ERROR ESTIMATES FOR DEEP LEARNING METHODS IN FLUID DYNAMICS

ANIMIKH BISWAS, JING TIAN, AND SULEYMAN ULUSOY

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 (DNN) approximation of the solution, with a fixed complexity.

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 [9, 11, 17]. Recently, its use has proliferated in computational sciences and physical modeling such as the modeling of turbulence [7, 26, 27, 16, 28, 29]. Moreover, machine learning methods (which are *mesh-free*) have gained wide applicability in obtaining numerical solutions of various types of partial differential equations (PDEs); see [2, 3, 10, 20, 21, 22, 23, 20] 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 stepsize. 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].

Many infinite dimensional dynamical systems of practical interest arise in the context of geophysical flows related to the atmosphere and ocean. The Navier-Stokes and Euler 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. Especially, the Navier-Stokes equations can be used to model the Incompressible fluid flow and have been employed in describing many phenomena in science and engineering applications. For example, they are used in modeling the water flow in a pipe, air flow around a wing, ocean

Date: August 10, 2020.
Key words and phrases. partial differential equations, machine learning, deep learning, neural network, convergence, error estimate.

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. In this study, we mainly focus on using the neural networks techniques to solve two dimensional Navier-Stokes equations. However, we consider the elliptic case first to illustrate the fundamental issues involved.

In literature, most studies [10, 20, 21, 22] focus on the numerical efficacy study of designing the neural network algorithms. Concrete and complete mathematical analysis are 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 [23]. The goal of this paper is to provide a mathematically rigorous error analysis of deep learning methods employed in [10, 20, 21, 22] for the general elliptic and two-dimensional Navier-Stokes equations. Our goal in this paper is not to analyze all the details of different possibilities of neural network architecture. Instead, we would like to provide a mathematically rigorous analysis of the method, with *error estimates, and stability analysis*, similar in spirit to the probabilistic error analysis for machine learning algorithms for the Black-Scholes equations in [3]. We consider two different settings: the elliptic PDEs, mainly to fix ideas and illustrate our approach, and the Navier-Stokes equations, which is the main focus of this work. Although our results are proven in the context of the two-dimensional Navier-Stokes equations, we note that 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 computational algorithm employed in machine learning of PDEs (for instance in [10, 20, 21, 22]) 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. Nevertheless, we side step this issue by obtaining an explicit error estimate in terms of the attained value of the loss function (which takes the value zero for the true solution). 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.

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 section 4, we present the mathematical analysis of the neural network algorithm in the elliptic system. This also serves as a systematic introduction of our analysis. In section 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 $\sigma(wx + b)$, $x \in \mathbb{R}^d, w = (w_1, \cdots, w_d)$, $b \in \mathbb{R}$. 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:

$$\begin{aligned}
y_2(x) &= \sigma(W_1 x + b_1),\\
y_3(y_2) &= \sigma(W_2 y_2 + b_2),\\
&\cdot\\
&\cdot\\
&\cdot\\
y_{n_l}(y_{n_l-1}) &= \sigma(W_{n_l-1} y_{n_l-1} + b_{n_l-1}),\\
y_{n_l+1}(y_{n_l}) &= \sigma(W_{n_l} y_{n_l} + b_{n_l}),\\
f_\theta &= y_{n_l+1}(y_{n_l}(\cdots(y_2(x)))),
\end{aligned} \tag{2.2}$$

where θ ensembles all the weights and parameters.

$$\theta = \{W_1, W_2, \cdots, W_{n_l}, b_1, \cdots, b_{n_l}\}. \tag{2.3}$$

In practice, different neural network architectures are possible such as those involving recurrent cells [14], convolutional layers [17], sparse convolutional neural networks [15], pooling layers, residual connections [11].

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 [13]; see [19, 30] 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 approxmation 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 [12], 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. A relevant theorem from [30], relating the accuracy of the approximation of a DNN with the complexity of the DNN and the regularity of the function being approximated, is given below.

Theorem 2.1. *Suppose that $\sigma \in C^\infty(\mathbb{R})$, $\sigma^{(v)}(0) \neq 0$ for $\nu = 0, 1, ...,$ 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 with complexity $n \in \mathbb{N}$ exists such that*

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

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

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

3. Main results

Let $\mathcal{F}_N$ be a DNN with complexity N, which is a finite dimensional function space on a bounded domain. 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

$$M := \|u\|_{H^2(\Omega)} \le c\|f\|_{L^2(\Omega)}.$$

Consequently, the minimization problem

$$\inf_{u\in \text{ appropriate Sobolev class}} \left\{\|(\mathcal{L}u)(x) - f(x)\|^2_{L^2(\Omega)} + \|u|_{\partial\Omega}\|^2_{L^2(\partial\Omega)}\right\} \tag{3.2}$$

