

Enhancing nonlinear solvers for the Navier–Stokes equations with continuous (noisy) data assimilation

Bosco García-Archilla^{a,1}, Xuejian Li^b, Julia Novo^{c,2}, Leo G. Rebholz^{b,*,3}

^a Departamento de Matemática Aplicada II, Universidad de Sevilla, Sevilla, Spain

^b School of Mathematical and Statistical Sciences, Clemson University, Clemson, SC, 29364, United States of America

^c Departamento de Matemáticas, Universidad Autónoma de Madrid, Spain

ARTICLE INFO

Keywords:

Navier Stokes equations
Picard iteration
Newton iteration
Continuous data assimilation

ABSTRACT

We consider nonlinear solvers for the incompressible, steady (or at a fixed time step for unsteady) Navier–Stokes equations in the setting where partial measurement data of the solution is available. The measurement data is incorporated/assimilated into the solution through a nudging term addition to the Picard iteration that penalized the difference between the coarse mesh interpolants of the true solution and solver solution, analogous to how continuous data assimilation (CDA) is implemented for time dependent PDEs. This was considered in the paper (Li et al. 2023), and we extend the methodology by improving the analysis to be in the L^2 norm instead of a weighted H^1 norm where the weight depended on the coarse mesh width, and to the case of noisy measurement data. For noisy measurement data, we prove that the CDA-Picard method is stable and convergent, up to the size of the noise. Numerical tests illustrate the results, and show that a very good strategy when using noisy data is to use CDA-Picard to generate an initial guess for the classical Newton iteration.

1. Introduction

We consider herein nonlinear solvers for the Navier–Stokes equations (NSE), which are widely used for modeling incompressible fluid flow. In particular, we study the steady state NSE, which are given on a domain $\Omega \subset \mathbb{R}^d$, $d = 2, 3$ by

$$\begin{cases} -\nu \Delta u + u \cdot \nabla u + \nabla p = f & \text{in } \Omega, \\ \nabla \cdot u = 0 & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases} \quad (1.1)$$

where u is the fluid velocity, p is the pressure, ν is the kinematic viscosity, and f is an external forcing term. The Reynolds number $Re := \frac{1}{\nu}$ is a parameter that describes how complex a flow is; higher Re is associated with more complex physics and non-unique steady solutions. While our study is restricted to the steady system (1.1) with homogeneous Dirichlet boundary conditions, the

* Corresponding author.

E-mail addresses: bosco@esi.us.es (B. García-Archilla), xuejial@clemson.edu (X. Li), julia.novo@uam.es (J. Novo), rebholz@clemson.edu (L.G. Rebholz).

¹ Research is supported by grants PID2021-123200NB-I00 and PID2022-136550NB-I00 funded by MCIN/AEI/10.13039/501100011033 and by ERDF. A way of making Europe, by the European Union.

² Research is supported by grant PID2022-136550NB-I00 funded by MCIN/AEI/10.13039/501100011033 and by ERDF. A way of making Europe, by the European Union.

³ Research, is partially supported by National Science Foundation grant DMS 2152623.

results are extendable to solving the time dependent NSE at a fixed time step in a time stepping scheme as well as nonhomogeneous mixed Dirichlet/Neumann boundary conditions.

Probably the most common nonlinear iteration for solving (1.1) is the Picard iteration, which is given by (suppressing the boundary conditions)

$$\begin{aligned} -\nu \Delta u_{k+1} + u_k \cdot \nabla u_{k+1} + \nabla p_{k+1} &= f, \\ \nabla \cdot u_{k+1} &= 0. \end{aligned}$$

The Picard iteration is globally convergent under a smallness condition on the Reynolds number/problem parameters $\alpha := M\nu^{-2}\|f\|_{H^{-1}} < 1$, where M is a domain-size dependent constant arising from Sobolev/Ladyzhenskaya inequalities [1–3]. The constant α is an upper bound on Picard's H^1 linear convergence rate [2] and is believed reasonably sharp. Moreover, $\alpha < 1$ also implies that steady NSE solutions are unique.

In [4], the CDA-Picard iteration was proposed, in order to improve convergence in the setting where solution data is available, e.g. from measurements or observed data. The CDA-Picard algorithm takes the form

$$\begin{aligned} -\nu \Delta u_{k+1} + u_k \cdot \nabla u_{k+1} + \nabla p_{k+1} + \mu I_H(u_{k+1} - u) &= f, \\ \nabla \cdot u_{k+1} &= 0, \end{aligned}$$

where $\mu > 0$ is a (user-defined) nudging parameter, and I_H is an appropriate coarse mesh interpolant (defined precisely in Section 2). We note as is typical in CDA, the true solution u remains unknown, but $I_H(u)$ is known from given or observed solution data. The general idea is that, if one knows the solution at some points in space, then one should be able to improve convergence and robustness of a solver. Indeed, in [4], it is proven analytically and illustrated numerically how CDA can accelerate convergence and enlarge the convergence basin for Picard.

The use of CDA above is based on the Azouani, Olson, and Titi algorithm from 2014 [5], and as noted in [4] its intent is for time dependent problems. It has been used on a wide variety of problems including Navier–Stokes equations and turbulence [5–7], Benard convection [8], a planetary geostrophic model [9], the Cahn–Hilliard equation [10], and many others. CDA has gained immense interest since its development and recent research includes sensitivity analyses [11], numerical analyses [10,12–20], efficient nudging methods [14], effects of noisy data [21], and using it for parameter recovery [22,23]. To our knowledge, the CDA-Picard method from [4] is the first time CDA was used to improve efficiency and robustness of nonlinear solvers for steady PDEs.

The purpose of this paper is to extend the study of CDA-Picard in two ways. First, we provide an improved convergence analysis for CDA-Picard based on the L^2 norm, instead of the weighted H^1 norm that depends on the coarse mesh width H used in the original analysis from [4]. Second, we extend the study of CDA-Picard to the more realistic case of noisy data, both analytically and numerically. We prove that for the case of noisy data, CDA-Picard is stable and convergent provided enough measurement data is known. However, the convergence is only for the L^2 residual (which is the same as the difference in successive iterates if one considers CDA-Picard as a fixed point iteration, i.e. $\|u_{k+1} - u_k\|_{L^2} = \|g_{\text{picard}}(u_k) - u_k\|_{L^2}$), as the L^2 error $\|u_{\text{nse}} - u_k\|$ is found to scale with the size of the noise in the data. Our analysis holds for any Reynolds number and forcing, i.e. it is valid in the case of non-unique steady NSE solutions, but the amount of required measurement data increases as Re and other problem parameters increase. However, we note that proofs for larger Re require different techniques and not surprisingly have slightly worse scalings. In addition to the analysis, we give results of numerical tests on two challenging test problems which illustrate the analytical results. Finally, we show numerically that a very effective solution method when given noisy data is to run CDA-Picard until its residual is sufficiently small (so that its error is as small as possible, which is on the order of the noise level), and then use its final iterate as the initial guess to usual Newton; we call this method CDA-Picard + Newton.

The paper is organized as follows. Section 2 gives preliminaries, sets notation for the paper, and gives background on CDA-Picard. Section 3 gives the improved analysis for CDA-Picard using the L^2 norm instead of the weighted H^1 norm, and Section 4 analyzes CDA-Picard for the case of noisy data. Here we analyze stability, residual convergence, and error, both for smaller Re and larger Re . Section 5 contains numerical results which illustrate the theory and also test CDA-Picard + Newton. Finally, conclusions are drawn in Section 6.

2. Preliminaries

The domain Ω will be an open connected set in $\mathbb{R}^d$ ($d = 2$ or 3) that is a convex polygon or has smooth boundary. We use $(\cdot, \cdot)$ to denote the L^2 inner product that induces the L^2 norm $\|\cdot\|$, and all other norms will have appropriate subscripts.

We denote the natural NSE function spaces for the velocity and pressure by

$$Q := \{v \in L^2(\Omega) : \int_{\Omega} v dx = 0\}, \quad (2.1)$$

$$X := \{v \in H^1(\Omega) : v = 0 \text{ on } \partial\Omega\}. \quad (2.2)$$

We use $\langle \cdot, \cdot \rangle$ to denote the duality between H^{-1} and X , and use $\|\cdot\|_{-1}$ to denote the dual norm $\|\cdot\|_{H^{-1}}$. The divergence free subspace of X is denoted by

$$V := \{v \in X : (\nabla \cdot v, q) = 0 \ \forall q \in Q\}.$$

Recall the Poincare inequality holds on X : there exists a constant C_P dependent only on the domain satisfying $\|\phi\| \leq C_P \|\nabla \phi\|$ for all $\phi \in X$.

2.1. Steady NSE preliminaries

The weak form of the NSE (1.1) is: Given $f \in H^{-1}$, find $(u, p) \in X \times Q$ such that

$$\begin{cases} v(\nabla u, \nabla v) + b(u, u, v) + (p, \nabla \cdot v) = \langle f, v \rangle & \forall v \in X, \\ (\nabla \cdot u, q) = 0 & \forall q \in Q, \end{cases} \quad (2.3)$$

where

$$b(u, w, v) = ((u \cdot \nabla)w, v) \quad \forall u, w, v \in X.$$

Since the pair (X, Q) satisfies the inf-sup condition:

$$\inf_{q \in Q} \sup_{v \in X} \frac{(q, \nabla \cdot v)}{\|\nabla v\| \|q\|} \geq \beta > 0, \quad (2.4)$$

we can instead consider the equivalent system [2]: Given $f \in H^{-1}$, find $u \in V$ satisfying

$$v(\nabla u, \nabla v) + b(u, u, v) = \langle f, v \rangle \quad \forall v \in V. \quad (2.5)$$

It is important to note that $b(u, v, v) = 0$ for all $u, v \in X$. The following inequalities hold for b [1,24]: there exists a constant M dependent only on the domain Ω such that

$$b(u, w, v) \leq M \|u\|^{\frac{1}{2}} \|\nabla u\|^{\frac{1}{2}} \|\nabla w\| \|\nabla v\|, \quad (2.6)$$

$$b(u, w, v) \leq M \|\nabla u\| \|\nabla w\| \|\nabla v\|. \quad (2.7)$$

Although this paper considers analysis in V , in discretizations using $(X_h, Q_h) \subset (X, Q)$ but where $\nabla \cdot X_h \not\subset Q_h$, the analysis of this paper will not immediately transfer to such discretizations due to the trilinear term. A fix for this is to instead use the classical skew-symmetric form of the trilinear term,

$$b^*(u, v, w) = b(u, v, w) + \frac{1}{2}((\nabla \cdot u)v, w)$$

and also grad-div stabilization. These modifications will allow all of our results to extend to the case of mixed finite elements whose velocities are only discretely divergence free. All of our computations use divergence-free Scott–Vogelius elements where $\nabla \cdot X_h \subset Q_h$ does hold.

