MA916 Research Study Group

Data Assimilation

Report

Charles Brett
Andrew Lam
David McCormick
Michael Scott

May 13, 2011

Abstract

In this report, we consider 3DVar data assimilation for the 2D Navier–Stokes equations. We prove analytic results to show that, for certain choices of the covariance operators, different estimators with the same observations converge geometrically, and that estimators come within bounded distance from the true solution provided the observation noise is uniformly bounded. We back this up with numerical evidence, and predict that the theorems should hold for other choices of the covariance operators.

Contents

1 Introduction

Data assimilation is a technique used to reconcile predictions from a model and observations of reality step-by-step. In our case the model we consider will be the Navier–Stokes equations on the torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ (equivalently, we consider $[0,1) \times [0,1)$ with periodic boundary conditions).

In a perfect world, we would measure the initial state of the system, run the model forwards in time, and the predictions would agree with our observations of the system. For example, suppose we have a sequence of observations $\{y_n\}_{n\in\mathbb{N}}$. Starting from some initial condition u_0, we run the model forward from to obtain predictions $\{u_n\}_{n\in\mathbb{N}}$ of the state of the system at time n. Ideally, u_n would "agree" with y_n.

The fundamental question is thus: how do we reconcile u_n and y_n when they do not agree? There are two main reasons why discrepancies might arise:

- *observation error*: we usually cannot observe the system exactly, but rather have to contend with noisy data;
- *model error*: our model might not be the correct model for the system we are observing.

Our goal is to *estimate* u_n given $\{y_j\}_{j=1}^n$, and only uncertain knowledge of the initial condition u_0. In doing so we get a better prediction of the state of the system at time n than merely running the model forward from u_0, since we incorporate the observations up to time n; this enables us to use this estimate of u_n to run the model forward and produce better predictions for the future as well.

The paradigm example of data assimilation is weather forecasting: we have many complex models to predict the state of the world's weather, but we cannot know the exact state of the system at any one time; we thus have to contend with an initial condition u_0 which is unknown, and only partial observations y_n at each timestep. What we want is to predict the weather at time $n+1$ given the observations up to time n; we do so by *estimating* the current state of the system u_n based on the observations, and then run the model forward to obtain a prediction of u_{n+1}.

Indeed, the particular method of data assimilation which we use here, known as 3DVar, originated in weather forecasting. The technique of 3DVar (though not originally called that) was first proposed by Andrew Lorenc at the UK Met Office in 1986 [**9**], and was picked up by the National Oceanic and Atmospheric Administration (NOAA) in the US in an article by Parrish and Derber in 1992 [**12**]. More details of the implementation of 3DVar by the Met Office can be found in [**10**], and by the European Centre for Medium-Range Weather Forecasts (ECMWF) in [**3**].

Although data assimilation is widely used in practice, much of the analysis of it, in the context of fluid mechanics, works with finite-dimensional dynamical models. Our aim is to work directly with a PDE model, the Navier–Stokes equations, and thereby confront the high-dimensional nature of the problem head-on. There has been little research of this type for the Navier–Stokes equations, an exception being the recent preprint of Hayden *et al.* [**5**]; we hope to fill this void.

In section 2, we formally introduce 3DVar and the requisite background. In section 3, we analyse 3DVar for both the (linear) Stokes equations and the (non-linear) Navier–Stokes equations; we prove the following results in certain cases:

- that different estimators with the *same* observations converge geometrically to the same limit as $n \to \infty$, and

- that estimators come within bounded distance of the *true* solution, provided the observation noise is uniformly bounded.

For both the Stokes and the Navier–Stokes equations, we consider two choices of the covariance operators, given by the parameter $\alpha = \pm 1$. Here, α controls the weighting of our prediction on the observations and the model: for $\alpha < 0$, the observations have more weight than the model, while for $\alpha > 0$ the model has more weight than the observations. We prove results for both $\alpha = \pm 1$ for the Stokes equations, and go on prove results for the Navier–Stokes equations in the case $\alpha = -1$. In section 4, we report on our numerical simulations, in which we confirm numerically the analytic results in the case $\alpha = -1$, and give numerical evidence that the theorems should also hold in the case $\alpha = 1$. Finally, in section 5 we give some conclusions and indicate directions for further research.

2 Background

2.1 Notation and Setup

At the outset, let us fix the notation and set everything in context for the rest of the report. As in the survey paper [**1**], we largely follow the unified notation of [**8**]. For clarity, we write $\mathbb{N}$ for the natural numbers $\{1, 2, 3, \dots\}$, and $\mathbb{Z}^+ := \mathbb{N} \cup \{0\}$ for the non-negative integers $\{0, 1, 2, 3, \dots\}$.

Let X and Y be Hilbert spaces with inner products $\langle \cdot, \cdot \rangle$ and norms $\| \cdot \|$. If L is a self-adjoint, positive-definite operator on either X or Y, we define

$$\langle \cdot, \cdot \rangle_L := \langle \cdot, L^{-1} \cdot \rangle, \qquad \| \cdot \|_L := \| L^{-1/2} \cdot \|.$$

X will represent the space in which the model takes values, and Y will represent the space in which we observe the data.

Let $\Psi \colon X \to X$ be Lipschitz on bounded sets, and let $\mathcal{H} \colon X \to Y$ be a bounded linear operator. We define $\{u_n\}_{n \in \mathbb{Z}^+}$, $u_n \in X$ by

$$u_{n+1} := \Psi(u_n). \tag{2.1}$$

Thus, Ψ is the operator which evolves the system at time n to the system at time $n + 1$. Note that, throughout this report, we consider the case where there is *no* model error.

Correspondingly, let $\{\bar{\xi}_n\}_{n \in \mathbb{N}}$ be a sequence of independent and identically distributed Y-valued random variables: this will represent the noise corresponding to observation error. We define $\{y_n\}_{n \in \mathbb{N}}$, $y_n \in Y$ by

$$y_{n+1} := \mathcal{H} u_{n+1} + \bar{\xi}_{n+1} \tag{2.2}$$

for $n \in \mathbb{Z}^+$. Thus, $\mathcal{H}$ is the observation operator.

Goal: *Estimate u_n, given $\{y_j\}_{j=1}^n$, and only uncertain knowledge of u_0.*

There are many different approaches to data assimilation, one very useful approach being the Bayesian approach, whereby each u_n and y_n is considered not as a single point of X or Y, but rather a probability distribution. We will not go into detail here, but many details can be found in the survey paper of Stuart [**14**].