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

$$L = \alpha^2\|\mathcal{L}u_N - f\|^2_{L^2(\Omega)} + \beta^2\|u_N|_{\partial\Omega}\|^2_{L^2(\partial\Omega)}, u_N \in \mathcal{F}_N, \tag{3.3}$$

under the restriction that $\|u_N\|_{H^2(\Omega)} \le \widetilde{M}$ (i.e. $\|u_N\|_{H^2(\Omega)}$ is bounded) for suitable $\widetilde{M}$ (e.g. $\widetilde{M} = 2M$), 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 [21]. 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

$$\inf_{u\in\mathcal{F}_N, \|u_N\|_{H^2(\Omega)}\le\widetilde{M}} \left\{\sum_{j=1}^m \alpha^2 |(\mathcal{L}u)(x_j) - f(x_j)|^2 + \sum_{j=1}^n \beta^2 |u(y_j) - g(y_j)|^2\right\}. \tag{3.4}$$

Note that (3.4) 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

$$\inf_{u\in\mathcal{F}_N, \|u_N\|_{H^2(\Omega)}\le\widetilde{M}} \left\{\|(\mathcal{L}u)(x) - f(x)\|^2_{L^2(\Omega)} + \|u|_{\partial\Omega} - g\|^2_{L^2(\partial\Omega)}\right\}. \tag{3.5}$$

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 3.1. We can also use an unrestricted optimization in (3.4) or (3.5). 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 actual solution are close to each other, we can control the loss function.

Theorem 3.1. *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*

$$\alpha^2 \|\mathcal{L}u_N - f\|^2_{L^2(\Omega)} + \beta^2 \|u_N|_{\partial\Omega}\|^2_{L^2(\partial\Omega)} \leq c\varepsilon^2 \text{ and } \|u_N\|_{H^2(\Omega)} \leq \widetilde{M},$$

where $\widetilde{M}$ can be taken to be $2M = 2\|u\|_{H^2(\Omega)}$.

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.

Theorem 3.2. *Let u be a solution of* (3.1) *and $\alpha, \beta, \epsilon > 0$. Assume that $u_N \in \mathcal{F}_N$ is such that*

$$\alpha^2 \|\mathcal{L}u_N - f\|^2_{L^2(\Omega)} + \beta^2 \|u_N|_{\partial\Omega}\|^2_{L^2(\partial\Omega)} \leq \varepsilon^2, \tag{3.6}$$

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

$$\|u - u_N\|_{H^1(\Omega)} \leq O((M + \widetilde{M})^{1/2} \varepsilon^{1/2}),$$

and

$$\|u - u_N\|_{L^2(\Omega)} \leq O((M + \widetilde{M})^{1/3} \varepsilon^{2/3}).$$

Observe that in the Theorem above, as expected, the error estimate in the L^2-norm is stronger than the error estimate in the H^1-norm. We show in the theorem below how the error estimate in the L^2-norm can be improved further by altering the loss function.

Theorem 3.3. *Let u be a solution of* (3.1) *and let $u_N \in \mathcal{F}_N$ be such that*

$$\alpha^2 \|\mathcal{L}u_N - f\|^2_{L^2(\Omega)} + \beta^2 \|u_N|_{\partial\Omega}\|^2_{H^{1/2}(\partial\Omega)} \leq \varepsilon^2, \tag{3.7}$$

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

$$\|u - u_N\|_{L^2(\Omega)} < O(\varepsilon).$$

3.2. **Incompressible Navier-Stokes Equations.** The incompressible Navier-Stokes equations (NSE) are given by

$$\begin{aligned} \partial_t u - \Delta u + u \cdot \nabla u + \nabla p &= f, \\ \nabla \cdot u &= 0, \\ u|_{\partial\Omega} &= 0, \\ u(x,0) &= u_0(x), x \in \Omega. \end{aligned} \tag{3.8}$$

In (3.8), u denotes the velocity of the fluid and p the pressure. 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 3.4. *Assume that u is a strong solution of the 2D NSE (3.8) and $\tilde{u}_N \in \mathcal{F}_N$ such that*

$$\begin{aligned} &\|\tilde{u}_N|_{\partial\Omega}\|^4_{L^4([0,T];H^{1/2}(\partial\Omega))} + \|\tilde{u}_N(x,0) - u_0(x)\|^2_{L^2(\Omega)} \\ &+ \|\partial_t \tilde{u}_N - \Delta\tilde{u}_N + \tilde{u}_N \cdot \nabla\tilde{u}_N + \nabla\tilde{p}_N - f\|^2_{L^2(\Omega\times[0,T])} \\ &+ \|\nabla \cdot \tilde{u}_N\|^4_{L^4([0,T];L^2(\Omega))} + \lambda\|\tilde{u}_N\|^4_{L^4([0,T];H^1(\Omega))} \le \varepsilon^2. \end{aligned} \tag{3.9}$$

Then

$$\|u - \tilde{u}_N\|_{L^4([0,T];L^2(\Omega))} \le O\left(\varepsilon^{1/2} + \frac{\varepsilon}{\lambda^{1/4}}\right). \tag{3.10}$$

The reverse direction of Theorem 3.4 has also been proved.