Recall the classical well-posedness result for Eq. (2.5) [1,24]:

Lemma 2.1. *Let $\alpha = M v^{-2} \|f\|_{-1}$. For any $f \in H^{-1}$ and $v > 0$, solutions to (2.5) exist and satisfy*

$$\|\nabla u\| \leq v^{-1} \|f\|_{-1}. \quad (2.8)$$

Furthermore, if $\alpha < 1$, the solution is unique.

The restriction $\alpha < 1$ is usually referred to as the small data condition for steady NSE, but we refer to it herein as a smallness condition since we use the term ‘data’ to mean observations or measurements (and not the PDE parameters).

We will assume throughout this work that NSE solutions we consider satisfy

$$\|\nabla u\|_{\infty} =: K_1 < \infty, \quad (2.9)$$

and also that f is sufficiently regular to allow K_1 to be finite.

2.2. Picard and Newton iterations

Herein we will mostly study the Picard iteration and variations of it, but Newton will also be used in the numerical tests. These iterations can be written in their V -formulations as follows: Find $u_{k+1} \in V$ satisfying for all $v \in V$ that

$$a(u_{k+1}, v) + b(u_k, u_{k+1}, v) = \langle f, v \rangle \quad (\text{Picard}), \quad (2.10)$$

$$a(u_{k+1}, v) + b(u_k, u_{k+1}, v) + b(u_{k+1}, u_k, v) = \langle f, v \rangle + b(u_k, u_k, v) \quad (\text{Newton}). \quad (2.11)$$

We recall some basic results about the Picard iteration (2.10) from [2].

Lemma 2.2. *The Picard method (2.10) is unconditionally stable. Furthermore, if $\alpha < 1$ then the sequence $\{u_k\}$ generated by Picard converges to the NSE solution u as $k \rightarrow \infty$ with an α -linear rate.*

2.3. CDA and CDA-Picard

We now give CDA preliminaries and define CDA-Picard. Denote by $\tau_H(\Omega)$ a coarse mesh of Ω used to represent an interpolant of the true solution using measurements or observables. We assume that I_H satisfies the usual properties as other CDA applications: there exists a constant C_I independent of H satisfying

$$\|I_H v - v\| \leq C_I H \|\nabla v\| \quad \forall v \in X, \quad (2.12)$$

$$\|I_H v\| \leq C_I \|v\| \quad \forall v \in X. \quad (2.13)$$

Some examples of such interpolation operators include Bernardi–Girault [25], Scott–Zhang [26], and the L^2 projection onto piecewise constants [27].

The CDA-Picard iteration is given in its V -formulations by

$$a(u_{k+1}, v) + b(u_k, u_{k+1}, v) + \mu(I_H u_{k+1} - I_H u, I_H v) = \langle f, v \rangle \quad (\text{CDA-Picard}). \quad (2.14)$$

The term $\mu I_H(u_{k+1} - u)$ is a CDA nudging (penalty) driving the Step $k + 1$ solution u_{k+1} toward to the observations, and $\mu > 0$ is a relaxation parameter that emphasizes the observations' accuracy. We note that CDA-Picard uses a type of variational crime on the nudging term, by using I_H in the second argument of the nudging terms. While this is consistent if I_H is an L^2 projection, in general it is not consistent but using I_H additionally on the test function is key to allowing for less restrictions on parameters and no upper bound on μ [14,15,28].

Convergence of CDA-Picard is proven in [4] in terms of the following weighted H^1 norm.

$$\|v\|_* = \left(\|\nabla v\|^2 + \frac{1}{2C_I^2 H^2} \|v\|^2 \right)^{\frac{1}{2}}. \quad (2.15)$$

The result is given next. It shows that convergence is guaranteed for any α provided enough data is available, and that the convergence rate $O(H^{1/2}\alpha)$ is improved by including data measurements.

Theorem 2.1 (Convergence of CDA-Picard [4]). *Let u be a steady NSE solution and suppose $I_H(u)$ is known. If $\sqrt{2C_I H \alpha^2} < 1$ and $\mu \geq \frac{\nu}{4C_I^2 H^2}$, then for any initial guess u_0 the CDA-Picard iteration (2.14) converges to u linearly with rate (at least) $\sqrt{2C_I H \alpha}$:*

$$\|u - u_{k+1}\|_* \leq \sqrt{2C_I H \alpha} \|u - u_k\|_*. \quad (2.16)$$

3. An improved analysis for the CDA-Picard iteration

We give below a new convergence analysis for CDA-Picard under the assumption that the data measurements are accurate (the next section considers the case of noisy measurement data). While the result from Theorem 2.1 shows that data improves convergence, the $\|\cdot\|_*$ -norm is not an ideal norm to use because it is a weighted H^1 norm where the weight depends on H . So decreasing H will improve the contraction ratio, but at the same time the norm is closer to the L^2 norm and further from the H^1 norm. Hence it would be better to show this kind of convergence result in a norm not dependent on H , and below we give a new convergence results for CDA-Picard in the L^2 norm.

Theorem 3.1 (L^2 Convergence of CDA-Picard). *Let u be a steady NSE solution and suppose $I_H(u)$ is known. Set $\gamma := \frac{1}{K_1} \min \left\{ \frac{\nu}{C_I^2 H^2}, \mu \right\}$, and suppose $\mu > 2K_1$ and $H < \frac{\nu^{1/2}}{(2K_1)^{1/2} C_I}$. Then it holds*

$$\|u_{k+1} - u\| \leq \frac{\sqrt{2}}{\sqrt{\gamma}} \|u_k - u\|. \quad (3.1)$$

Since $\gamma > 2$, CDA-Picard converges linearly in L^2 and the convergence is faster whenever H decreases and μ increases.

Remark 3.1. In the finite element setting where subspace $(X_h, Q_h) \subset (X, Q)$ but where $\nabla \cdot X_h \not\subset Q_h$ (e.g. Taylor–Hood elements), the trilinear term $b(u, v, w)$ needs to be modified by a consistent skew-symmetrization (e.g. adding $\frac{1}{2}((\nabla \cdot u)v, w)$) for stability or convergence results to hold. Provided grad-div stabilization is also used, the result above will still hold with just slight modifications. In this case, one reaches the bound (3.1) with norm $\|\cdot\|_{\square}$ instead of $\|\cdot\|$ where $\|v\|_{\square} = (\|v\|^2 + \|\nabla \cdot v\|^2)^{1/2}$. Our numerical tests use divergence-free Scott–Vogelius elements where $\nabla \cdot X_h \subset Q_h$ does hold.

Proof. Denoting by $e_k = u - u_k$ and arguing as in [4] it is easy to get

$$\nu \|\nabla e_{k+1}\|^2 + \mu \|I_H e_{k+1}\|^2 = -b(e_k, u, e_{k+1}).$$

The nonlinear term treatment is a key difference from the analysis of [4], and begins with Hölder's inequality with $L^2 - L^\infty - L^2$ followed by Young's inequality and (2.9),

$$\nu \|\nabla e_{k+1}\|^2 + \mu \|I_H e_{k+1}\|^2 \leq \|e_k\| \|\nabla u\|_\infty \|e_{k+1}\|$$

$$\begin{aligned}
&\leq \frac{1}{2} \|\nabla u\|_\infty \|e_{k+1}\|^2 + \frac{1}{2} \|\nabla u\|_\infty \|e_k\|^2 \\
&\leq \frac{1}{2} K_1 \|e_{k+1}\|^2 + \frac{1}{2} K_1 \|e_k\|^2.
\end{aligned}$$

Denoting $L = \frac{1}{2} K_1$ and arguing as in [15] we can write

$$L \|e_{k+1}\|^2 \leq 2L \|I_H e_{k+1}\|^2 + 2L \|(I - I_H) e_{k+1}\|^2.$$

Since by assumption $\mu - 2L \geq \mu/2$ (which gives $\mu \geq 4L$), we get the bound

$$\nu \|\nabla e_{k+1}\|^2 + \frac{\mu}{2} \|I_H e_{k+1}\|^2 - 2L \|(I - I_H) e_{k+1}\|^2 \leq \frac{1}{2} K_1 \|e_k\|^2,$$

thanks to (2.9). Arguing again as in [15], applying (2.12) and using the assumption on H we find

$$\nu \|\nabla e_{k+1}\|^2 - 2L \|(I - I_H) e_{k+1}\|^2 \geq \nu \|\nabla e_{k+1}\|^2 - 2LC_I^2 H^2 \|\nabla e_{k+1}\|^2 \geq \frac{\nu}{2} \|\nabla e_{k+1}\|^2.$$

Combining these estimates now yields

$$\nu \|\nabla e_{k+1}\|^2 + \mu \|I_H e_{k+1}\|^2 \leq K_1 \|e_k\|^2.$$

The bound (2.12) gives us

$$\frac{\nu}{C_I^2 H^2} \|(I - I_H) e_{k+1}\|^2 + \mu \|I_H e_{k+1}\|^2 \leq \nu \|\nabla e_{k+1}\|^2 + \mu \|I_H e_{k+1}\|^2 \leq K_1 \|e_k\|^2,$$

and from the definition of γ and taking into account

$$\|I_H e_{k+1}\|^2 + \|(I - I_H) e_{k+1}\|^2 \geq \frac{1}{2} \|e_{k+1}\|^2,$$

we conclude that

$$\frac{\gamma}{2} \|e_{k+1}\|^2 \leq \|e_k\|^2.$$

Reducing implies

$$\|u_{k+1} - u\| \leq \frac{\sqrt{2}}{\sqrt{\gamma}} \|u_k - u\|,$$

which is the stated estimate. ■

4. CDA-Picard with noisy data

We consider now the case where the data used for nudging has noise (error) denoted by ϵ . That is, instead of knowing $u(x_i)$ at all the measurement points, we instead know $u(x_i) + \epsilon_i$. Hence we will write $I_H \epsilon$ to denote the coarse mesh interpolant constructed from $\{\epsilon_i\}_{i=1}^N$. Typically, these errors can be considered random, and may result for example from errors in the measurement process. In practice one may even know the relative size or distribution of the errors.