2.2 3DVar

Here, we consider the 3DVar approach to data assimilation. Let us denote by $\hat{u}_n$ our estimate of u_n. Given $\hat{u}_n$ we define $\hat{u}_{n+1}$ by

$$\hat{u}_{n+1} := \arg\min_{u \in X} J(u),$$

where

$$J(u) := \frac{1}{2}\|\mathcal{H}u - y_{n+1}\|_\Gamma^2 + \frac{1}{2}\|u - \Psi(\hat{u}_n)\|_C^2.$$

(Recall that arg min returns the argument that minimises the function, rather than the minimum value itself.) Note that the first term is a norm in Y, and the second term is a norm in X. For example, if $\Gamma = \varepsilon^2 I$ and $C = \delta^2 I$, then we have

$$J(u) = \frac{1}{2\varepsilon^2}\|\mathcal{H}u - y_{n+1}\|^2 + \frac{1}{2\delta^2}\|u - \Psi(\hat{u}_n)\|^2;$$

minimising this is then equivalent to minimising

$$\frac{1}{2}\|\mathcal{H}u - y_{n+1}\|^2 + \frac{\varepsilon^2}{2\delta^2}\|u - \Psi(\hat{u}_n)\|^2;$$

so we see that in this case the only thing that matters is the ratio of ε to δ.

A statistical interpretation of the above equation is as follows: we assign a prior distribution on u_{n+1} to be a Normal distribution with mean $\Psi(\hat{u}_n)$ and variance-covariance matrix C. We receive data of the form $y_{n+1} = \mathcal{H}u_{n+1} + \bar{\xi}_{n+1}$, where $\bar{\xi}_{n+1}$ has a Normal distribution with mean 0 and variance-covariance matrix Γ. In light of the data, our likelihood function is

$$y_{n+1}|u_{n+1} \sim N(\mathcal{H}u_{n+1}, \Gamma).$$

We are interested in the posteriori distribution $\mathbb{P}(u_{n+1}|y_{n+1})$ which we can relate to the prior and the likelihood via Bayes' rule:

$$\mathbb{P}(u_{n+1}|y_{n+1}) \propto \mathbb{P}(y_{n+1}|u_{n+1})\mathbb{P}(u_{n+1}),$$

hence

$$\begin{aligned}\mathbb{P}(u_{n+1}|y_{n+1}) &\propto \exp\left(-\frac{1}{2}\|\Gamma^{-1/2}(\mathcal{H}u_{n+1} - y_{n+1})\|^2\right)\exp\left(-\frac{1}{2}\|C^{-1/2}(u_{n+1} - \Psi(\hat{u}_n))\|^2\right)\\ &= \exp(-J(u_{n+1})).\end{aligned}$$

Because J is quadratic, the resulting posterior distribution is Gaussian, and the posterior mean minimises J. (Strictly speaking, the preceding analysis is finite dimensional, but it can be generalised to our Hilbert space setting using ideas described in Stuart [**14**].

The functional choice and our choice of Γ and C encodes our belief that the estimators should be guided by the observations (the $\|\mathcal{H}u - y_{n+1}\|_\Gamma$ term) or the model (the $\|u - \Psi(\hat{u}_{n+1})\|_\Gamma$ term).

In section 3 we consider two natural choices of C and Γ applied to two fundamental models of 2D imcompressible fluid mechanics: the Stokes equations and the Navier–Stokes equations. We prove analytic results for two of the cases mentioned.

First we prove the following that shows us how to update $\hat{u}_{n+1}$ given $\hat{u}_n$ and the observation y_{n+1}.

Proposition 2.1. *Let* $\hat{u}_{n+1} = \arg\min_{u\in X} J(u)$*, where*

$$J(u) := \frac{1}{2}\|\mathcal{H}u - y_{n+1}\|_\Gamma^2 + \frac{1}{2}\|u - \Psi(\hat{u}_n)\|_C^2.$$

Then $\hat{u}_{n+1}$ *is obtained as the solution to*

$$(C^{-1} + \mathcal{H}^*\Gamma^{-1}\mathcal{H})\hat{u}_{n+1} = C^{-1}\Psi(\hat{u}_n) + \mathcal{H}^*\Gamma^{-1}y_{n+1}.$$

Note that J is well-defined if y_{n+1} and $\Psi(\hat{u}_n)$ are in Image(Γ) and Image(C) respectively.

Proof. Let $\delta J(u)$ denote the variation of first order of $J(\cdot)$; that is

$$\langle \delta J(u), w\rangle := \lim_{t\to 0} \frac{J(u+tw) - J(u)}{t},$$

where $w \in X$ is arbitrary. As $J(\cdot)$ is a quadratic form with positive coefficients, there exists a unique $u \in X$ such that $\langle \delta J(u), w\rangle = 0$ for all $w \in X$; such u is then the minimiser $\hat{u}_{n+1}$ (see [**6**], theorem 1.43). To this end for $u \in X$, $w \in X$ arbitrary and $t \neq 0$, after expanding out we obtain

$$\begin{aligned} J(u+tw) - J(u) = \frac{1}{2}t^2\|\Gamma^{-1/2}\mathcal{H}w\| + t\langle \Gamma^{-1/2}\mathcal{H}w, \Gamma^{-1/2}(\mathcal{H}u - y_{n+1})\rangle \\ + t\langle C^{-1/2}w, C^{-1/2}(u - \Psi(\hat{u}_n))\rangle + \frac{1}{2}t^2\|C^{-1/2}w\|. \end{aligned}$$

On dividing the above by $t \neq 0$ and sending $t \to 0$ and equating the result to zero we have

$$\langle \Gamma^{-1/2}\mathcal{H}w, \Gamma^{-1/2}(\mathcal{H}u - y_{n+1})\rangle + \langle C^{-1/2}w, C^{-1/2}(u - \Psi(\hat{u}_n))\rangle = 0.$$

Hence, since Γ and C are positive definite and self-adjoint (so that $\Gamma^* = \Gamma$ and $C^* = C$), we see that

$$\langle w, \mathcal{H}^*\Gamma^{-1}(\mathcal{H}u - y_{n+1}) + C^{-1}(u - \Psi(\hat{u}_n))\rangle = 0.$$

Since w was arbitrary, $\mathcal{H}^*\Gamma^{-1}(\mathcal{H}u - y_{n+1}) + C^{-1}(u - \Psi(\hat{u}_n)) = 0$; rearranging yields

$$(C^{-1} + \mathcal{H}^*\Gamma^{-1}\mathcal{H})u = C^{-1}\Psi(\hat{u}_n) + \mathcal{H}^*\Gamma^{-1}y_{n+1}.$$

Noting that as such u is unique and is a minimiser we have

$$(C^{-1} + \mathcal{H}^*\Gamma^{-1}\mathcal{H})\hat{u}_{n+1} = C^{-1}\Psi(\hat{u}_n) + \mathcal{H}^*\Gamma^{-1}y_{n+1}$$

as required. □

3 Analysis of 3DVar

In this section, we consider the analytic properties of 3DVar, in two cases: in section 3.1 we consider the Stokes equations (the linearised form of the Navier–Stokes equations), before moving on to the full-blown Navier–Stokes equations in section 3.2.