Theorem 3.5. *Given any $\varepsilon > 0$, we can find $\tilde{u}_N \in \mathcal{F}_N$, such that*

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

Furthermore, we prove that our scheme is approximately stable.

Theorem 3.6. *Assume $\tilde{u}_{N_1} \in \mathcal{F}_{N_1}$ is the approximate solution of*

$$\begin{aligned} \frac{\partial}{\partial t} u_1 - \Delta u_1 + u_1 \cdot \nabla u_1 + \nabla p_1 &= f_1, \\ \nabla \cdot u_1 &= 0, \\ u_1|_{\partial\Omega} &= 0, \\ u_1(x,0) &= u_{0,1}(x). \end{aligned} \tag{3.12}$$

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

$$\begin{aligned} \frac{\partial}{\partial t} u_2 - \Delta u_2 + u_2 \cdot \nabla u_2 + \nabla p_2 &= f_2, \\ \nabla \cdot u_2 &= 0, \\ u_2|_{\partial\Omega} &= 0, \\ u_2(x,0) &= u_{0,2}(x). \end{aligned} \tag{3.13}$$

Here, $\tilde{u}_{N_1}$ and $\tilde{u}_{N_2}$ satisfy (3.9) with corresponding f_1 and f_2. Then, we have

$$\begin{aligned} &\|\tilde{u}_{N_1} - \tilde{u}_{N_2}\|_{L^4([0,T];L^2(\Omega))} \le \\ &O\left(\varepsilon^{1/2} + \frac{\varepsilon}{\lambda^{1/4}} + \|u_{0,1} - u_{0,2}\|_{L^2(\Omega)} + \|f_1 - f_2\|_{L^4([0,T];L^2(\Omega))}\right). \end{aligned}$$

4. Elliptic Equations: Proofs of Main Theorems

Proof of Theorem 3.1:

Proof. We remark first that given any $\epsilon > 0$, by Theorem 2.1, there exists a DNN $\mathcal{F}_N$ of complexity N and $u_N \in \mathcal{F}_N$ such that $\|u - u_N\|_{H^2(\Omega)} \le \epsilon$.

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

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

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

We also have

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

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

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

Combining (4.1) and (4.2), we have

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

Finally,

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

$\square$

Let us consider the converse of Theorem 3.1. 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.

Lemma 4.1. *If* $\|u_N\|_{H^2(\Omega)} \le \widetilde{M}$, $\|\nabla(u - u_N)\|^2_{L^2(\Omega)} \le c\varepsilon(M + \widetilde{M})\left(\frac{1}{\alpha} + \frac{1}{\beta}\right)$, *and* $\|(u - u_N)|_{\partial\Omega}\|^2_{L^2(\partial\Omega)} < \frac{\varepsilon^2}{\beta^2}$, *then*

$$\|u - u_N\|^2_{L^2(\Omega)} < O((M + \widetilde{M})\varepsilon). \tag{4.3}$$

Proof. Since

$$\begin{aligned}\|u - u_N\|_{L^2(\Omega)} &\le \|u\|_{L^2(\Omega)} + \|u_N\|_{L^2(\Omega)}\\ &\le \|u\|_{L^2(\Omega)} + \|u_N\|_{H^2(\Omega)},\end{aligned}$$

and

$$\|u\|_{L^2(\Omega)} \le \|u\|_{H^2(\Omega)} \le M;\ \|u_N\|_{H^2(\Omega)} \le \widetilde{M},$$

we have

$$\|u - u_N\|_{L^2(\Omega)} \le M + \widetilde{M} \text{ and } \|u - u_N\|_{H^2(\Omega)} \le M + \widetilde{M}. \tag{4.4}$$

Consider

$$\begin{aligned}\|(u-u_N)|_{\partial\Omega}\|_{H^{1/2}(\partial\Omega)} &\leq \|(u-u_N)|_{\partial\Omega}\|^{2/3}_{L^2(\partial\Omega)}\|(u-u_N)|_{\partial\Omega}\|^{1/3}_{H^{3/2}(\partial\Omega)} \\ &\leq c\frac{\varepsilon^{2/3}}{\beta^{2/3}}\|(u-u_N)\|^{1/3}_{H^2(\Omega)} \\ &\leq c(M+\widetilde{M})^{1/3}\frac{\varepsilon^{2/3}}{\beta^{2/3}}.\end{aligned} \tag{4.5}$$

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.