We denote by $\epsilon \in V$ any divergence-free function such that $I_H \epsilon(x_i) = \epsilon_i$, and assume that

$$\|\nabla(u + \epsilon)\| \leq (1 + \delta) \|\nabla u\|. \quad (4.1)$$

Remark 4.1. We assume the solution and noise is such that $\delta \leq O(1)$. While ϵ could scale with H^{-1} in a worst case analysis, the analysis below considers fixed H and thus our assumption on δ reduces to assuming that the signal to noise ratio is small. A more in depth study of this assumption, including from a probabilistic approach, could be a topic of future research.

It will be useful in our analysis to define

$$\alpha_\delta = (1 + \delta)\alpha = (1 + \delta)M\nu^{-2} \|f\|_{-1}, \quad (4.2)$$

and

$$C_\delta = 2\sqrt{1 + (1 + \delta)^2}. \quad (4.3)$$

Algorithm 4.2 (CDA-Picard For Noisy Data). The CDA-Picard iteration with noise is defined by: find $u_{k+1} \in V$ such that

$$\nu(\nabla u_{k+1}, \nabla v) + b(u_k, u_{k+1}, v) + \mu(I_H u_{k+1} - I_H(u + \epsilon), I_H v) = (f, v) \quad \forall v \in V. \quad (4.4)$$

4.1. Stability

Let us denote by

$$\hat{\lambda} = \min \left\{ \frac{\nu}{4C_I^2 H^2}, \frac{\mu}{2} \right\}. \quad (4.5)$$

We now prove two stability results. The first is for any data, while the second one is sharper but requires a smallness condition of the PDE parameters (not the measurement data).

Lemma 4.1 (Stability Result 1). *For any $\nu > 0$, $\alpha > 0$ and $\mu > 0$ and $\hat{\lambda}$ defined in (4.5) the $k + 1$ iterate satisfies the bound*

$$\frac{\nu}{2} \|\nabla u_{k+1}\|^2 + \hat{\lambda} \|u_{k+1}\|^2 \leq \nu^{-1} \|f\|_{-1}^2 + 2\mu C_I^2 C_P^2 \nu^{-2} \|f\|_{-1}^2 + 2\mu \|I_H \epsilon\|^2. \quad (4.6)$$

Furthermore, if $\mu \leq \frac{\nu}{2C_I^2 H^2}$, then

$$\|u_{k+1}\|^2 \leq 2\mu^{-1} \nu^{-1} \|f\|_{-1}^2 + 4C_I^2 C_P^2 \nu^{-2} \|f\|_{-1}^2 + 4\|I_H \epsilon\|^2. \quad (4.7)$$

Proof. Taking $v = u_{k+1}$ in (4.4) vanishes the nonlinear term and leaves

$$\nu \|\nabla u_{k+1}\|^2 + \mu \|I_H u_{k+1}\|^2 = (f, u_{k+1}) + \mu (I_H u, I_H u_{k+1}) + \mu (I_H \epsilon, I_H u_{k+1}).$$

Using the definition of the H^{-1} norm on the forcing term and Cauchy–Schwarz on the other right hand side terms, followed by applying Young’s inequality on all of them, we obtain

$$\nu \|\nabla u_{k+1}\|^2 + \mu \|I_H u_{k+1}\|^2 \leq \nu^{-1} \|f\|_{-1}^2 + 2\mu \|I_H u\|^2 + 2\mu \|I_H \epsilon\|^2.$$

Breaking up the viscous term on the left hand side into two pieces, using property (2.12) of I_H on this and property (2.13) on the second right hand side term and then Poincaré, we get that

$$\frac{\nu}{2} \|\nabla u_{k+1}\|^2 + \frac{\nu}{2C_I^2 H^2} \|u_{k+1} - I_H u_{k+1}\|^2 + \mu \|I_H u_{k+1}\|^2 \leq \nu^{-1} \|f\|_{-1}^2 + 2\mu C_I^2 C_P^2 \|\nabla u\|^2 + 2\mu \|I_H \epsilon\|^2.$$

With the definition of $\hat{\lambda}$ (4.5) we get from the triangle inequality a lower bound on the left hand side which provides the bound

$$\frac{\nu}{2} \|\nabla u_{k+1}\|^2 + \hat{\lambda} \|u_{k+1}\|^2 \leq \nu^{-1} \|f\|_{-1}^2 + 2\mu C_I^2 C_P^2 \nu^{-2} \|f\|_{-1}^2 + 2\mu \|I_H \epsilon\|^2,$$

thanks to the bound on u from (2.8).

The bound (4.7) can be deduced from (4.6) by dropping the first left hand side term, reducing the min term to $\frac{\mu}{2}$, multiplying both sides by $\frac{2}{\mu}$ and reducing. ■

The following stability bound assumes the PDE parameters and noise are sufficiently small, and shows that the stability bound is almost the same as the steady NSE solution bound (2.8), up to an $O(1)$ constant factor.

Lemma 4.2 (Stability Result 2). *Let α_δ and C_δ be as in (4.2), (4.3) with $\alpha_\delta < \frac{1}{3}$. Then for k sufficiently large,*

$$\|\nabla u_{k+1}\| \leq 2\nu^{-1} \|f\|_{-1} \sqrt{1 + (1 + \delta)^2} = C_\delta \nu^{-1} \|f\|_{-1}.$$

Proof. Set $\hat{u} = u + \epsilon$, and choose $v = u_{k+1} - \hat{u}$ in (4.4) to get

$$\frac{\nu}{2} (\|\nabla u_{k+1}\|^2 - \|\nabla \hat{u}\|^2 + \|\nabla(u_{k+1} - \hat{u})\|^2) + \mu \|I_H(u_{k+1} - \hat{u})\|^2 = (f, u_{k+1} - \hat{u}) + b(u_k, u_{k+1}, \hat{u}),$$

using the skew-symmetric property of b . Using the assumed bound on $\hat{u}$ (4.1), (2.7) and (2.8), we obtain

$$\begin{aligned} \frac{\nu}{2} \|\nabla u_{k+1}\|^2 + \frac{\nu}{2} \|\nabla(u_{k+1} - \hat{u})\|^2 + \mu \|I_H(u_{k+1} - \hat{u})\|^2 \\ \leq \frac{\nu}{2} \|\nabla \hat{u}\|^2 + \|f\|_{-1} \|\nabla(u_{k+1} - \hat{u})\| + M \|\nabla u_k\| \|\nabla u_{k+1}\| \|\nabla \hat{u}\| \\ \leq \frac{\nu^{-1}(1 + \delta)^2}{2} \|f\|_{-1}^2 + \frac{\nu^{-1}}{2} \|f\|_{-1}^2 + \frac{\nu}{2} \|\nabla(u_{k+1} - \hat{u})\|^2 + M \nu^{-1} \|f\|_{-1} (1 + \delta) \|\nabla u_k\| \|\nabla u_{k+1}\| \\ \leq \frac{\nu^{-1}(1 + \delta)^2}{2} \|f\|_{-1}^2 + \frac{\nu^{-1}}{2} \|f\|_{-1}^2 + \frac{\nu}{2} \|\nabla(u_{k+1} - \hat{u})\|^2 + \nu \alpha_\delta \|\nabla u_k\| \|\nabla u_{k+1}\|, \end{aligned}$$

with the last step thanks to the definition of α_δ (4.2). This reduces to

$$\nu \|\nabla u_{k+1}\|^2 + 2\mu \|I_H(u_{k+1} - \hat{u})\|^2 \leq \nu^{-1} \|f\|_{-1}^2 (1 + (1 + \delta)^2) + 2\nu \alpha_\delta \|\nabla u_k\| \|\nabla u_{k+1}\|.$$

Next, drop the second term on the left hand size, and then apply Young’s inequality to the last right hand side term via

$$2\nu \alpha_\delta \|\nabla u_k\| \|\nabla u_{k+1}\| \leq \frac{\nu}{2} \|\nabla u_{k+1}\|^2 + 2\nu \alpha_\delta^2 \|\nabla u_k\|^2,$$

to obtain the bound

$$\frac{\nu}{2} \|\nabla u_{k+1}\|^2 \leq \nu^{-1} \|f\|_{-1}^2 (1 + (1 + \delta)^2) + 2\nu\alpha_\delta^2 \|\nabla u_k\|^2.$$

This reduces to

$$\|\nabla u_{k+1}\|^2 \leq 2\nu^{-2} \|f\|_{-1}^2 (1 + (1 + \delta)^2) + 4\alpha_\delta^2 \|\nabla u_k\|^2. \quad (4.8)$$

Elementary real analysis tells us that if $a_{n+1} \leq ba_n + c$, where $a_i (i = 0, 1, \dots, n+1)$, $b, c > 0$, then $a_{n+1} \leq b^{n+1}a_0 + c \frac{1-b^{n+1}}{1-b}$. We can apply this result to (4.8), yielding

$$\|\nabla u_{k+1}\|^2 \leq (4\alpha_\delta^2)^{k+1} \|\nabla u_0\|^2 + 2\nu^{-2} \|f\|_{-1}^2 (1 + (1 + \delta)^2) \frac{1 - (4\alpha_\delta^2)^k}{1 - 4\alpha_\delta^2}.$$

Since $\alpha_\delta < \frac{1}{3}$, then for k sufficiently large, we obtain

$$\|\nabla u_{k+1}\| \leq 2\nu^{-1} \|f\|_{-1} \sqrt{1 + (1 + \delta)^2}.$$

This completes the proof. ■

4.2. Convergence of the CDA-Picard with noisy data iteration residuals

We now discuss convergence of the CDA-Picard iteration with noisy data. We are able to essentially mimic the proofs of CDA-Picard for accurate data, to prove that the residuals of the method for noisy data converges. Note that we can write $u_{k+1} - u_k = g_{\text{picard}}(u_k) - u_k$ where g_{picard} is the solution operator of the linear problem (4.4) at step k (the stability bounds above are sufficient to show g_{picard} is well defined), and hence the residuals of the iteration can be measured with the difference in successive iterates. Of course, since there is noise in the data, the limit solution of this iteration will not be particularly meaningful. Still, if one is going to run an iteration, then it is helpful to know what is the expected behavior of that iteration.