3.1 The Stokes problem

3.1.1 Forward Model

We consider the Stokes equations (the linear version of the Navier–Stokes equations) on the torus $\mathbb{T}^2 := \mathbb{R}^2/\mathbb{Z}^2 = [0,1) \times [0,1)$ with periodic boundary conditions:

$$\partial_t u - \nu \Delta u + \nabla p = f \qquad \text{for all } (x,t) \in \mathbb{T}^2 \times (0,\infty) \tag{3.1a}$$

$$\nabla \cdot u = 0 \qquad \text{for all } (x,t) \in \mathbb{T}^2 \times (0,\infty) \tag{3.1b}$$

$$u(x,0) = u_0(x) \qquad \text{for all } x \in \mathbb{T}^2 \tag{3.1c}$$

Here $u\colon \mathbb{T}^2 \times (0,\infty) \to \mathbb{R}^2$ is a time-dependent vector field representing the velocity, $p\colon \mathbb{T}^2 \times (0,\infty) \to \mathbb{R}$ is a time-dependent scalar field representing the pressure, $f\colon \mathbb{T}^2 \to \mathbb{R}^2$ is a time-*independent* vector field representing the forcing, and ν^{-1} is the Reynolds number. In numerical simulations (see section 4), we typically represent the solution not by the velocity field u but by the stream function ζ; they are related by $u = \nabla^\perp \zeta$, where $\nabla^\perp = (\partial_2, -\partial_1)^{\mathrm{T}}$.

We approach the equations using Fourier analysis. Consider the orthonormal basis of $\{u \in L^2(\mathbb{T}^2) : \int_{\mathbb{T}^2} u \,\mathrm{d}x = 0\}$ given by $\phi_k\colon \mathbb{R}^2 \to \mathbb{R}$, $\phi_k(x) := \exp(2\pi i k \cdot x)$ for $k \in \mathbb{Z}^2 \setminus 0$. (We abuse notation so that $\mathbb{Z}^2 \setminus 0 = \mathbb{Z}^2 \setminus \{(0,0)\}$.) In order to simplify the equations, we reduce the equations by only considering divergence-free u in the following space:

$$H := \left\{ u \in (L^2(\mathbb{T}^2))^2 : \nabla \cdot u = 0, \int_{\mathbb{T}^2} u(x) \,\mathrm{d}x = 0 \right\}.$$

Given $k = (k_1, k_2)^{\mathrm{T}}$, define $k^\perp := (k_2, -k_1)^{\mathrm{T}}$. Then an orthonormal basis of H is given by $\psi_k\colon \mathbb{R}^2 \to \mathbb{R}^2$, $\psi_k(x) := \frac{k^\perp}{|k|} \exp(2\pi i k \cdot x)$ for $k \in \mathbb{Z}^2 \setminus 0$; these basis functions are in H since

$$\nabla \cdot \psi_k(x) = \frac{\partial \psi_k}{\partial x_1} + \frac{\partial \psi_k}{\partial x_2} = \frac{1}{|k|}(k_2(2\pi i k_1) - k_1(2\pi i k_2)) = 0$$

and

$$\langle \psi_j, \psi_k \rangle = \int_{\mathbb{T}^2} \frac{j^\perp}{|j|} \frac{k^\perp}{|k|} \exp(2\pi i j \cdot x) \overline{\exp(2\pi i k \cdot x)} \,\mathrm{d}x = \delta_{jk}.$$

Any vector $g \in \mathbb{C}^2$ can be decomposed uniquely into $g = g^\perp \frac{k^\perp}{|k|} + g^\parallel \frac{k}{|k|}$, where $g^\perp = \frac{1}{|k|}\langle g, k^\perp \rangle$ and $g^\parallel = \frac{1}{|k|}\langle g, k \rangle$. We define the projection operator P from $\left\{u \in (L^2(\mathbb{T}^2))^2 : \int_{\mathbb{T}^2} u(x) \,\mathrm{d}x = 0\right\}$ into H as follows: for any $g \in \left\{u \in (L^2(\mathbb{T}^2))^2 : \int_{\mathbb{T}^2} u(x) \,\mathrm{d}x = 0\right\}$, we define

$$Pg := P \sum_{k \in \mathbb{Z}^2 \setminus 0} g_k \phi_k(x) = \sum_{k \in \mathbb{Z}^2 \setminus 0} g_k^\perp \psi_k(x).$$

Writing

$$u = \sum_{k \in \mathbb{Z}^2 \setminus 0} u_k(t) \psi_k(x), \qquad p = \sum_{k \in \mathbb{Z}^2 \setminus 0} p_k(t) \phi_k(x), \qquad f = \sum_{k \in \mathbb{Z}^2 \setminus 0} f_k(t) \phi_k(x),$$

we see that, since u, p, f are real functions, we have the following reality constraints: $u_{-k} = -\overline{u_k}$, $p_{-k} = \overline{p_k}$ and $f_{-k} = \overline{f_k}$. Using the Fourier decomposition of u, we define the *fractional Sobolev spaces*:

$$H^s := \left\{ u \in H : \sum_{k \in \mathbb{Z}^2 \setminus 0} (4\pi^2 |k|^2)^s |u_k|^2 < \infty \right\}$$

with the norm $\|u\|_s := \left(\sum_k (4\pi^2|k|^2)^s |u_k|^2\right)^{1/2}$, where $s \in \mathbb{R}$. We use the notation $\|u\|$ for the standard norm on H; note that $\|u\| = \|u\|_0$.

Now, for $u \in (H^2(\mathbb{T}^2))^2$, we see that

$$\Delta u = \sum_{k \in \mathbb{Z}^2 \setminus 0} u_k(t) \Delta \psi_k(x) = \sum_{k \in \mathbb{Z}^2 \setminus 0} -4\pi^2 |k|^2 u_k(t) \psi_k(x).$$

Thus, substituting into the Stokes equation gives

$$\sum_{k \in \mathbb{Z}^2 \setminus 0} \frac{\mathrm{d}u_k}{\mathrm{d}t} \psi_k - \nu \sum_{k \in \mathbb{Z}^2 \setminus 0} u_k \Delta \psi_k + \sum_{k \in \mathbb{Z}^2 \setminus 0} p_k \nabla \phi_k = \sum_{k \in \mathbb{Z}^2 \setminus 0} f_k \phi_k.$$

Note that $\nabla \phi_k = \nabla(\exp(2\pi i k \cdot x)) = 2\pi i k \phi_k$, so for a fixed $k \in \mathbb{Z}^2 \setminus 0$ we observe that

$$\frac{\mathrm{d}u_k}{\mathrm{d}t} \psi_k - \nu u_k \Delta \psi_k + p_k 2\pi i |k| \frac{k}{|k|} \phi_k = f_k^\perp \psi_k + f_k^\parallel \frac{k}{|k|} \phi_k.$$