$$\|l_\Omega(u-u_N)\|_{H^1(\Omega)} \leq C_{l_\Omega}\|(u-u_N)|_{\partial\Omega}\|_{H^{1/2}(\partial\Omega)} \leq c(M+\widetilde{M})^{1/3}\frac{\varepsilon^{2/3}}{\beta^{2/3}},$$

where C_{l_Ω} is the constant from the lifting operator. Moreover

$$\begin{aligned}\|(u-u_N)-l_\Omega(u-u_N)\|_{L^2(\Omega)} &\leq c\|\nabla((u-u_N)-l_\Omega(u-u_N))\|_{L^2(\Omega)} \\ &\leq c\|\nabla(u-u_N)\|_{L^2(\Omega)} + c\|\nabla(l_\Omega(u-u_N))\|_{L^2(\Omega)} \\ &\leq c\varepsilon^{1/2}(M+\widetilde{M})^{1/2}\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)^{1/2} + c\|l_\Omega(u-u_N)\|_{H^1(\Omega)} \\ &\leq c\varepsilon^{1/2}(M+\widetilde{M})^{1/2}\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)^{1/2} + c(M+\widetilde{M})^{1/3}\frac{\varepsilon^{2/3}}{\beta^{2/3}}.\end{aligned}$$

Therefore

$$\begin{aligned}\|u-u_N\|_{L^2(\Omega)} &\leq \|(u-u_N)-l_\Omega(u-u_N)+l_\Omega(u-u_N)\|_{L^2(\Omega)} \\ &\leq \|(u-u_N)-l_\Omega(u-u_N)\|_{L^2(\Omega)} + \|l_\Omega(u-u_N)\|_{L^2(\Omega)} \\ &\leq c\varepsilon^{1/2}(M+\widetilde{M})^{1/2}\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)^{1/2} + c(M+\widetilde{M})^{1/3}\frac{\varepsilon^{2/3}}{\beta^{2/3}} \\ &= O((M+\widetilde{M})^{1/2}\varepsilon^{1/2}).\end{aligned}$$

□

Proof of Theorem 3.2:

Proof. First, we have

$$\|(u-u_N)|_{\partial\Omega}\|^2_{L^2(\partial\Omega)} = \|u_N|_{\partial\Omega}\|^2_{L^2(\partial\Omega)} \leq \frac{\varepsilon^2}{\beta^2}. \tag{4.6}$$

Moreover

$$\begin{aligned}&\|\nabla(u-u_N)\|^2_{L^2(\Omega)} \\ &= \int_\Omega (\nabla(u-u_N))^2 dx \\ &= \int_{\partial\Omega} \nabla(u-u_N)|_{\partial\Omega}\cdot(u-u_N)|_{\partial\Omega}dx - \int_\Omega (u-u_N)\cdot(\Delta(u-u_N))dx \\ &\leq \|\nabla(u-u_N)|_{\partial\Omega}\|_{L^2(\partial\Omega)}\cdot\|(u-u_N)|_{\partial\Omega}\|_{L^2(\partial\Omega)} + \|u-u_N\|_{L^2(\Omega)}\cdot\|\Delta(u-u_N)\|_{L^2(\Omega)}.\end{aligned} \tag{4.7}$$

Since

$$\|\Delta(u-u_N)\|_{L^2(\Omega)} \leq c\|\mathcal{L}u_N - f\|_{L^2(\Omega)} \leq c\frac{\varepsilon}{\alpha},$$

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

and

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

we have

$$\|u - u_N\|_{L^2(\Omega)} \leq M + \widetilde{M}.$$

Moreover

$$\|(u - u_N)|_{\partial\Omega}\|_{L^2(\partial\Omega)} \leq \frac{\varepsilon}{\beta},$$

and

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

Therefore, we have

$$\|\nabla(u - u_N)\|^2_{L^2(\Omega)} \leq c\varepsilon(M + \widetilde{M}) \left(\frac{1}{\alpha} + \frac{1}{\beta}\right). \tag{4.8}$$

Therefore, we can apply Lemma 4.1 and obtain (4.3). Combining (4.8) and (4.3), we have

$$\|u - u_N\|_{H^1(\Omega)} = \|u - u_N\|_{L^2(\Omega)} + \|\nabla(u - u_N)\|_{L^2(\Omega)} = O((M + \widetilde{M})^{1/2}\varepsilon^{1/2}).$$

We can improve the rate of $\|u - u_N\|_{L^2(\Omega)}$ by using a different approach: Denote $\mathcal{L}u_N = f_\varepsilon$ and $u_N|_{\partial\Omega} = Tr(u_N) = g_\varepsilon$. From (3.6), we have $\|f - f_\varepsilon\|_{L^2(\Omega)} \leq \frac{\varepsilon}{\alpha}$, from (4.5), we have $\|g_\varepsilon\|_{H^{1/2}(\partial\Omega)} < c(M + \widetilde{M})^{1/3}\frac{\varepsilon^{2/3}}{\beta^{2/3}}$. Let $\tilde{u}_N = u_N - l_\Omega g_\varepsilon$. Then

$$\begin{aligned} \mathcal{L}\tilde{u}_N &= f_\varepsilon - \mathcal{L}l_\Omega g_\varepsilon, \\ \tilde{u}_N|_{\partial\Omega} &= 0. \end{aligned} \tag{4.9}$$