We first prove a convergence result for the residual, under a smallness condition on the PDE parameters, but for any $H > 0$ or $\mu \geq 0$. Then we prove a general convergence result with no smallness assumptions on PDE parameters, but with H small enough and μ large enough. Since these results follow very similarly to the convergence results for the case of no noise from [4] (once two successive iterations are subtracted, the noise term drops), we omit the proofs here but move them to the appendix for interested readers.

Theorem 4.3 (Convergence of the Residuals for Small PDE Parameters and Noise but Any H or μ). Suppose $\alpha_\delta < \frac{1}{3}$. Then for sufficiently large k and any H and μ , we have the bound

$$\|\nabla(u_{k+1} - u_k)\| \leq C_\delta \alpha \|\nabla(u_k - u_{k-1})\|.$$

Remark 4.4. For PDE parameters and noise sufficiently small so that $C_\delta \alpha = 2\sqrt{1 + (1 + \delta)^2} \alpha < 1$, the iteration will converge (asymptotically) linearly.

We now state a result for residual convergence for any data and noise size, but with H small enough and μ large enough. Recall the definition of $\hat{\lambda}$ in (4.5).

Theorem 4.5 (Convergence of the Residuals for General PDE Parameters and Noise, Provided H Is Small Enough and μ Is Large Enough). Suppose μ is chosen so that $\mu \geq \frac{\nu}{2C_I^2 H^2}$ and that H is small enough so that $\rho = 2C_I C_\delta^2 \alpha^2 H < 1$. Then

$$\frac{\nu}{4} \|\nabla(u_{k+1} - u_k)\|^2 + \hat{\lambda} \|u_{k+1} - u_k\|^2 \leq 2C_I C_\delta^2 \alpha^2 H \left(\frac{\nu}{4} \|\nabla(u_k - u_{k-1})\|^2 + \hat{\lambda} \|u_k - u_{k-1}\|^2 \right).$$

Remark 4.6. This result is similar to the case of convergence when no noise is present, and the noise effect is seen in the C_δ term. This result implies that smaller H and larger μ will help the iteration converge to a limit solution. However, the limit solution will likely be inaccurate if the noise level is high.

4.3. Error analysis of the CDA-Picard with noisy data iteration

We now consider error analysis of the CDA Picard iteration with noisy data. Even though we have proved above that the iteration itself will converge (under appropriate assumptions on the data, noise, and CDA parameters), it is not expected that the limit solution will be accurate. Hence we consider here the error itself, and find two results, which both show the error is bounded by the size of the noise. We denote the error at step k by $e_k := u_k - u$, where u is the NSE solution from which the measurement data were obtained.

Our first result shows that with enough measurement points and sufficiently small PDE parameters, an L^2 error estimate can be proven that depends on the size of the error but is independent of μ .

Theorem 4.7 (Error at Step k). Let $0 < r < 1$. Pick parameters μ and H such that $\frac{4\nu\alpha^4}{r^2} \leq \mu \leq \frac{\nu}{C_I^2 H^2}$ and $H \leq \frac{r}{2C_I \alpha^2}$. Then for k sufficiently large, the error in step k of the iteration satisfies

$$\|e_k\| \leq \sqrt{\frac{4}{1-r}} \|I_H \epsilon\|. \quad (4.9)$$

Remark 4.8. Note that for case of no noise in observations (i.e. formally $\epsilon = 0$), this will reduce to the same scenario as [Theorem 3.1](#), and the result (4.9) aligns with [Theorem 3.1](#) in sense that both $e_k \rightarrow 0$ as k gets large.

Proof. We begin by subtracting the scheme from the NSE (the solution u from which the measurements were taken), and then testing with e_{k+1} , to obtain

$$\nu \|\nabla e_{k+1}\|^2 + \mu \|I_H e_{k+1}\|^2 = -b(e_k, u, e_{k+1}) - \mu (I_H \epsilon, I_H e_{k+1}). \quad (4.10)$$

Bounding the last right hand side term with Cauchy–Schwarz and Young, we get that

$$-\mu (I_H \epsilon, I_H e_{k+1}) \leq \frac{\mu}{2} \|I_H \epsilon\|^2 + \frac{\mu}{2} \|I_H e_{k+1}\|^2,$$

and so

$$\nu \|\nabla e_{k+1}\|^2 + \frac{\mu}{2} \|I_H e_{k+1}\|^2 = -b(e_k, u, e_{k+1}) + \frac{\mu}{2} \|I_H \epsilon\|^2. \quad (4.11)$$

Proceeding similar to above, we can lower bound the left side via

$$\begin{aligned} \nu \|\nabla e_{k+1}\|^2 + \frac{\mu}{2} \|I_H e_{k+1}\|^2 &= \frac{\nu}{2} \|\nabla e_{k+1}\|^2 + \frac{\nu}{2} \|\nabla e_{k+1}\|^2 + \frac{\mu}{2} \|I_H e_{k+1}\|^2 \\ &\geq \frac{\nu}{2} \|\nabla e_{k+1}\|^2 + \frac{\nu}{2C_I^2 H^2} \|e_{k+1} - I_H e_{k+1}\|^2 + \frac{\mu}{2} \|I_H e_{k+1}\|^2, \end{aligned}$$

thanks to the property (2.12). Setting

$$\lambda = \min \left\{ \frac{\mu}{4}, \frac{\nu}{4C_I^2 H^2} \right\},$$

we use the triangle inequality to get the bound

$$\nu \|\nabla e_{k+1}\|^2 + \frac{\mu}{2} \|I_H e_{k+1}\|^2 \geq \frac{\nu}{2} \|\nabla e_{k+1}\|^2 + \lambda \|e_{k+1}\|^2.$$

Using this in (4.11) and then bounding the right hand side of (4.11) using (2.6), as well as bounding $\|\nabla u\|$ with (2.8) and the definition of α to get

$$\begin{aligned} \frac{\nu}{2} \|\nabla e_{k+1}\|^2 + \lambda \|e_{k+1}\|^2 &\leq -b(e_k, u, e_{k+1}) + \frac{\mu}{2} \|I_H \epsilon\|^2 \\ &\leq M \|e_k\|^{1/2} \|\nabla e_k\|^{1/2} \|\nabla u\| \|\nabla e_{k+1}\| + \frac{\mu}{2} \|I_H \epsilon\|^2 \\ &\leq \nu \alpha \|e_k\|^{1/2} \|\nabla e_k\|^{1/2} \|\nabla e_{k+1}\| + \frac{\mu}{2} \|I_H \epsilon\|^2 \\ &\leq \frac{\nu}{4} \|\nabla e_{k+1}\|^2 + \nu \alpha^2 \|e_k\| \|\nabla e_k\| + \frac{\mu}{2} \|I_H \epsilon\|^2. \end{aligned}$$

Applying Cauchy–Schwarz on the middle right hand side term and reducing yields for any $B > 0$ that

$$\begin{aligned} \frac{\nu}{4} \|\nabla e_{k+1}\|^2 + \lambda \|e_{k+1}\|^2 &\leq \nu \frac{B}{2} \|\nabla e_k\|^2 + \frac{\nu \alpha^4}{2B} \|e_k\|^2 + \frac{\mu}{2} \|I_H \epsilon\|^2 \\ &\leq r \left(\nu \frac{B}{2r} \|\nabla e_k\|^2 + \frac{\nu \alpha^4}{2Br} \|e_k\|^2 \right) + \frac{\mu}{2} \|I_H \epsilon\|^2, \end{aligned}$$

where $0 < r < 1$. Setting $B = \frac{r}{2}$ and noting the assumptions on μ and H from the theorem statement imply that

$$\frac{\nu \alpha^4}{r^2} \leq \lambda = \min \left\{ \frac{\mu}{4}, \frac{\nu}{4C_I^2 H^2} \right\},$$

and thus since $r < 1$

$$\begin{aligned} \frac{\nu}{4} \|\nabla e_{k+1}\|^2 + \lambda \|e_{k+1}\|^2 &\leq r \left(\frac{\nu}{4} \|\nabla e_k\|^2 + \lambda \|e_k\|^2 \right) + \frac{\mu}{2} \|I_H \epsilon\|^2 \\ &\leq r^2 \left(\frac{\nu}{4} \|\nabla e_{k-1}\|^2 + \lambda \|e_{k-1}\|^2 \right) + (1+r) \frac{\mu}{2} \|I_H \epsilon\|^2 \\ &\leq \dots \\ &\leq r^{k+1} \left(\frac{\nu}{4} \|\nabla e_0\|^2 + \lambda \|e_0\|^2 \right) + \frac{\mu}{2(1-r)} \|I_H \epsilon\|^2. \end{aligned}$$