We now apply the projection P to both sides. Noting that the ψ_k are divergence-free, we see that $P \frac{\mathrm{d}u_k}{\mathrm{d}t} \phi_k = \frac{\mathrm{d}u_k}{\mathrm{d}t} \psi_k$, $P f_k^\perp \psi_k = f_k^\perp \psi_k$, and $P(\nu u_k \Delta \psi_k) = \nu u_k (-A\psi_k)$. In addition,

$$P\left(f_k^\parallel \frac{k}{|k|} \phi_k - p_k 2\pi i k \phi_k \right) = 0,$$

since the orthogonal component is zero. As a consequence we obtain $p_k = \frac{1}{2\pi i |k|} f_k^\parallel$, so the non-divergence-free part of the forcing is "absorbed" by the pressure. As such, we are left with the following ODE:

$$\frac{\mathrm{d}u_k}{\mathrm{d}t} + 4\pi^2 \nu |k|^2 u_k = f_k^\perp. \tag{3.2}$$

Thus, for each $k \in \mathbb{Z}^2 \setminus 0$ we have

$$u_k(t) = e^{-4\pi^2 \nu |k|^2 t} u_k(0) + \int_0^t e^{-4\pi^2 \nu |k|^2 (t-s)} f_k^\perp \, \mathrm{d}s.$$

Using the projection, we define the *Stokes operator*: for u in the space

$$\left\{ u \in (H^2(\mathbb{T}^2))^2 : \int_{\mathbb{T}^2} u(x) \, \mathrm{d}x = 0 \right\} \subset H,$$

we define the Stokes operator A by $Au := -P\Delta u$. Notice that $A\psi_k = 4\pi^2 |k|^2 \psi_k$. Thus, with abuse of notation for f,

$$u(t) = e^{-\nu A t} u_0 + \int_0^t e^{-\nu A (t-s)} f \, \mathrm{d}s,$$

and $u_0 \in H^s$ where $0 \le s \le 1$.

For the rest of this section, fix a time-step $h > 0$ and a forcing $f \in H$ which is independent of time. Let $\Psi \colon H^s \to H^s$ be the solution operator which takes an initial condition $u_0 \in H$ and returns the solution of the Stokes equations at time h; i.e.,

$$\Psi(u_0) := e^{-\nu A h} u_0 + \int_0^h e^{-\nu A (h-s)} f \, \mathrm{d}s.$$

We must check that Ψ is well-defined:

Proposition 3.1. *The map $\Psi \colon H^s \to H^s$ is well defined.*

Proof. It suffices to prove that if $u_0 \in H^s$ then the solution $u(h) \in H^s$. Since the smallest eigenvalue of the Stokes operator A is $4\pi^2$, we observe that $\|e^{-\nu Ah}\|_{\mathcal{L}(H^s,H^s)} = e^{-4\pi^2\nu h} < 1$, and so we have that

$$\|u(h)\|_s^2 < 2\|u_0\|_s^2 + 2h\|f\|^2 < \infty$$

by assumption, since $f \in H$. □

We now show that Ψ is Lipschitz with Lipschitz constant less than 1, and is thus a contraction.

Proposition 3.2. *Fix $s \in [0,1]$. There exists a constant $K = e^{-4\pi^2\nu h}$ such that, for any two initial conditions $u_0, v_0 \in H^s$, we have that*

$$\|\Psi(u_0) - \Psi(v_0)\|_s \leq K\|u_0 - v_0\|_s.$$

Proof. Since we have an analytic formula for the solution, direct calculation reveals that $\Psi(u_0) - \Psi(v_0) = e^{-\nu Ah}(u_0 - v_0)$, and so taking the H^s-norm one has

$$\|\Psi(u_0) - \Psi(v_0)\|_s \leq \|e^{-\nu Ah}\|_{\mathcal{L}(H^s,H^s)}\|u_0 - v_0\|_s = e^{-4\pi^2\nu h}\|u_0 - v_0\|_s,$$

since the smallest eigenvalue of the Stokes operator is $4\pi^2$. This completes the proof. □

We have thus defined Ψ as in equation (2.1), which takes us from one time-step to the next; that is, given an initial condition u_0 we define the sequence u_n inductively by $u_{n+1} := \Psi(u_n)$.

3.1.2 3DVar

Throughout this report, for simplicity, we will consider the observation operator $\mathcal{H} = I$, the identity on H. As in section 2.2, we seek the minimiser $\hat{u}_{n+1}$ of

$$(C^{-1} + \Gamma^{-1})\hat{u}_{n+1} = C^{-1}\Psi(\hat{u}_n) + \Gamma^{-1}y_{n+1},$$

We consider the case when C^{-1} and Γ^{-1} are both powers of the Stokes operator A, i.e. $C^{-1} = \frac{1}{\delta^2}A^\gamma$ and $\Gamma^{-1} = \frac{1}{\varepsilon^2}A^\beta$. Substituting this in yields

$$\left(\frac{1}{\delta^2}A^\gamma + \frac{1}{\varepsilon^2}A^\beta\right)\hat{u}_{n+1} = \frac{1}{\delta^2}A^\gamma\Psi(\hat{u}_n) + \frac{1}{\varepsilon^2}A^\beta y_{n+1}.$$

Multiplying through by $\varepsilon^2 A^{-\beta}$, we see that

$$\left(\frac{\varepsilon^2}{\delta^2}A^{\gamma-\beta} + I\right)\hat{u}_{n+1} = \frac{\varepsilon^2}{\delta^2}A^{\gamma-\beta}\Psi(\hat{u}_n) + y_{n+1}.$$

Define $\eta = \varepsilon/\delta$ and set $\alpha = \gamma - \beta$. Then we have

$$\hat{u}_{n+1} = \eta^2(I + \eta^2 A^\alpha)^{-1}A^\alpha\Psi(\hat{u}_n) + (I + \eta^2 A^\alpha)^{-1}y_{n+1}.$$

Since A only appears to the power of α, we can consider all possible cases of C^{-1} and Γ^{-1} just by choosing an appropriate value of α. We will consider two main cases:

$$\alpha = 1: \quad \text{equivalent to } C^{-1} = \frac{1}{\delta^2}A, \quad \Gamma^{-1} = \frac{1}{\varepsilon^2}I;$$
$$\alpha = -1: \quad \text{equivalent to } C^{-1} = \frac{1}{\delta^2}I, \quad \Gamma^{-1} = \frac{1}{\varepsilon^2}A.$$

Before we proceed to analysing the cases $\alpha = \pm 1$, we prove some preliminary results.

We now show that the inverse map $(\sigma I + \tau A)^{-1}$ exists for any $\sigma, \tau > 0$, and hence that we can indeed apply proposition 2.1.

