq F(G, u_0)T\varepsilon^2. \tag{5.28}$$

Moreover

$$\begin{aligned}\|\tilde{u}_N|_{\partial\Omega}\|^4_{L^4([0,T];H^{1/2}(\partial\Omega))} &= \|\tilde{u}_N|_{\partial\Omega} - u|_{\partial\Omega}\|^4_{L^4([0,T];H^{1/2}(\partial\Omega))} \\ &\leq c\|\tilde{u}_N - u\|^4_{L^4([0,T];H^1(\Omega))} \\ &\leq c\varepsilon^4. \end{aligned} \tag{5.29}$$

In summary, combining (5.23)–(5.29), we have

$$\begin{aligned}&\|\tilde{u}_N|_{\partial\Omega}\|^4_{L^4([0,T];H^{1/2}(\partial\Omega))} + \|\tilde{u}_N(x,0) - u_0(x)\|^2_{L^2(\Omega)} \\ &\quad + \|\partial_t \tilde{u}_N - \Delta \tilde{u}_N + \tilde{u}_N \cdot \nabla \tilde{u}_N + \nabla \tilde{p}_N - f\|^2_{L^2(\Omega\times[0,T])} \\ &\quad + \|\nabla \cdot \tilde{u}_N\|^4_{L^4([0,T];L^2(\Omega))} + \lambda\|\tilde{u}_N\|^4_{L^4([0,T];H^1(\Omega))} \leq F(G, u_0)T\varepsilon^2.\end{aligned}$$

Proof of Theorem 6 Considering

$$\begin{aligned}&\|\tilde{u}_{N_1} - \tilde{u}_{N_2}\|_{L^4([0,T];L^2(\Omega))} \\ &\quad = \|\tilde{u}_{N_1} - u_1 + u_2 - \tilde{u}_{N_2} + u_1 - u_2\|_{L^4([0,T];L^2(\Omega))} \\ &\quad \leq \|\tilde{u}_{N_1} - u_1\|_{L^4([0,T];L^2(\Omega))} + \|u_2 - \tilde{u}_{N_2}\|_{L^4([0,T];L^2(\Omega))} \\ &\qquad + \|u_1 - u_2\|_{L^4([0,T];L^2(\Omega))}.\end{aligned}$$

From (3.12)

$$\|\tilde{u}_{N_1} - u_1\|_{L^4([0,T];L^2(\Omega))} \leq \left((ce^{F(G,u_{0,1})T} + C_{Tr})\varepsilon^2 + ce^{F(G,u_{0,1})T}\frac{\varepsilon^4}{\lambda}\right)^{1/4},$$

and

$$\|u_2 - \tilde{u}_{N_2}\|_{L^4([0,T];L^2(\Omega))} \leq \left((ce^{F(G,u_{0,2})T} + C_{Tr})\varepsilon^2 + ce^{F(G,u_{0,2})T}\frac{\varepsilon^4}{\lambda}\right)^{1/4}.$$

From the stability of solutions of NSE, we have $\|u_1 - u_2\|_{L^4([0,T];L^2(\Omega))} \leq c(\|u_{0,1} - u_{0,2}\|_{L^2(\Omega)} + \|f_1 - f_2\|_{L^4([0,T];L^2(\Omega))})$.

Therefore, we have

$$\|u_{N_1} - u_{N_2}\|_{L^4([0,T];L^2(\Omega))} \leq C_1 + C_2 + C_3,$$

where $C_1 = \left((ce^{F(G,u_{0,1})T} + C_{Tr})\varepsilon^2 + ce^{F(G,u_{0,1})T}\frac{\varepsilon^4}{\lambda}\right)^{1/4}$, $C_2 = \left((ce^{F(G,u_{0,2})T} + C_{Tr})\varepsilon^2 + ce^{F(G,u_{0,2})T}\frac{\varepsilon^4}{\lambda}\right)^{1/4}$, $C_3 = c(\|u_{0,1} - u_{0,2}\|_{L^2(\Omega)} + \|f_1 - f_2\|_{L^4([0,T];L^2(\Omega))})$.

This implies that our scheme is approximately stable.

Acknowledgements We thank the referee and the editorial board members for comments and suggestions. J. Tian's work is supported in part by the AMS Simons Travel Grant and the Jess and Mildred Fisher Endowed Professor fund of Mathematics from the Fisher College of Science and Mathematics at Towson University.