Thus for k sufficiently large such that

$$r^{k+1} \left(\frac{\nu}{4} \|\nabla e_0\|^2 + \lambda \|e_0\|^2 \right) \leq \frac{\mu}{2(1-r)} \|I_H \epsilon\|^2,$$

and using the assumption that $\mu \leq \frac{\nu}{C_I^2 H^2}$ and dropping the positive term $\frac{\nu}{4} \|\nabla e_{k+1}\|^2$,

$$\frac{\mu}{4} \|e_{k+1}\|^2 = \min \left\{ \frac{\mu}{4}, \frac{\nu}{4C_I^2 H^2} \right\} \|e_k\|^2 = \lambda \|e_k\|^2 \leq \frac{\mu}{1-r} \|I_H \epsilon\|^2,$$

which reduces to

$$\|e_k\| \leq \sqrt{\frac{4}{1-r}} \|I_H \epsilon\|, \quad (4.12)$$

finishing the proof. ■

The theorem above proved an error bound under (essentially) a smallness condition on the PDE parameters. The theorem below proves an error bound for any data, and does not require the NSE solution be unique. However, we do require a single NSE solution is used for measurement data. We denote

$$\bar{\lambda} := \min \left\{ \frac{\nu}{C_I^2 H^2}, \frac{\mu}{2} \right\} \quad (4.13)$$

in the theorem.

Theorem 4.9 (Error at Step k For General PDE Parameters). *Let u be a steady NSE solution and $I_H(u)$ be its interpolant on X_H . Pick parameters μ and H such that $\bar{\lambda} > 2$, $H \leq \frac{\nu^{1/2}}{\sqrt{2}K_1 C_I}$, and $\mu \geq 4K_1^2$. Then it holds that*

$$\|e_k\|^2 \leq \left(\frac{2}{\bar{\lambda}} \right)^k \|e_0\|^2 + \frac{2\mu}{\bar{\lambda}-2} \|I_H \epsilon\|^2. \quad (4.14)$$

Moreover, if $\mu \leq \frac{\nu}{C_I^2 H^2}$ (equivalently $H \leq \frac{\sqrt{2\nu^{1/2}}}{\mu^{1/2} C_I}$) then

$$\|e_k\|^2 \leq \left(\frac{2}{\bar{\lambda}} \right)^k \|e_0\|^2 + \frac{4\mu}{\mu-4} \|I_H \epsilon\|^2. \quad (4.15)$$

Remark 4.10. Hence we observe linear convergence of the error in the L^2 norm up to $\left(\frac{2\mu}{\bar{\lambda}-2} \right)^{1/2} \|I_H \epsilon\|$, provided μ is large enough and H is small enough to satisfy the assumptions. Compared to Theorem 4.7, the improvement of this theorem is that there is no restriction on PDE parameters Re , f , and M . The tradeoff is that the term representing the noise is scaled by μ , which depends on $\|\nabla u\|_{L^\infty}^2$, see (4.14). However, assuming $\mu \leq \frac{\nu}{C_I^2 H^2}$, as in Theorem 4.7, or equivalently H small enough, $H \leq \frac{\sqrt{2\nu^{1/2}}}{\mu^{1/2} C_I}$, we get (4.15) in which the constant $\frac{4\mu}{\mu-4}$ is bounded as $\mu \rightarrow \infty$.

Remark 4.11. The inequality (4.14) gives similar result compared to the CDA convergence results with noise in [29][Theorem 4.1, 4.5], in which the solution error is scaled by the parameter μ . If we assume either sufficiently small H or small α , then the estimates (4.12) and (4.15) above are able to remove this scaling.

Proof. We begin the proof the same was as for the proof of Theorem 4.7 and thus begin with the error equation

$$\nu \|\nabla e_{k+1}\|^2 + \mu \|I_H e_{k+1}\|^2 = -b(e_k, u, e_{k+1}) - \mu(I_H \epsilon, I_H e_{k+1}).$$

Applying Young's inequality on the last right hand side term (just as in Theorem 4.7, but now estimating the nonlinear term using Hölder's inequality with $L^2 - L^\infty - L^2$) gives using (2.9)

$$\begin{aligned} \nu \|\nabla e_{k+1}\|^2 + \frac{\mu}{2} \|I_H e_{k+1}\|^2 &\leq \|e_k\| \|\nabla u\|_\infty \|e_{k+1}\| + \frac{\mu}{2} \|I_H \epsilon\|^2 \\ &\leq \frac{1}{2} \|\nabla u\|_\infty^2 \|e_{k+1}\|^2 + \frac{1}{2} \|e_k\|^2 + \frac{\mu}{2} \|I_H \epsilon\|^2 \\ &\leq \frac{1}{2} K_1^2 \|e_{k+1}\|^2 + \frac{1}{2} \|e_k\|^2 + \frac{\mu}{2} \|I_H \epsilon\|^2. \end{aligned}$$

with the last step thanks to Young's inequality. Denoting $L := \frac{1}{2} K_1^2$ and arguing as in [15] we can write

$$L \|e_{k+1}\|^2 \leq 2L \|I_H e_{k+1}\|^2 + 2L \|(I - I_H) e_{k+1}\|^2. \quad (4.16)$$

Since $\mu \geq 4K_1^2$ is assumed, we have that $\mu \geq 8L$ and so $\mu/2 - 2L \geq \mu/4$ and thus

$$\nu \|\nabla e_{k+1}\|^2 + \frac{\mu}{4} \|I_H e_{k+1}\|^2 - 2L \|(I - I_H) e_{k+1}\|^2 \leq \frac{1}{2} \|e_k\|^2 + \frac{\mu}{2} \|I_H \epsilon\|^2.$$

Now using (2.12) and the assumption on H we get that

$$\nu \|\nabla e_{k+1}\|^2 - 2L \|(I - I_H) e_{k+1}\|^2 \geq \nu \|\nabla e_{k+1}\|^2 - 2LC_I^2 H^2 \|\nabla e_{k+1}\|^2 \geq \frac{\nu}{2} \|\nabla e_{k+1}\|^2,$$

which combines with the previous inequality to yield

$$\nu \|\nabla e_{k+1}\|^2 + \frac{\mu}{2} \|I_H e_{k+1}\|^2 \leq \|e_k\|^2 + \mu \|I_H \epsilon\|^2.$$

Applying (2.12) again to lower bound the left hand side (similar to in the proofs above) gives us that

$$\begin{aligned} \nu C_I^{-2} H^{-2} \|(I - I_H)e_{k+1}\|^2 + \frac{\mu}{2} \|I_H e_{k+1}\|^2 &\leq \nu \|\nabla e_{k+1}\|^2 + \frac{\mu}{2} \|I_H e_{k+1}\|^2 \\ &\leq \|e_k\|^2 + \mu \|I_H \epsilon\|^2. \end{aligned}$$

Next, using the definition of $\bar{\lambda}$ and (4.16) (after dividing both sides by $2L$) yields

$$\frac{\bar{\lambda}}{2} \|e_{k+1}\|^2 \leq \|e_k\|^2 + \mu \|I_H \epsilon\|^2,$$

so that

$$\|u_{k+1} - u\|^2 \leq \frac{2}{\bar{\lambda}} \|u_k - u\|^2 + \frac{2\mu}{\bar{\lambda}} \|I_H \epsilon\|^2.$$

Denoting $r = \frac{2}{\bar{\lambda}} < 1$, (by assumption of the theorem) we can step the iteration backward and obtain the bound

$$\begin{aligned} \|u_{k+1} - u\|^2 &\leq r^{k+1} \|u_0 - u\|^2 + (1 + r + \dots + r^k) \mu r \|I_H \epsilon\|^2 \\ &\leq r^{k+1} \|u_0 - u\|^2 + \frac{\mu r}{1 - r} \|I_H \epsilon\|^2 \\ &\leq \left(\frac{2}{\bar{\lambda}}\right)^{k+1} \|u_0 - u\|^2 + \frac{2\mu}{\bar{\lambda} - 2} \|I_H \epsilon\|^2. \end{aligned}$$

This proves (4.14). In case $\mu \leq \frac{\nu}{C_I^2 H^2}$ then $\bar{\lambda} = \mu/2$ which gives (4.15). ■

5. Numerical experiments

In this section we illustrate the above theory and the effectiveness of the proposed method on the 2D driven cavity at $Re=3000$ and 10,000, and the 3D driven cavity at $Re=200$ and 1000. CDA-Picard L^2 error is shown in all tests to converge to approximately the level of the signal to noise ratio while its L^2 residual (the difference between iterates) converges linearly to 0. In addition, we propose and test the strategy of using CDA-Picard with noisy data to create an initial guess for usual Newton. This is shown to be a very effective solver for these test problems.

5.1. 2D driven cavity with varying signal to noise ratio

For our first test we consider the CDA-Picard method for the NSE applied to the benchmark 2D driven cavity problem on $\Omega = (0, 1)^2$. Boundary conditions are zero on the sides and bottom, and on the top we enforce $u = \langle 1, 0 \rangle^T$ which models a lid moving horizontally with unit velocity. No external forcing is used in this problem ($f = 0$) and the viscosity ν will be chosen as the inverse of the Reynolds number, and we test with $Re=3000$, and 10000 (the first of which is believed to satisfy the smallness condition $\alpha < 1$ and the second of which does not [3]).

Computations are performed using (P_2, P_1^{disc}) Scott-Vogelius elements on a barycenter refinement (also named the Alfeld split in the Guzman-Neilan vernacular) of a uniform $\frac{1}{64}$ triangulation that has southwest-northeast diagonals. It is known from [30] that this element choice is inf-sup stable on this type of structured mesh. This discretization provides 98,818 total velocity degrees of freedom (dof) and 73,728 pressure dof. For all of these tests, an initial iterate of $u_0 = 0$ is chosen. Plots of the NSE solutions found on this discretization for these problems are shown in Fig. 1, and they agree well with existing literature [31,32].