Proposition 3.3. *Fix $s \in [0,1]$. For any $\sigma, \tau > 0$, the inverse map $(\sigma I + \tau A)^{-1} \colon H^s \to H^{s+2} \subset H^s$ exists, and satisfies*

$$\|(\sigma I + \tau A)^{-1}\|_{\mathcal{L}(H^s, H^s)} \leq \frac{1}{\sigma + 4\pi^2 \tau}.$$

Proof. Define the symbol of a differential operator $\mathcal{P} = \sum_{|\alpha| \leq l} a_\alpha D^\alpha$ by $\mathrm{Sym}(\mathcal{P})(k) = \sum_{|\alpha| \leq l} a_\alpha k^\alpha$. Then we have

$$\mathrm{Sym}(\sigma I + \tau A)(k) = \mathrm{Sym}(\sigma I - \tau \Delta)(k) = \sigma + 4\pi^2 \tau |k|^2 \geq \sigma + 4\pi^2 \tau > 0,$$

since $|k|^2 > 1$ for every $k \in \mathbb{Z}^2 \setminus 0$. Hence $(\sigma I + \tau A)^{-1}$ exists by the theory of pseudo-differential operators (cf. the book of Hörmander [**7**]), and

$$\mathrm{Sym}((\sigma I + \tau A)^{-1})(k) = \frac{1}{\sigma + 4\pi^2 \tau |k|^2} \leq \frac{1}{\sigma + 4\pi^2 \tau}. \qquad \square$$

Corollary 3.4. *Fix $s \in [0,1]$. For any $\sigma, \tau > 0$, the map $(\sigma I + \tau A)^{-1} \tau A \colon H^s \to H^s$ is a bounded linear operator, and satisfies*

$$\|(\sigma I + \tau A)^{-1} \tau A\|_{\mathcal{L}(H^s, H^s)} \leq 1.$$

Proof. It is clear that $(\sigma I + \tau A)^{-1} \tau A$ is linear, and from the previous proposition we see that

$$\mathrm{Sym}((\sigma I + \tau A)^{-1} \tau A)(k) = \frac{4\pi^2 |k|^2 \tau}{\sigma + 4\pi^2 |k|^2 \tau} \leq 1$$

for all $k \in \mathbb{Z}^2 \setminus 0$, as required. $\square$

3.1.3 Analysis of cases

In this section, we will consider the analysis of cases $\alpha = 1$ and $\alpha = -1$. Note that everything in this section is true for every $\alpha \in \mathbb{R}$; we do not present the proof here, but only consider the cases of $\alpha = 1$ and $\alpha = -1$ as an exemplar of the analytical techniques needed.

Case 1: $\alpha = 1$ Recalling proposition 2.1 we have that the estimator at time $(n+1)h$ is given by

$$\hat{u}_{n+1} = \eta^2 (I + \eta^2 A)^{-1} A \Psi(\hat{u}_n) + (I + \eta^2 A)^{-1} y_{n+1}.$$

Propositions 3.2 and 3.3 and corollary 3.4 provide all the necessary theory to prove that if we start with two different initial estimators $\hat{u}_0, \hat{v}_0$ but use the *same* observations y_n for every $n \in \mathbb{N}$, then the estimators converge exponentially to a common limit:

Theorem 3.5. *Fix $s \in [0,1]$. Let $K = e^{-4\pi^2 \nu h} < 1$ as in proposition 3.2. Given initial estimators $\hat{u}_0, \hat{v}_0 \in H^s$, and observations $y_n \in H^s$ for every $n \in \mathbb{N}$, we have that*

$$\|\hat{u}_{n+1} - \hat{v}_{n+1}\|_s \leq K^{n+1} \|\hat{u}_0 - \hat{v}_0\|_s$$

for every $n \in \mathbb{Z}^+$.

Proof. We first show by induction that $\hat{u}_n \in H^s$ for all $n \in \mathbb{Z}^+$. Suppose that $\hat{u}_n \in H^s$. By proposition 3.3 and corollary 3.4, we have that

$$\begin{aligned} \|\hat{u}_{n+1}\|_s &\leq \|\eta^2 (I + \eta^2 A)^{-1} A \Psi(\hat{u}_n)\|_s + \|(I + \eta^2 A)^{-1} y_{n+1}\|_s \\ &\leq \|\hat{u}_n\|_s + \frac{1}{1 + 4\pi^2 \eta^2} \|y_{n+1}\|_s < \infty. \end{aligned}$$

Hence $\hat{u}_{n+1} \in H^s$, as required. By proposition 2.1, subtracting and taking the H^s-norm we have

$$\|\hat{u}_{n+1} - \hat{v}_{n+1}\|_s = \|\eta^2(I + \eta^2 A)^{-1} A[\Psi(\hat{u}_n) - \Psi(\hat{v}_n)]\|_s.$$

By corollary 3.4 and proposition 3.2, we see that

$$\|\hat{u}_{n+1} - \hat{v}_{n+1}\|_s \leq \|\Psi(\hat{u}_n) - \Psi(\hat{v}_n)\|_s \leq K\|\hat{u}_n - \hat{v}_n\|_s,$$

where $K = e^{-4\pi^2\nu h}$ as in proposition 3.2. Hence

$$\|\hat{u}_{n+1} - \hat{v}_{n+1}\| \leq K^{n+1}\|\hat{u}_0 - \hat{v}_0\|_s. \qquad \square$$

We proceed to prove the following estimate on the distance between the true solution and the estimator. We define $e_n := \|\hat{u}_n - u_n\|_s$.

Theorem 3.6. *Fix $s \in [0, 1]$. Suppose that $u_0, \hat{u}_0 \in H^s$ are the true and estimated initial condition respectively. Suppose further that the noise $\bar{\xi}_n$ is uniformly bounded in H^s, that is $M := \sup \|\bar{\xi}_n\|_s < \infty$. Then*

$$e_{n+1} \leq K^{n+1} e_0 + \frac{M}{(1 - K)(1 + 4\pi^2\eta^2)},$$

where $K = e^{-4\pi^2\nu h}$ as in proposition 3.2.

Proof. Recall that

$$u_{n+1} = \Psi(u_n) = (\eta^2 A + I)^{-1}(\eta^2 A + I)\Psi(u_n).$$

By proposition 2.1,

$$\hat{u}_{n+1} = (\eta^2 A + I)^{-1}\eta^2 A\Psi(\hat{u}_n) + (\eta^2 A + I)^{-1} y_{n+1}.$$

Subtracting, taking the H^s-norm and recalling that $y_{n+1} = \Psi(u_n) + \bar{\xi}_{n+1}$ we have

$$\|\hat{u}_{n+1} - u_{n+1}\|_s \leq \|\Psi(\hat{u}_n) - \Psi(u_n)\|_s + \|(\eta^2 A + I)^{-1}\bar{\xi}_{n+1}\|_s,$$

since $\|(\eta^2 A + I)^{-1}\eta^2 A\|_{\mathcal{L}(H^s, H^s)} \leq 1$, by corollary 3.4. By proposition 3.3, $\|(\eta^2 A + I)^{-1}\|_{\mathcal{L}(H^s, H^s)} \leq \frac{1}{1+4\pi^2\eta^2}$; so, setting $K = e^{-4\pi^2\nu h}$ and using proposition 3.2 yields

$$\|\hat{u}_{n+1} - u_{n+1}\|_s \leq K\|\hat{u}_n - u_n\|_s + \frac{M}{1 + 4\pi^2\eta^2}.$$

Iterating yields

$$\|\hat{u}_{n+1} - u_{n+1}\|_s \leq K^{n+1}\|\hat{u}_0 - u_0\|_s + \frac{M}{1 + 4\pi^2\eta^2}\sum_{j=0}^{n} K^j \leq K^{n+1}\|\hat{u}_0 - u_0\|_s + \frac{M}{(1 - K)(1 + 4\pi^2\eta^2)}. \qquad \square$$

Case 2: $\alpha = -1$ Here, we prove the analogue of theorem 3.5 and theorem 3.6, but with a different choice of operators C and Γ. By proposition 2.1, estimator $\hat{u}_{n+1}$ with observations y_{n+1} satisfies

$$\hat{u}_{n+1} = \eta^2(\eta^2 I + A)^{-1}\Psi(\hat{u}_n) + (\eta^2 I + A)^{-1} A y_{n+1}.$$

We now proceed to prove an analogue of theorem 3.5.