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

$$\|u - \tilde{u}_N\|_{H^1(\Omega)} \leq C_{\mathcal{L}^{-1}}\|(f - f_\varepsilon) + \mathcal{L}l_\Omega g_\varepsilon\|_{H^{-1}(\Omega)}. \tag{4.10}$$

Therefore, we have

$$\begin{aligned} &\|u - u_N\|_{H^1(\Omega)} \\ &= \|u - (u_N - l_\Omega g_\varepsilon) - l_\Omega g_\varepsilon\|_{H^1(\Omega)} \\ &\leq \|u - \tilde{u}_N\|_{H^1(\Omega)} + \|l_\Omega g_\varepsilon\|_{H^1(\Omega)} \\ &\leq C_{\mathcal{L}^{-1}}\|(f - f_\varepsilon) + \mathcal{L}l_\Omega g_\varepsilon)\|_{H^{-1}(\Omega)} + \|l_\Omega\|\|g_\varepsilon\|_{H^{1/2}(\partial\Omega)}. \end{aligned} \tag{4.11}$$

Since

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

and

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

Thus

$$\begin{aligned} \|u - u_N\|_{L^2(\Omega)} &\leq \|u - u_N\|_{H^1(\Omega)} \\ &\leq c\|f - f_\varepsilon\|_{L^2(\Omega)} + c\|g_\varepsilon\|_{H^{1/2}(\partial\Omega)}. \end{aligned} \tag{4.12}$$

Therefore, we have

$$\|u - u_N\|_{L^2(\Omega)} \leq \frac{c}{\alpha}\varepsilon + \frac{c(M + \widetilde{M})^{1/3}}{\beta^{2/3}}\varepsilon^{2/3} =: O((M + \widetilde{M})^{1/3}\varepsilon^{2/3}). \tag{4.13}$$

□

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 $\|u - u_N\|_{L^2(\Omega)}$ which is stated in Theorem 3.3.

Proof of Theorem 3.3:

Proof. Same as before, $\mathcal{L}u_N = f_\varepsilon$ and $u_N|_{\partial\Omega} = Tr(u_N) = g_\varepsilon$. From (3.7), we have $\|f - f_\varepsilon\|_{L^2(\Omega)} \leq \dfrac{\varepsilon}{\alpha}$, and $\|g_\varepsilon\|_{H^{1/2}(\partial\Omega)} \leq \dfrac{\varepsilon}{\beta}$. Let $\tilde{u}_N = u_N - l_\Omega g_\varepsilon$.

$$\begin{aligned} \mathcal{L}\tilde{u}_N &= f_\varepsilon - \mathcal{L}l_\Omega g_\varepsilon, \\ \tilde{u}_N|_{\partial\Omega} &= 0. \end{aligned} \tag{4.14}$$

Similar to the proof of Theorem 3.2, we have

$$\|u - u_N\|_{H^1(\Omega)} \leq C_{\mathcal{L}^{-1}}\|(f - f_\varepsilon) + \mathcal{L}l_\Omega g_\varepsilon)\|_{H^{-1}(\Omega)} + \|l_\Omega\|\|g_\varepsilon\|_{H^{1/2}(\partial\Omega)},$$

and

$$\begin{aligned} \|u - u_N\|_{L^2(\Omega)} &\leq \|u - u_N\|_{H^1(\Omega)} \\ &\leq c\|f - f_\varepsilon\|_{L^2(\Omega)} + c\|g_\varepsilon\|_{H^{1/2}(\partial\Omega)}. \end{aligned}$$

Therefore, we have

$$\|u - u_N\|_{L^2(\Omega)} \leq \frac{c}{\alpha}\varepsilon + \frac{c}{\beta}\varepsilon = O(\varepsilon).$$

□

5. Navier-Stokes Equations: Proof of Main Theorems

5.1. **Functional analytic framework.** Let Ω be a bounded domain in $\mathbb{R}^2$ and

$$H := \left\{u \in L^2(\Omega),\ \nabla \cdot u \in L^2(\Omega),\ \nabla \cdot u = 0\right\},$$

$$V := \left\{u \in H^1_0(\Omega),\ \nabla \cdot u = 0\right\}.$$

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 (3.8), the functional form of the NSE can be written as

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

We recall the definition of strong solutions from [24]:
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 (3.8) as in [5, 24], 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 5.1. *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, 25] for more details).