We obtain noisy data measurements by first computing the NSE solution directly on the given discretization for each Re (using the AA-Picard method from [3]). Measurement points are taken to be the vertices of the mesh that are closest to the midpoints of an $N \times N$ uniform square grid of Ω , where $N=10$ for $Re=3000$ and $N=20$ for $Re=10,000$. Signal to noise ratios (snr) were chosen to be 0.001, 0.01 and 0.05, and for each snr , random numbers were generated at each measurement point that come from a normal distribution with mean 0 and standard deviation $snr \times u_{max}$ (u_{max} is the max velocity of the true solution, which is 1 for this test problem). These noisy measurement data points were then used with CDA-Picard. For $Re=3000$, 100 total measurement points were used and for $Re=10,000$, 400 total measurement points were used (100 was not enough to yield convergence for $Re=10,000$).

Convergence results for $Re = 3000$ are shown in Fig. 2 for each choice of snr , and for both small and large nudging parameters ($\mu = 1$ and $\mu = 10,000$). We observe that in all cases, the L^2 residuals converge approximately linearly, with little difference between large and small μ and with the total number of iterations to convergence increasing slightly as the snr increases. The L^2 error plots, as expected, only converged up to approximately the level of the snr .

Results for $Re = 10000$ are shown in Fig. 3. The L^2 error again converges approximately to the level of the snr . The convergence of the L^2 residual shows somewhat different behavior than the $Re = 3000$ case. While as expected residual convergence is slower with higher Re , we observe better residual convergence for $snr = 0.01$ than for 0.001, which is not expected. We also observe for $snr = 0.05$ that the larger nudging parameter gives somewhat worse behavior. Of course, the interest in the L^2 residual is only useful in how it is associated with error.

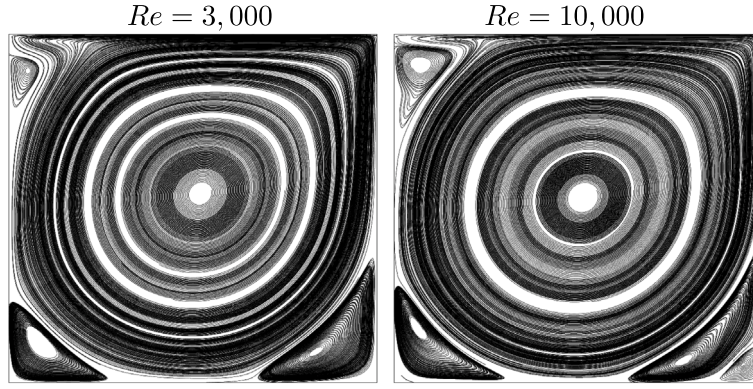


Fig. 1. The plot above shows streamlines of the solution of the 2D driven cavity problem at $Re = 3,000$ (left) and $10,000$ (right).

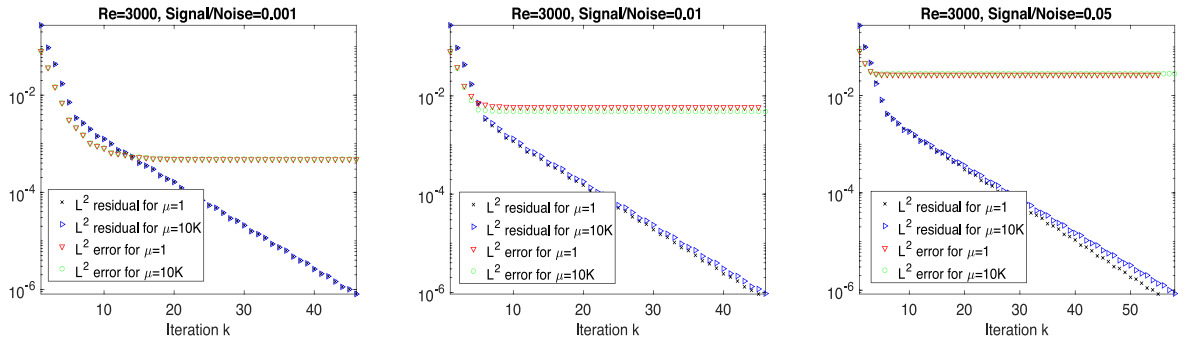


Fig. 2. Shown above are the error and residual plots for $Re=3000$ driven cavity tests with varying snr .

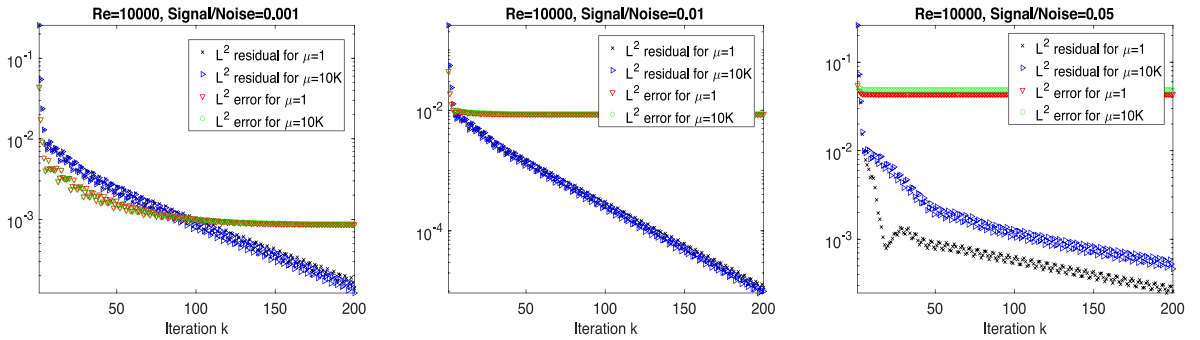


Fig. 3. Shown above are the error and residual plots for $Re=10,000$ driven cavity tests with varying snr .

We also noticed that it is practical to set the convergence tolerance of CDA-Picard algorithm as $10^{-2} \times snr$, since after this threshold further CDA-Picard iterations does not help reducing the error anymore. This aligns well with our theory that the CDA-Picard converges to a limit with distance proportional to snr away from the true solution. We only need to set the tolerance up to this solution precision, and using a scalar such as 10^{-2} to multiply snr ensures that all relevant information in observations has been thoroughly captured during iterations. In addition, it can be better to use a smaller μ (mainly depending on observation data H and the Reynolds number) in the CDA-Picard algorithm to avoid negative effects of the observation noise, especially when the noise is large. Such phenomena appears particularly in the third plot of Figs. 2 and 3. This fact also aligns well with our theory (inequality (4.14) in Theorem 4.9).

5.2. 3D driven cavity with varying signal to noise ratio

Our next test for CDA-Picard with noisy data uses the 3D lid-driven cavity benchmark problem. This problem is the 3D analogue of the test problem above: the domain is the unit cube, there is no forcing, homogeneous Dirichlet boundary conditions are enforced

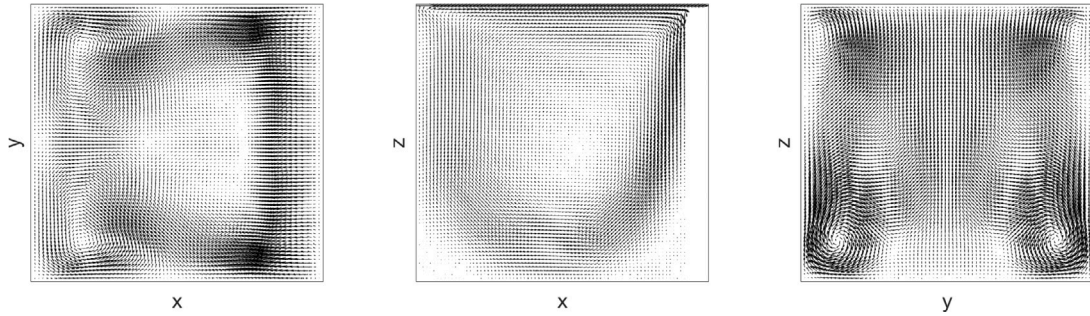


Fig. 4. Shown above are the midsliceplane plots of solutions for the 3D driven cavity simulations at $Re=1000$.

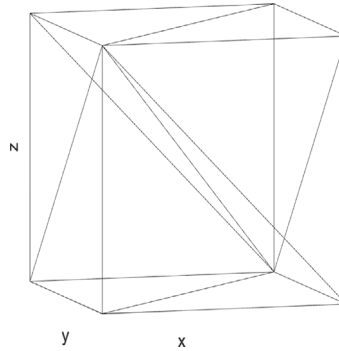


Fig. 5. Shown above is the method used to split a rectangular box into 6 tetrahedra.

on the walls and at the top of the box $u = \langle 1, 0, 0 \rangle$ represents the moving lid. The viscosity will be chosen as the inverse of the Reynolds number, and we will use $Re=200$ and 1000 . The tests will use $u_0 = 0$ for the initial condition for the iterations.

The mesh is constructed using Chebychev points on $[0,1]$ to first construct a $\mathcal{M} \times \mathcal{M} \times \mathcal{M}$ grid of rectangular boxes (we use $\mathcal{M}=11$ for $Re=200$ and $\mathcal{M}=13$ for $Re=1000$). Each box is then split into 6 tetrahedra in the manner of Fig. 5, and then each of these tetrahedra is split into 4 tetrahedra with a barycenter refinement (Alfeld split). This mesh is equipped with (P_3, P_2^{disc}) Scott–Vogelius elements, and this provides for approximately 796K total dof for the $Re=200$ tests and 1.3 million total dof for the $Re=1000$ tests. We note this velocity–pressure pair is known to be LBB stable on such a mesh construction from [33]. Solution plots found with this discretization matched those from the literature [34], and we show midsliceplanes of the $Re=1000$ solution in Fig. 4.