Theorem 3.7. *Fix $s \in [0,1]$. There exists a constant $C' = \frac{\eta^2 e^{-4\pi^2\nu h}}{\eta^2+4\pi^2} < 1$ such that, given initial estimators $\hat{u}_0, \hat{v}_0 \in H^s$, and observations $y_n \in H^s$ for every $n \in \mathbb{N}$, we have that*

$$\|\hat{u}_{n+1} - \hat{v}_{n+1}\|_s \leq (C')^{n+1}\|\hat{u}_0 - \hat{v}_0\|_s$$

for every $n \in \mathbb{Z}^+$.

Proof. We first show by induction that $\hat{u}_n \in H^s$ for all $n \in \mathbb{Z}^+$. Suppose that $\hat{u}_n \in H^s$. By proposition 3.3 and corollary 3.4, we have that

$$\begin{aligned}\|\hat{u}_{n+1}\|_s &\leq \|\eta^2(\eta^2 I + A)^{-1}\Psi(\hat{u}_n)\|_s + \|(\eta^2 I + A)^{-1}Ay_{n+1}\|_s \\ &\leq \frac{\eta^2}{\eta^2+4\pi^2}\|\hat{u}_n\|_s + \|y_{n+1}\|_s < \infty.\end{aligned}$$

Hence $\hat{u}_{n+1} \in H^s$, as required. By proposition 2.1, subtracting and taking the H^s-norm we have

$$\|\hat{u}_{n+1} - \hat{v}_{n+1}\|_s = \|\eta^2(\eta^2 I + A)^{-1}[\Psi(\hat{u}_n) - \Psi(\hat{v}_n)]\|_s.$$

By proposition 3.3 and proposition 3.2, we see that

$$\|\hat{u}_{n+1} - \hat{v}_{n+1}\|_s \leq \frac{\eta^2}{\eta^2+4\pi^2}\|\Psi(\hat{u}_n) - \Psi(\hat{v}_n)\|_s \leq \frac{C\eta^2}{\eta^2+4\pi^2}\|\hat{u}_n - \hat{v}_n\|_s,$$

where $C = e^{-4\pi^2\nu h}$ as in proposition 3.2. Setting $C' = \frac{C\eta^2}{\eta^2+4\pi^2}$, we see that

$$\|\hat{u}_{n+1} - \hat{v}_{n+1}\| \leq (C')^{n+1}\|\hat{u}_0 - \hat{v}_0\|_s. \qquad \square$$

Finally, we prove the analogue of theorem 3.6.

Theorem 3.8. *Fix $s \in [0,1]$. Suppose that $u_0, \hat{u}_0 \in H^s$ are the true and estimated initial condition respectively. Suppose further that the noise $\bar{\xi}_n$ is uniformly bounded in H^s, that is $M := \sup\|\bar{\xi}_n\|_s < \infty$. Then*

$$e_{n+1} \leq K^{n+1}e_0 + \frac{M}{1-K},$$

where $K = e^{-4\pi^2\nu h}$ as in proposition 3.2.

Proof. Following in a similar fashion to the proof of theorem 3.6 one has

$$\|\hat{u}_{n+1} - u_{n+1}\|_s \leq \|\Psi(\hat{u}_n) - \Psi(u_n)\|_s + \|\bar{\xi}_{n+1}\|_s,$$

noting that $\|(\eta^2 I + A)^{-1}A\|_{\mathcal{L}(H^s,H^s)}$, $\|\eta^2(\eta^2 I + A)^{-1}\|_{\mathcal{L}(H^s,H^s)} \leq 1$ by proposition 3.3 and corollary 3.4. Setting $K = e^{-4\pi^2\nu h}$ and using proposition 3.2 yields

$$\|\hat{u}_{n+1} - u_{n+1}\|_s \leq K\|\hat{u}_n - u_n\|_s + M,$$

so iterating yields

$$\|\hat{u}_{n+1} - u_{n+1}\|_s \leq K^{n+1}\|\hat{u}_0 - u_0\|_s + M\sum_{j=0}^{n} K^j \leq K^{n+1}\|\hat{u}_0 - u_0\|_s + \frac{M}{1-K}. \qquad \square$$

In the above cases, we used the fact that we have an analytic solution to the Stokes equations:

$$u(t) = e^{-\nu At}u_0 + \int_0^t e^{-\nu A(t-s)}f\,\mathrm{d}s$$

and $\|e^{-\nu Ah}\|_{\mathcal{L}(H^s,H^s)} = e^{-4\pi^2\nu h} < 1$. This made the analysis easy and so we could show that estimators of the same observation agree if n is suitably large and we gave a bound on the distance from the true solution and the estimator. We also see that if the noise is switched off, then we get agreement of the true solution and of the estimator.

3.2 The Navier–Stokes equations

In this section we prove analogous theorems for the Navier–Stokes equations for the case $\alpha = -1$; the case $\alpha = 1$ is made significantly harder by the nonlinear term in the Navier-Stokes equations and so we have omitted this case (though we do so reluctantly, and not for want of trying!).

Here, unlike in the Stokes equations, there is a key difference between $\alpha = 1$ and $\alpha = -1$. In the case $\alpha = 1$, the operator $\eta^2(I + \eta^2 A)^{-1}A$ in front of the Ψ term has operator norm 1, while in the case $\alpha = -1$, the operator $\eta^2(\eta^2 I + A)^{-1}$ in front of the Ψ term has operator norm strictly less than 1. While in the Stokes equations we have that Ψ is a contraction, for the Navier–Stokes equations we only have that Ψ is Lipschitz, thus making the analysis a lot harder.