5.3. **Proofs.** Consider an approximate solution $\tilde{u}_N \in \mathcal{F}_N$, i.e. $\tilde{u}_N$ satisfies (3.9) 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 5.2. *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

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

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

Then

$$\sup_{[0,T]} \|u(x,t) - u_N(x,t)\|_{L^2(\Omega)}^2 \le O\left(\varepsilon + \frac{\varepsilon^2}{\sqrt{\lambda}}\right).$$

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

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

Since

$$|(B(w,u),w)| \le c\|A^{1/2}u\|_{L^2}\|w\|_{L^2}\|A^{1/2}w\|_{L^2} \le c\frac{\|A^{1/2}u\|_{L^2}^2\|w\|_{L^2}^2}{2} + \frac{\|A^{1/2}w\|_{L^2}^2}{2},$$

and

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

we have

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

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

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

where, G is the Grashof number (defined as $G = \|f\|$ [8]) and $F(G)$ is a function of G, we have

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

Moreover

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

and

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

We have

$$\begin{aligned} \|w(t)\|_{L^2}^2 &\le e^{cF(G)t}\|w(0)\|_{L^2}^2 + e^{cF(G)t}\int_0^t ||\varphi||_{V'}^2 ds \\ &\le O\left(\varepsilon + \frac{\varepsilon^2}{\sqrt{\lambda}}\right) e^{cF(G)t}. \end{aligned}$$

Therefore

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

□

Lemma 5.3. *Assume*

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

and with the Hodge decomposition (5.3), we have

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

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 & Magenes [18], 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 (5.11) and (5.8), 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 (5.10) and (5.8),

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

Therefore,

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

□

Note that (5.9) also implies

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

Proof of Theorem 3.4

Proof. Applying the Leray projection operator $\mathbb{P}$ to (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

$$\int_0^T \|\mathbb{P}(v_N \cdot \nabla\tilde{u}_N)\|_{V'}^2 \, dt \le c \int_0^T \|\nabla\tilde{u}_N\|_{L^2(\Omega)}^2 \|v_N\|_{L^2(\Omega)} \|\nabla v_N\|_{L^2(\Omega)} \, dt$$

$$\le c \left(\int_0^T \|\nabla\tilde{u}_N\|_{L^2(\Omega)}^4 dt\right)^{1/2} \left(\int_0^T \|v_N\|_{L^2(\Omega)}^4 dt\right)^{1/4} \left(\int_0^T \|\nabla v_N\|_{L^2(\Omega)}^4 dt\right)^{1/4}.$$

Since

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

we have

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

Therefore

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

Applying Lemma 5.3, we have

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

Therefore

$$\int_0^T \|\mathbb{P}(v_N\cdot\nabla\tilde u_N)\|_{V'}^2\,dt \le \frac{c\varepsilon^2}{\sqrt{\lambda}}. \tag{5.17}$$

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

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

Therefore

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

Since

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

and

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

Therefore

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

Combining (5.17) and (5.18), we have

$$\int_0^T \|\psi\|_{V'}^2\,dt \le O\left(\frac{\varepsilon^2}{\sqrt{\lambda}}\right). \tag{5.19}$$

Moreover, since

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

we have

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

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

$$\int_0^T \|\mathbb{P}(\Delta v_N)\|_{V'}^2\,dt \le c\varepsilon,\ \int_0^T \|\mathbb{P}\tilde f\|_{V'}^2\,dt \le \int_0^T \|\mathbb{P}\tilde f\|_{L^2}^2\,dt \le \varepsilon^2,\ \int_0^T \|\psi\|_{V'}^2\,dt \le O\left(\frac{\varepsilon^2}{\sqrt{\lambda}}\right),$$

we obtain

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

Since $\|u_N(x,0)-u_0(x)\|_{L^2}^2 \le \|\tilde{u}_N(x,0)-u_0(x)\|_{L^2}^2 \le \varepsilon^2$. Applying Lemma 5.2, we have

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

Moreover, since

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

Therefore

$$\begin{aligned}\int_0^T \|u-\tilde{u}_N\|_{L^2(\Omega)}^4 dt &\le \int_0^T \|u-u_N\|_{L^2(\Omega)}^4 dt + \int_0^T \|v_N\|_{L^2(\Omega)}^4 dt \\ &\le O\left(\varepsilon^2+\frac{\varepsilon^4}{\lambda}\right)+O\left(\varepsilon^2\right)=O\left(\varepsilon^2+\frac{\varepsilon^4}{\lambda}\right).\end{aligned}$$

So

$$\left(\int_0^T \|u-\tilde{u}_N\|_{L^2(\Omega)}^4 dt\right)^{1/4} \le O\left(\varepsilon^{1/2}+\frac{\varepsilon}{\lambda^{1/4}}\right).$$