The measurement data is constructed similar to the 2D case above. Measurement points are taken to be the tetrahedra vertices closest to the centers of a uniform grid of $N \times N \times N$ cubes. Measurement data is taken to be the true solution (found using AA-Picard method of [3]) with noise added at each component from random numbers with normal distribution having mean 0 and standard deviation snr . We again take snr to be 0.001, 0.01 and 0.05 for these tests. These tests used $N=4$ (64 total measurement points) for $Re=200$ and $N=9$ (729 total measurement points) for $Re=1000$; for the latter case, $N=8$ was not sufficiently large to yield convergence (recall the theory above requires H sufficiently small). The nudging parameter $\mu=1$ was used, and we note that results from larger nudging parameters gave very similar results.

For $Re=200$, convergence of the L^2 residuals and L^2 errors for $snr = 0.001, 0.01$ and 0.05 are shown in Fig. 6, and for $Re=1000$ they are shown in Fig. 7. We observe approximately linear convergence of the L^2 residual down to the set tolerance of 10^{-8} in all cases, and also in all cases we observe the L^2 error to converge only to the level of the snr . Note that the $Re=1000$ tests converge faster than the case of $Re = 200$ but this is likely because many more measurement points were used (729 compared to 64).

5.3. Newton with CDA-Picard generated initial guess

For our last test, we consider using CDA-Picard with noisy data for the purpose of generating an initial guess for the (usual) Newton iteration. From the tests and analysis above, we observe that CDA-Picard's L^2 residual will converge (provided enough measurement data), however the L^2 error will only converge up to approximately the level of the signal to noise ratio. Also observed from the tests above, the L^2 residual and error are close to each other in value, until the L^2 error bottoms out near the snr while the L^2 residual continues to converge linearly. While noisy data will create a lower bound on the error for CDA-Picard iterates that is on the order of the signal to noise ratio, these iterates may be sufficiently close to the root that it can allow for Newton to converge if they are used as initial guesses. Indeed, we find below that this strategy can be quite effective.

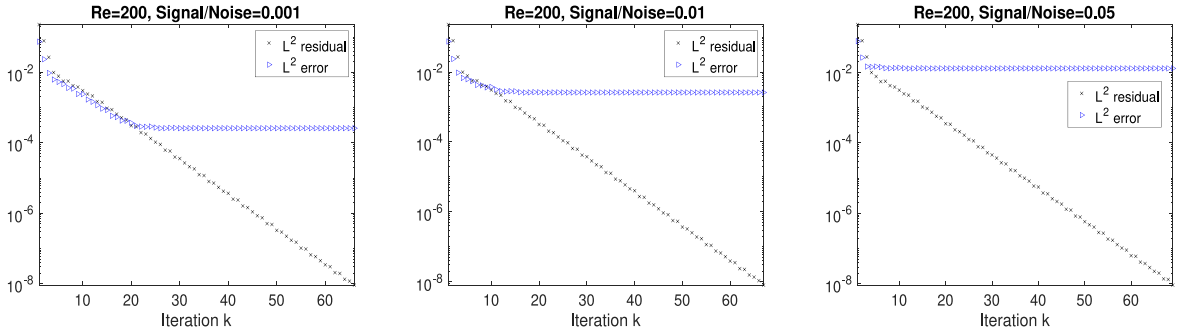


Fig. 6. Shown above are the error and residual plots for $Re=200$ 3D driven cavity tests with varying snr .

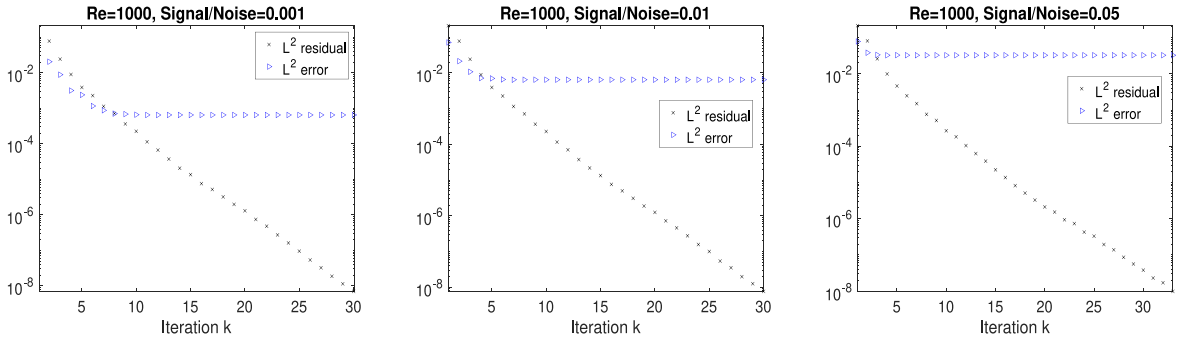


Fig. 7. Shown above are the error and residual plots for $Re=1000$ 3D driven cavity tests with varying snr .

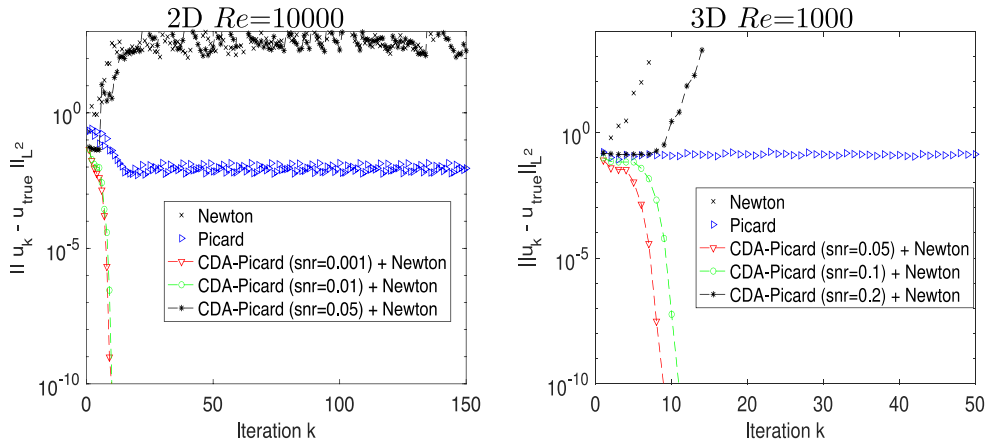


Fig. 8. Shown above are L^2 error vs iteration number, for the 2D driven cavity at $Re=10,000$ (left) and the 3D driven cavity at $Re=1000$ (right) for Newton, Picard, and CDA-Picard + Newton.

We now test this idea for the 2D driven cavity at $Re=10,000$ and 3D driven cavity with $Re=1000$ problems tested above, with the same discretizations and parameters. We note that both Newton and Picard iterations fail for these problems if an initial guess of $u_0 = 0$ is used. We run CDA-Picard with noisy data (and varying snr) until the L^2 residual falls below 10^{-2} , and then use the final iterate as the initial guess for the Newton iteration. We refer to this method as CDA-Picard + Newton below. We note that we also tested a 10^{-3} tolerance and got the same results, but 10^{-1} was not sufficiently accurate for Newton to converge. We also note that $\mu = 1$ was the nudging parameter, but results using $\mu = 10000$ gave very similar results.

Convergence results are shown in Fig. 8 as L^2 error vs iteration count for CDA-Picard + Newton for varying snr , usual Newton and usual Picard. The usual Newton, usual Picard, and CDA-Picard all use $u_0 = 0$ as the initial guess. We observe that usual Picard fails to converge for both tests, although it remains stable (as expected [2]). The usual Newton iteration also fails and the iterates get very large; in the 2D case direct solvers are used and the iteration is able to continue but in 3D the linear solvers fail after the size of the iterate grows above 10^4 (we use method of [35,36]). For CDA-Picard + Newton, however, convergence is achieved

provided the snr is not too large. For the 2D problem, convergence is achieved with $snr = 0.001$ and 0.01 , but the method fails to converge with $snr = 0.05$. For the 3D problem, convergence is achieved for $snr = 0.05$ and 0.1 but not for 0.2 . Hence as expected, in both cases there is a level of noise which is too high to provide a good enough initial guess for Newton to converge. Overall, these results show CDA-Picard + Newton is quite effective, provided the noise level is not too large.

6. Conclusions and future directions

We have extended the CDA-Picard methodology in this paper. We have improved the analysis by providing convergence results in more appropriate norms than the original analysis provides, and also extended the methodology and analysis to the more realistic case of noisy data. Our analytical and numerical results show that the L^2 error scales with the relative size of the noise. For large relative noise levels, however, we showed that using the CDA-Picard iteration to generate initial guesses for the usual Newton iteration can be very effective.

For future work, we plan to consider applications of CDA-Picard to multiphysics problems related to the NSE. Additionally, we plan to look for methods to reduce the amount of data required by the method, in particular in the 3D case.