3.2.1 The Forward Model

We now consider the full nonlinear Navier–Stokes equations in two dimensions on the torus $\mathbb{T}^2$, using the same notation as in the Stokes case:

$$\partial_t u - \nu\Delta u + u \cdot \nabla u + \nabla p = f \qquad \text{for all } (x,t) \in \mathbb{T}^2 \times (0,\infty) \tag{3.3a}$$

$$\nabla \cdot u = 0 \qquad \text{for all } (x,t) \in \mathbb{T}^2 \times (0,\infty) \tag{3.3b}$$

$$u(x,0) = u_0(x) \qquad \text{for all } x \in \mathbb{T}^2 \tag{3.3c}$$

Analogous to the previous section, we consider the following ODE in H:

$$\frac{\mathrm{d}}{\mathrm{d}t}u + \nu Au + B(u,u) = f$$

with initial condition $u(0) = u_0$. Here, as in the Stokes equations, $A = -P\Delta$ and acts as negative Laplacian on its domain, while the term $B(u,u) = P(u \cdot \nabla u)$ is the bilinear form found by projecting the nonlinear term $u \cdot \nabla u$ onto H; finally, with abuse of notation f is the original forcing, projected onto H. (Full details can be found in Cotter *et al.* [**2**], and in chapter 9 of Robinson [**13**].)

We define the operator $\Psi\colon H^s \to H^s$ as the solution map that takes an initial condition u_0 and maps it to the solution of the Navier–Stokes equations at time h. In order to show that this is, indeed, well-defined (i.e. it maps H^s into H^s), we require various bounds on solutions of the Navier–Stokes equations, which are taken without proof from Cotter *et al.* [**2**].

First, we require lemma 5.1 from Cotter *et al.* [**2**]:

Lemma 3.9. *For each $s \in [0,1]$ there exists constants $c_s, c_s' > 0$, independent of t, such that:*

$$\text{for } s = 0, \qquad |\langle (u\cdot\nabla)v, w\rangle| \le c_0 \|u\|^{1/2}\|u\|_1^{1/2}\|v\|_1\|w\|^{1/2}\|w\|_1^{1/2}; \tag{3.4}$$

$$\text{for } 0 < s < 1, \qquad |\langle (u\cdot\nabla)v, A^s w\rangle| \le c_s \|u\|_s\|v\|_1\|w\|_{1+s}; \tag{3.5}$$

$$\text{for } 0 < s \le 1, \qquad |\langle (u\cdot\nabla)v, A^s w\rangle| \le c_s' \|u\|^{1/2}\|u\|_s^{1/2}\|v\|_1^{1/2}\|v\|_{1+s}^{1/2}\|w\|_{1+s}. \tag{3.6}$$

We also note the following orthogonality relations from section 5.1 of Cotter *et al.* [**2**]:

$$\langle (u\cdot\nabla)v, v\rangle = 0; \tag{3.7}$$

$$\langle (u\cdot\nabla)u, Au\rangle = 0, \tag{3.8}$$

where equation (3.8) holds as we are working in a two-dimensional periodic domain.

Finally, we take lemma 5.2 from Cotter *et al.* [**2**] (with h in place of t).

Lemma 3.10. *Fix $s \in [0,1]$. Suppose that u solves the Navier–Stokes equations with initial condition $u_0 \in H^s$ and forcing $f \in H$ independent of time. Then there exists a constant $C = C(\nu, h, s) > 0$ such that*

$$\|u(h)\|_s^2 + \nu \int_0^h \|u(\tau)\|_{s+1}^2 \,\mathrm{d}\tau \leq C(\|u_0\|_s^2 + \|f\|^2) \tag{3.9}$$

In particular, lemma 3.10 shows that Ψ maps H^s into H^s, and is thus well-defined. We use these results to prove that Ψ is, in fact, Lipschitz on bounded sets; contrast this with the case of the Stokes equations, where we proved (in proposition 3.2) that Ψ is *globally* Lipschitz.

Proposition 3.11. *Fix $s \in [0,1]$, and let $R > 0$. There exists a constant $K = K(R, f, \nu, h, s) > 0$ such that, for any two initial conditions $u_0, v_0 \in H^s$ with $\|u_0\|_s, \|v_0\|_s \leq R$, we have that There exists a constant K which depends on R but is independent of u_0, v_0 such that*

$$\|\Psi(u_0) - \Psi(v_0)\|_s \leq K\|u_0 - v_0\|_s.$$

Proof. We prove that, given two initial conditions u_0 and v_0, the solutions u and v at time h satisfy $\|u(h) - v(h)\|_s \leq K\|u_0 - v_0\|_s$, where u and v satisfy

$$\frac{\mathrm{d}}{\mathrm{d}t}u + \nu Au + B(u,u) = f,$$
$$\frac{\mathrm{d}}{\mathrm{d}t}v + \nu Av + B(v,v) = f,$$

respectively. Set $w = u - v$. A calculation shows that $B(u,u) - B(v,v) = B(u,w) + B(w,u) - B(w,w)$, so that w satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}w + \nu Aw + B(u,w) + B(w,u) - B(w,w) = 0.$$

Multiplying by w and taking the H^s-inner product we obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|w\|_s^2 + \nu\|w\|_{1+s}^2 + \langle B(u,w), A^s w\rangle + \langle B(w,u), A^s w\rangle - \langle B(w,w), A^s w\rangle = 0.$$

In the case $s = 0$, $\langle B(u,w), w\rangle = \langle B(w,w), w\rangle = 0$ by the orthogonality relation (3.7), so using (3.4) we see that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|w\|^2 + \nu\|w\|_1^2 \leq |\langle B(w,u), w\rangle| \leq c_0\|u\|_1\|w\|\|w\|_1 \leq \frac{\nu}{2}\|w\|_1^2 + \frac{c_0^2}{2\nu}\|u\|_1^2\|w\|^2$$

using Young's inequality. Hence,

$$\frac{\mathrm{d}}{\mathrm{d}t}\|w\|^2 - \frac{c_0^2}{\nu}\|u\|_1^2\|w\|^2 \leq -\nu\|w\|_1^2 \leq 0,$$

so by Grönwall's inequality we have that

$$\|w(h)\|^2 \exp\left(-\frac{c_0^2}{\nu}\int_0^h \|u(\tau)\|_1^2 \,\mathrm{d}\tau\right) - \|w(0)\|^2 \leq 0,$$

and hence, by lemma 3.10, we have

$$\begin{aligned}
\|w(h)\|^2 &\leq \exp\left(\frac{c_0^2}{\nu}\int_0^h \|u(\tau)\|_1^2 \,\mathrm{d}\tau\right)\|w(0)\|^2 \\
&\leq \exp\left(\frac{C(\|u_0\|^2 + \|f\|^2)c_0^2}{\nu^2}\right)\|w(0)\|^2 \\
&\leq \exp\left(\frac{C(R^2 + \|f\|^2)c_0^2}{\nu^2}\right)\|w(0)\|^2,
\end{aligned}$$

since $\|u_0\|, \|v_0\| \leq R$. Thus, setting $K := \sqrt{\exp\left(\frac{C(R^2+\|f\|^2)c_0^2}{\nu^2}\right)}$ we see that $\|u(h)-v(h)\| \leq K\|u_0-v_0\|$, as required.