□

Lemma 5.4. *Given $\varepsilon > 0$, assume that (u,p) satisfies (3.8), then, there exists $(\tilde{u}_N,\tilde{p}_N)\in\mathfrak{F}_N$ satisfying*

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

Proof. From Theorem 2.1, 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 3.5:

Proof. From Lemma 5.4, given $\varepsilon>0$, assume that u is a strong solution of (3.8) and there is an N such that $\tilde{u}_N\in\mathfrak{F}_N$ satisfying $\sup_{t\in[0,T]}\|u-\tilde{u}_N\|_{L^2(\Omega)}\le\varepsilon$, $\|u-\tilde{u}_N\|_{H^{1,2}(\Omega\times[0,T])}\le\varepsilon$, $\left(\int_0^T \|u-\tilde{u}_N\|_{W^{1,4}(\Omega)}^4 dt\right)^{1/4}\le\varepsilon$, $\|p-\tilde{p}_N\|_{L^2([0,T];H^1(\Omega))}\le\varepsilon$.
Now, we estimate the left hand side of (3.11) term by term:

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

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

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

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

Consider

$$\begin{aligned}\lambda\|\tilde{u}_N\|^4_{L^4([0,T];H^1(\Omega))} &\le c\lambda\|\tilde{u}_N-u\|^4_{L^4([0,T];H^1(\Omega))}+c\lambda\|u\|^4_{L^4([0,T];H^1(\Omega))}\\ &\le c\lambda\varepsilon^4+c\lambda F(G).\end{aligned}$$

Picking λ small enough, we have

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

Consider

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

We have

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

and

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

Moreover

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

We have

$$\|\nabla\tilde{u}_N\|_{L^4(\Omega\times[0,T])}\le\|\nabla\tilde{u}_N-\nabla u\|_{L^4(\Omega\times[0,T])}+\|\nabla u\|_{L^4(\Omega\times[0,T])}\le c\varepsilon+cF(G)\le cF(G).$$

Moreover,

$$\|\tilde{u}_N-u\|_{L^4(\Omega\times[0,T])}\le c\varepsilon,\ \|u\|_{L^4(\Omega\times[0,T])}\le cF(G),\ \|\nabla\tilde{u}_N-\nabla u\|_{L^4(\Omega\times[0,T])}\le c\varepsilon.$$

So

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

Therefore

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

Moreover

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

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

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

□

Proof of Theorem 3.6:

Proof. Considering

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

From (3.10), $\|\tilde{u}_{N_1} - u_1\|_{L^4([0,T];L^2(\Omega))} \leq O\left(\varepsilon^{1/2} + \dfrac{\varepsilon}{\lambda^{1/4}}\right)$ and $\|u_2 - \tilde{u}_{N_2}\|_{L^4([0,T];L^2(\Omega))} \leq O\left(\varepsilon^{1/2} + \dfrac{\varepsilon}{\lambda^{1/4}}\right)$, from the stability of solutions of NSE, we have $\|u_1 - u_2\|_{L^4([0,T];L^2(\Omega))} \leq O(\|u_{0,1} - u_{0,2}\|_{L^2(\Omega)} + \|f_1 - f_2\|_{L^4([0,T];L^2(\Omega))})$.

Therefore

$$\|u_{N_1} - u_{N_2}\|_{L^4([0,T];L^2(\Omega))} \leq O\left(\varepsilon^{1/2} + \frac{\varepsilon}{\lambda^{1/4}} + \|u_{0,1} - u_{0,2}\|_{L^2(\Omega)} + \|f_1 - f_2\|_{L^4([0,T];L^2(\Omega))}\right).$$

This implies that our scheme is approximately stable. □

Acknowledgement

J. Tian's work is supported in part by the AMS Simons Travel Grant.

References

[1] A. Barron; Universal approximation bounds for superpositions of a sigmoidal function. *IEEE Transactions on Information theory*, **39** (1993), 930-945.

[2] C. Beck, W. E, and A. Jentzen; Machine learning approximation algorithms for high-dimensional fully nonlinear partial differential equations and second-order backward stochastic differential equations. *Journal of Nonlinear Science*, **29** (2019), 1563-1619.

[3] J. Berner, P. Grohs, and A. Jentzen; 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 Journal on Mathematics of Data Science*, **2** (2020), 631-657.

[4] G. Cybenko; Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems*, **2** (1989), 303-314.

[5] P. Constantin, C. Foias; Navier-Stokes Equations. *Chicago Lectures in Mathematics*, The University of Chicago Press, 1988.

[6] L. C. Evans; Partial Differential Equations. *Graduate Studies in Mathematics*, AMS, 1998.

[7] R. Fang, D. Sondak, and P. Protopapas; Succi, Sauro; Neural network models for the anisotropic Reynolds stress tensor in turbulent channel flow, *Journal of Turbulence,* (2019), 1468-5248. DOI: 10.1080/14685248.2019.1706742.