CRedit authorship contribution statement

Bosco García-Archilla: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Xuejian Li:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Julia Novo:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Leo G. Rebholz:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Leo Rebholz reports financial support was provided by National Science Foundation. Julia Novo and Bosco Garcia-Archilla report financial support was provided by the State Agency of Research and by the European Union. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix

Proof of Theorem 4.3. Subtracting the iteration at Step $k + 1$ from Step k , and testing with $u_{k+1} - u_k$ gives

$$\nu \|\nabla(u_{k+1} - u_k)\|^2 + \mu \|I_H(u_{k+1} - u_k)\|^2 \leq -b(u_k - u_{k-1}, u_k, u_{k+1} - u_k), \quad (\text{A.1})$$

thanks to $b(u_{k-1}, u_{k+1} - u_k, u_{k+1} - u_k) = 0$. Using (2.7) and Young's inequality, we obtain

$$\nu \|\nabla(u_{k+1} - u_k)\|^2 + \mu \|I_H(u_{k+1} - u_k)\|^2 \leq M \|\nabla(u_k - u_{k-1})\| \|\nabla u_k\| \|\nabla(u_{k+1} - u_k)\|.$$

Using Lemma 4.2 (for sufficiently large k) on the first right hand side term and dropping the second (positive) left hand side term, we get the stated bound

$$\|\nabla(u_{k+1} - u_k)\| \leq C_\delta \alpha \|\nabla(u_k - u_{k-1})\|. \quad \blacksquare$$

Proof of Theorem 4.5. Starting from (A.1) and using (2.6) and Young's inequality, we obtain

$$\nu \|\nabla(u_{k+1} - u_k)\|^2 + \mu \|I_H(u_{k+1} - u_k)\|^2 \leq M \|\nabla(u_k - u_{k-1})\|^{1/2} \|u_k - u_{k-1}\|^{1/2} \|\nabla u_k\| \|\nabla(u_{k+1} - u_k)\|.$$

Proceeding as in Lemma 4.1 to lower bound the left hand side and using Lemma 4.2 (for sufficiently large k) on the first right hand side term along with Young's inequality gives

$$\frac{\nu}{4} \|\nabla(u_{k+1} - u_k)\|^2 + \hat{\lambda} \|u_{k+1} - u_k\|^2 \leq C_\delta^2 \nu \alpha^2 \|\nabla(u_k - u_{k-1})\| \|u_k - u_{k-1}\|.$$

Applying Young's inequality again to the right hand side gives

$$\begin{aligned} \frac{\nu}{4} \|\nabla(u_{k+1} - u_k)\|^2 + \hat{\lambda} \|u_{k+1} - u_k\|^2 &\leq \rho \frac{\nu}{4} \|\nabla(u_k - u_{k-1})\|^2 + \frac{C_\delta^4 \nu \alpha^4}{\rho} \|u_k - u_{k-1}\|^2 \\ &\leq \rho \left(\frac{\nu}{4} \|\nabla(u_k - u_{k-1})\|^2 + \frac{C_\delta^4 \nu \alpha^4}{\rho^2} \|u_k - u_{k-1}\|^2 \right). \end{aligned}$$

Using the stated choice of ρ and assumption that it is less than 1, we obtain the stated result,

$$\frac{\nu}{4} \|\nabla(u_{k+1} - u_k)\|^2 + \hat{\lambda} \|u_{k+1} - u_k\|^2 \leq 2C_I C_\delta^2 \alpha^2 H \left(\frac{\nu}{4} \|\nabla(u_k - u_{k-1})\|^2 + \hat{\lambda} \|u_k - u_{k-1}\|^2 \right). \quad \blacksquare$$

References

- [1] W. Layton, An Introduction to the Numerical Analysis of Viscous Incompressible Flows, SIAM, Philadelphia, 2008.
- [2] V. Girault, P.-A. Raviart, Finite Element Methods for Navier–Stokes Equations: Theory and Algorithms, Springer-Verlag, 1986.
- [3] S. Pollock, L. Rebholz, M. Xiao, Anderson-accelerated convergence of Picard iterations for incompressible Navier-Stokes equations, *SIAM J. Numer. Anal.* 57 (2019) 615–637.
- [4] X. Li, E. Hawkins, L. Rebholz, D. Vargun, Accelerating and enabling convergence of nonlinear solvers for Navier-Stokes equations by continuous data assimilation, *Comput. Methods Appl. Mech. Engrg.* 416 (2023) 1–17.
- [5] A. Azouani, E. Olson, E.S. Titi, Continuous data assimilation using general interpolant observables, *J. Nonlinear Sci.* 24 (2014) 277–304.
- [6] P.C. Di Leoni, A. Mazzino, L. Biferale, Synchronization to big data: nudging the Navier-Stokes equations for data assimilation of turbulent flows, *Phys. Rev. X* 10 (011023) (2020).
- [7] A. Farhat, E. Lunasin, E.S. Titi, A Data Assimilation Algorithm: the Paradigm of the 3D Leray- α Model of Turbulence, in: London Mathematical Society Lecture Note Series, Cambridge University Press, 2019, pp. 253–273.
- [8] A. Farhat, M.S. Jolly, E.S. Titi, Continuous data assimilation for the 2D Bénard convection through velocity measurements alone, *Physica D* 303 (2015) 59–66.
- [9] A. Farhat, E. Lunasin, E.S. Titi, On the charney conjecture of data assimilation employing temperature measurements alone: The paradigm of 3D planetary geostrophic model, *Math. Clim. Weather Forecast.* 2 (1) (2016).
- [10] A.E. Diegel, L.G. Rebholz, Continuous data assimilation and long-time accuracy in a C^0 interior penalty method for the Cahn-Hilliard equation, *Appl. Math. Comput.* 424 (2022) 127042.
- [11] E. Carlson, A. Larios, Sensitivity analysis for the 2D Navier–Stokes equations with applications to continuous data assimilation, *J. Nonlinear Sci.* 31 (84) (2021).
- [12] A.H. Ibdah, C.F. Mondaini, E.S. Titi, Fully discrete numerical schemes of a data assimilation algorithm: Uniform-in-time error estimates, *IMA J. Numer. Anal.* 40 (4) (2019) 2584–2625, 11.
- [13] A. Larios, L. Rebholz, C. Zervas, Global in time stability and accuracy of IMEX-FEM data assimilation schemes for Navier–Stokes equations, *Comput. Methods Appl. Mech. Engrg.* 345 (2019) 1077–1093.
- [14] L.G. Rebholz, C. Zervas, Simple and efficient continuous data assimilation of evolution equations via algebraic nudging, *Numer. Methods Partial Differential Equations* 37 (3) (2021) 2588–2612.
- [15] B. García-Archilla, J. Novo, E. Titi, Uniform in time error estimates for a finite element method applied to a downscaling data assimilation algorithm, *SIAM J. Numer. Anal.* 58 (2020) 410–429.
- [16] M.S. Jolly, A. Pakzad, Data assimilation with higher order finite element interpolants, *Int. J. Num. Methods* 95 (2023) 472–490.
- [17] A. Biswas, J. Tian, S. Ulusoy, Error estimates for deep learning methods in fluid dynamics, *Numer. Math.* 151 (3) (2022) 753–777.
- [18] A. Biswas, K.R. Brown, V.R. Martinez, Mesh-free interpolant observables for continuous data assimilation, *Ann. Appl. Math.* 38 (3) (2022) 296–355.
- [19] M.A. Hammoud, E.S. Titi, I. Hoteit, O. Knio, Cdanet: A physics-informed deep neural network for downscaling fluid flows, *J. Adv. Modelling Earth Syst.* 14 (12) (2022).
- [20] R. Hu, Q. Lin, A. Raydan, S. Tang, Higher-order error estimates for physics-informed neural networks approximating the primitive equations, *Partial Differ. Equ. Appl.* 4 (4) (2023) 35, Paper (34).
- [21] C. Foias, C.F. Mondaini, E.S. Titi, A discrete data assimilation scheme for the solutions of the two-dimensional Navier-Stokes equations and their statistics, *SIAM J. Appl. Dyn. Syst.* 14 (4) (2016) 2109–2142.
- [22] E. Carlson, J. Hudson, A. Larios, Parameter recovery for the 2 dimensional Navier-Stokes equations via continuous data assimilation, *SIAM J. Sci. Comput.* 42 (1) (2020) A250–A270.
- [23] A. Biswas, J. Hudson, Determining the viscosity of the Navier-Stokes equations from observations of finitely many modes, *Inverse Problems* 39 (12) (2023) 29, Paper (125012).
- [24] R. Temam, Navier-Stokes Equations, Elsevier North-Holland, 1991.
- [25] C. Bernardi, V. Girault, A local regularization operator for triangular and quadrilateral finite elements, *SIAM J. Numer. Anal.* 35 (5) (1998) 1916–18930.
- [26] L.R. Scott, S. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions, *Math. Comp.* 54 (190) (1990) 483–493.
- [27] A. Ern, J.L. Guermond, Theory and Practice of Finite Elements, in: Applied Mathematical Sciences, vol. 159, Springer-Verlag, New York, 2004.
- [28] B. García-Archilla, J. Novo, Error analysis of fully discrete mixed finite element data assimilation schemes for the Navier-Stokes equations, *Adv. Comput. Math.* (2020) 46–61.
- [29] H. Bessaih, E. Olson, E.S. Titi, Continuous data assimilation with stochastically noisy data, *Nonlinearity* 28 (3) (2015) 729–753.
- [30] D. Arnold, J. Qin, Quadratic velocity/linear pressure Stokes elements, in: R. Vichnevetsky, D. Knight, G. Richter (Eds.), *Advances in Computer Methods for Partial Differential Equations VII*, IMACS, 1992, pp. 28–34.
- [31] C.-H. Bruneau, M. Saad, The 2D lid-driven cavity problem revisited, *Comput. & Fluids* 35 (3) (2006) 326–348.
- [32] E. Erturk, T.C. Corke, C. Gökcöl, Numerical solutions of 2D-steady incompressible driven cavity flow at high Reynolds numbers, *Internat. J. Numer. Methods Fluids* 48 (2005) 747–774.
- [33] S. Zhang, A new family of stable mixed finite elements for the 3D Stokes equations, *Math. Comput.* 74 (250) (2005) 543–554.
- [34] K.L. Wong, A.J. Baker, A 3D incompressible Navier–Stokes velocity–vorticity weak form finite element algorithm, *Internat. J. Numer. Methods Fluids* 38 (2) (2002) 99–123.
- [35] M. Benzi, M. Olshanskii, An augmented Lagrangian-based approach to the Oseen problem, *SIAM J. Sci. Comput.* 28 (2006) 2095–2113.
- [36] T. Heister, G. Rapin, Efficient augmented Lagrangian-type preconditioning for the oseen problem using grad-div stabilization, *Internat. J. Numer. Methods Fluids* 71 (2013) 118–134.