References

1. Barron, A.: Universal approximation bounds for superpositions of a sigmoidal function. IEEE Trans. Inf. Theory **39**, 930–945 (1993)
2. Beck, C., Jentzen, A.: Machine learning approximation algorithms for high-dimensional fully nonlinear partial differential equations and second-order backward stochastic differential equations. J. Nonlinear Sci. **29**, 1563–1619 (2019)
3. Berner, J., Grohs, P., Jentzen, A.: Analysis of the generalization error: empirical risk minimization over deep artificial neural networks overcomes the curse of dimensionality in the numerical approximation of black-scholes partial differential equations. SIAM J. Math. Data Sci. **2**, 631–657 (2020)
4. Cybenko, G.: Approximation by superpositions of a sigmoidal function. Math. Control Signals Syst. **2**, 303–314 (1989)
5. Constantin, P., Foias, C.: Navier-Stokes Equations, Chicago Lectures in Mathematics. The University of Chicago Press, Chicago (1988)
6. Evans, L.C.: Partial Differential Equations. Graduate Studies in Mathematics, AMS (1998)
7. Fang, R., Sondak,D., Protopapas, P., Succi, S.: Neural network models for the anisotropic Reynolds stress tensor in turbulent channel flow. J. Turbulence, **21**, 525–543 (2020)
8. Farwig, Reinhard, Galdi, Giovanni P., Sohr, Hermann: A new class of weak solutions of the Navier–Stokes equations with nonhomogeneous data. J. Math. Fluid Mech. **8**, 423–444 (2006)
9. Foias, C., Jolly, M., Manley, O., Rosa, R.: Statistical estimates for the Navier–Stokes equations and the Kraichnan theory of 2-D fully developed turbulence. J. Stat. Phys. **108**, 591–645 (2002)
10. Fursikov, A.V., Gunzburger, M.D., Hou, L.S.: Inhomogeneous boundary value problems for the three-dimensional evolutionary Navier–Stokes equations. J. Math. Fluid Mech. **4**, 45–75 (2002)
11. Goodfellow, I., Bengio, Y., Courville, A., Bengio, Y.: Deep Learning, vol. 1. MIT Press, Cambridge (2016)
12. Gulian, M., Raissi, M., Perdikaris, P., Karniadakis, G.E.: Machine learning of space-fractional differential equations. SIAM J. Sci. Comput. **41**, A2485–A2509 (2019)
13. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778 (2016)
14. Hornik, K., Stichcombe, M., Halbert, W.: Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. Neural Netw. **3**, 551–560 (1990)
15. Hornik, K.: Approximation capabilities of Muitilayer feedforward networks. Neural Netw. **4**, 251–257 (1991)
16. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**, 1735–1780 (1997)
17. Jia, X., Zhao, L., Zhang, L., He, J., Xu, J.: Modified Regularized Dual Averaging Method for Training Sparse Convolutional Neural Networks, arXiv preprint arXiv:1807.04222, 1097–1105 (2018)
18. Jiménez, J.: Machine-aided turbulence theory, J. Fluid Mech., **854** (2018)
19. Krizhevsky, A., Sutskever, I., Hinton, G. E.: Imagenet classification with deep convolutional neural networks, Adv. Neural Inf. Process. Syst. **25**, 1097–1105 (2012)
20. Lions, J., Magenes, E.: Non-Homogeneous Boundary Value Problems and Applications. Springer Science & Business Media, Berlin (2012)
21. Mehta, P.P., Pang, G., Song, F., Karniadakis, G.E.: Discovering a universal variable-order fractional model for turbulent Couette flow using a physics-informed neural network **22**, 1675–1688 (2019)

22. Mao, Z., Ameya, D.J., Karniadakis, G.E.: Physics-informed neural networks for high-speed flows. Comput. Methods Appl. Mech. Eng. **360**, 112789 (2020)
23. Mao, Z., Zhen, L., Karniadakis, G.E.: Nonlocal flocking dynamics: Learning the fractional order of pdes from particle simulations. Commun. Appl. Math. Comput. **1**, 597–619 (2019)
24. Nguyen-Thien, T., Tran-Cong, T.: Approximation of functions and their derivatives: A neural network implementation with applications. Appl. Math. Model. **23**, 687–704 (1999)
25. Pang, G., Lu, L., Karniadakis, G.E.: fPINNs: Fractional physics-informed neural networks. SIAM J. Sci. Comput. **41**, A2603–A2626 (2019)
26. Raissi, M., Wang, Z., Triantafyllou, M., Karniadakis, G.E.: Deep learning of vortex-induced vibrations. J. Fluid Mech. **861**, 119–137 (2019)
27. Raissi, M., Perdikaris, P., Karniadakis, G.E.: Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J. Comput. Phys. **378**, 386–707 (2019)
28. Raissi, M., Karniadakis, G.E.: Hidden physics models: machine learning of nonlinear partial differential equations. J. Comput. Phys. **357**, 125–141 (2018)
29. Raymond, J.-P.: Stokes and Navier–Stokes equations with nonhomogeneous boundary conditions. Ann. I. H. Poincaré AN **24**, 921–951 (2007)
30. Sirignano, J., Spiliopoulos, K.: DGM: A deep learning algorithm for solving partial differential equations. J. Comput. Phys. **375**, 1339–1364 (2018)
31. Temam, R.: Navier–Stokes equations and nonlinear functional analysis, Soc. Indus. Appl. Math., **66** (1995)
32. Temam, R.: Navier–Stokes equations: theory and numerical analysis, Am. Math. Soc., **343** (2001)
33. Thuerey, N., Wei, K., Prantl, L., Hu, X.: Deep learning methods for reynolds-averaged Navier–Stokes simulations of airfoil flows. AIAA J. **58**, 25–36 (2020)
34. Tracey, B., Duraisamy, K., Alonso, J.J.: A machine learning strategy to assist turbulence model development. AIAA Paper **1287**, 1287 (2015)
35. Wang, J.X., Wu, J.L., Xiao, H.: Physics-informed machine learning approach for reconstructing Reynolds stress modeling discrepancies based on dns data. Phys. Rev. Fluids **2**, 034603 (2017)
36. Wu, J.L., Xiao, H., Paterson, E.: Physics-informed machine learning approach for augmenting turbulence models: A comprehensive framework. Phys. Rev. Fluids **3**, 074602 (2018)
37. Xie, T., Cao, F.: The errors of simultaneous approximation of multivariate functions by neural networks. Comput. Math. Appl. **61**, 3146–3152 (2011)
38. Yang, L., Zhang, D., Karniadakis, G.E.: Physics-informed generative adversarial networks for stochastic differential equations. SIAM J. Sci. Comput. **42**, A292–A317 (2020)