[8] C. Foias, M. Jolly, O. Manley, and R. Rosa; Statistical estimates for the Navier–Stokes equations and the Kraichnan theory of 2-D fully developed turbulence. *Journal of Statistical Physics*, **108** (2002), 591–645.

[9] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio; Deep learning, MIT press Cambridge, 2016, Vol. 1.

[10] M. Gulian, M. Raissi, P. Perdikaris, and G. E. Karniadakis; Machine learning of space-fractional differential equations. *SIAM Journal on Scientific Computing,* **41** (2019), A2485A2509.

[11] K. He, X. Zhang, S. Ren, and J. Sun; Deep Residual Learning for Image Recognition. *Deep Residual Learning for Image Recognition*, 770-778, 2016.

[12] K. Hornik, M. Stichcombe, and W. Halbert; Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural networks*, **3** (1990), 551-560.

[13] K. Hornik; Approximation Capabilities of Muitilayer Feedforward Networks, *Neural networks*, **4** (1991), 251-257.

[14] S. Hochreiter and J. Schmidhuber; Long Short-Term Memory. *Neural computation,* **9** (1997), 1735-1780.

[15] X. Jia, L. Zhao, L. Zhang, J. He, and J. Xu; Modified Regularized Dual Averaging Method for Training Sparse Convolutional Neural Networks, *arXiv preprint arXiv:1807.04222*, 1097-1105, 2018.

[16] J. Jiménez; Machine-aided turbulence theory, *Journal of Fluid Mechanics,* **854** (2018).

[17] A. Krizhevsky, I. Sutskever, and G. E. Hinton; Imagenet Classification with Deep Convolutional Neural Networks, *Advances in neural information processing systems*, (2012), 1097-1105.

[18] J. Lions and E. Magenes; Non-homogeneous boundary value problems and applications. *Springer Science & Business Media*, 2012.

[19] T. Nguyen-Thien and T. Tran-Cong; Approximation of functions and their derivatives: A neural network implementation with applications, *Applied Mathematical Modelling*, **23** (1999), 687-704.

[20] M. Raissi,Z. Wang, M. Triantafyllou, and G. E. Karniadakis; Deep learning of vortex-induced vibrations. *Journal of Fluid Mechanics*, **861** (2019), 119-137.

[21] M. Raissi, P. Perdikaris, G. E. Karniadakis; Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, **378** (2019), 386–707.

[22] M. Raissi, G. E. Karniadakis; Hidden physics models: machine learning of nonlinear partial differential equations. *Journal of Computational Physics*, **357** (2018), 125-141.

[23] J. Sirignano, K. Spiliopoulos; DGM: A deep learning algorithm for solving partial differential equations. *Journal of Computational Physics*, **375** (2018), 1339–1364.

[24] R. Temam; Navier-Stokes equations and nonlinear functional analysis, **66**, 1995.

[25] R. Temam; Navier-Stokes Equations: theory and numerical analysis, *American Mathematical Society,* **343** (2001).

[26] N. Thuerey, K. Wei, L. Prantl, and X. Hu; Deep Learning Methods for Reynolds-Averaged NavierStokes Simulations of Airfoil Flows, *AIAA Journal,* **58** (2020), 25-36.

[27] B. Tracey, K. Duraisamy, and J. J. Alonso; A machine learning strategy to assist turbulence model development, *AIAA Paper,* **1287** (2015), p1287.

[28] J. X. Wang, J. L. Wu, and H. Xiao; Physics-informed machine learning approach for reconstructing Reynolds stress modeling discrepancies based on dns data, *Physical Review Fluids,* **2** (2017), pp. 034603.

[29] J. L. Wu, H. Xiao, and E. Paterson; Physics-informed machine learning approach for augmenting turbulence models: A comprehensive framework, *Physical Review Fluids,* **3** (2018), pp. 074602.

[30] T. Xie and F. Cao; The errors of simultaneous approximation of multivariate functions by neural networks. *Computers & Mathematics with Applications*, **61** (2011), 3146–3152.

(Animikh Biswas)
Department of Mathematics and Statistics,
University of Maryland Baltimore County,
1000 Hilltop Circle,
Baltimore, MD 21250, USA
E-mail address: `abiswas@umbc.edu`
URL: `https://userpages.umbc.edu/~abiswas/`

(Jing Tian (corresponding author))
Department of Mathematics,
Towson University,
7800 York Road,
Towson, MD 21252, USA
E-mail address: `jtian@towson.edu`

(Suleyman Ulusoy)
Department of Mathematics and Natural Sciences,
American University of Ras Al Khaimah,
P.O. Box 10021,
Ras Al Khaimah, UAE
E-mail address: `suleyman.ulusoy@aurak.ac.ae`
URL: `https://aurak.ac.ae/en/dr-suleyman-ulusoy/`