In the case $s = 1$, $\langle B(w,w), Aw\rangle = 0$ by the orthogonality relation (3.8), so using (3.6) we see that

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|w\|_1^2 + \nu\|w\|_2^2 &\leq |\langle B(u,w), Aw\rangle| + |\langle B(w,u), Aw\rangle| \\ &\leq c_1'\|u\|^{1/2}\|u\|_1^{1/2}\|w\|_1^{1/2}\|w\|_2^{3/2} + c_1'\|u\|_1^{1/2}\|u\|_2^{1/2}\|w\|^{1/2}\|w\|_1^{1/2}\|w\|_2 \\ &\leq \frac{\nu}{4}\|w\|_2^2 + \frac{27(c_1')^4}{4\nu^3}\|u\|^2\|u\|_1^2\|w\|_1^2 + \frac{\nu}{4}\|w\|_2^2 + \frac{(c_1')^2}{\nu}\|u\|_1\|u\|_2\|w\|\|w\|_1 \\ &\leq \frac{\nu}{2}\|w\|_2^2 + \left(\frac{27(c_1')^4}{4\nu^3}\|u\|_1^4 + \frac{(c_1')^2}{\nu}\|u\|_1\|u\|_2\right)\|w\|_1^2\end{aligned}$$

where we have used Young's inequality and the fact that $\|\cdot\| \leq \|\cdot\|_1$. Rearranging, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\|w\|_1^2 - \left(\frac{27(c_1')^4}{2\nu^3}\|u\|_1^4 + \frac{2(c_1')^2}{\nu}\|u\|_1\|u\|_2\right)\|w\|_1^2 \leq -\nu\|w\|_2^2 \leq 0,$$

so by Grönwall's inequality we have

$$\|w(h)\|_1^2 \exp\left(-\frac{27(c_1')^4}{2\nu^3}\int_0^h \|u(\tau)\|_1^4\,\mathrm{d}\tau - \frac{2(c_1')^2}{\nu}\int_0^h \|u(\tau)\|_1\|u(\tau)\|_2\,\mathrm{d}\tau\right) - \|w(0)\|_1^2 \leq 0,$$

and hence, since $\|u_0\|_s, \|v_0\|_s \leq R$,

$$\begin{aligned}\|w(h)\|_1^2 &\leq \|w(0)\|_1^2 \exp\left(\frac{27(c_1')^4}{2\nu^3}\int_0^h \|u(\tau)\|_1^4\,\mathrm{d}\tau + \frac{2(c_1')^2}{\nu}\int_0^h \|u(\tau)\|_1\|u(\tau)\|_2\,\mathrm{d}\tau\right) \\ &\leq \|w(0)\|_1^2 \exp\left(\frac{27h(c_1')^4C^4(\|u_0\|_1^2+\|f\|^2)^4}{2\nu^3} + \frac{2(c_1')^2C(\|u_0\|_1^2+\|f\|^2)}{\nu}\int_0^h \|u(\tau)\|_2\,\mathrm{d}\tau\right) \\ &\leq \|w(0)\|_1^2 \exp\left(\frac{27h(c_1')^4C^4(\|u_0\|_1^2+\|f\|^2)^4}{2\nu^3} + \frac{2(c_1')^2C^2(\|u_0\|_1^2+\|f\|^2)^2}{\nu}\right) \\ &\leq \|w(0)\|_1^2 \exp\left(\frac{27h(c_1')^4C^4(R^2+\|f\|^2)^4}{2\nu^3} + \frac{2(c_1')^2C^2(R^2+\|f\|^2)^2)(c_1')^2}{\nu}\right)\end{aligned}$$

by lemma 3.10; thus, setting $K = \sqrt{\exp\left(\frac{27h(c_1')^4C^4(R^2+\|f\|^2)^4}{2\nu^3} + \frac{2(c_1')^2C^2(R^2+\|f\|^2)^2)(c_1')^2}{\nu}\right)}$ we have that $\|u(h)-v(h)\|_1 \leq K\|u_0-v_0\|_1$, as required.

Finally, in the case $s \in (0,1)$, using (3.5) we see that

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|w\|_s^2 + \nu\|w\|_{1+s}^2 &\leq |\langle B(u,w), A^s w\rangle| + |\langle B(w,u), A^s w\rangle| + |\langle B(w,w), A^s w\rangle| \\ &\leq c_s\|u\|_s\|w\|_1\|w\|_{1+s} + c_s\|u\|_1\|w\|_s\|w\|_{1+s} + c_s\|w\|_1\|w\|_s\|w\|_{1+s} \\ &\leq c_s\|u\|_s\|w\|_s^s\|w\|_{1+s}^{2-s} + c_s\|u\|_1\|w\|_s\|w\|_{1+s} + c_s\|w\|_1\|w\|_s\|w\|_{1+s}\end{aligned}$$

using the Sobolev interpolation estimate $\|w\|_1 \leq \|w\|_s^s\|w\|_{1+s}^{1-s}$. Using Young's inequality, we obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|w\|_s^2 + \nu\|w\|_{1+s}^2 \leq \frac{\nu}{2}\|w\|_{1+s}^2 + \|w\|_s^2\left[\frac{3c_s^2}{2\nu}\left(\|u\|_1^2+\|w\|_1^2\right) + \frac{s}{2}\left(\frac{\nu}{3(2-s)}\right)^{(s-2)/s} c_s^{2/s}\|u\|_s^{2/s}\right].$$

Rearranging yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\|w\|_s^2 - \|w\|_s^2\left[\frac{3c_s^2}{\nu}\left(\|u\|_1^2+\|w\|_1^2\right) + sc_s^{2/s}\left(\frac{\nu}{3(2-s)}\right)^{(s-2)/s}\|u\|_s^{2/s}\right] \leq -\nu\|w\|_{1+s}^2 \leq 0,$$

so using Grönwall's inequality we obtain

$$\begin{aligned}\|w(h)\|_s^2 &\leq \|w(0)\|_s^2 \exp\left[\frac{3c_s^2}{\nu}\int_0^h \left(\|u(\tau)\|_1^2 + \|w(\tau)\|_1^2\right)\,\mathrm{d}\tau + sc_s^{2/s}\left(\frac{\nu}{3(2-s)}\right)^{(s-2)/s}\int_0^h \|u(\tau)\|_s^{2/s}\,\mathrm{d}\tau\right] \\ &\leq \|w(0)\|_s^2 \exp\left[\frac{3Cc_s^2}{\nu}(\|u_0\|^2 + \|w_0\|^2 + 2\|f\|^2) + sh(Cc_s(\|u_0\|_s^2 + \|f\|^2))^{2/s}\left(\frac{\nu}{3(2-s)}\right)^{(s-2)/s}\right] \\ &\leq \|w(0)\|_s^2 \exp\left[\frac{3Cc_s^2}{\nu}(3R^2 + 2\|f\|^2) + sh(Cc_s(R^2 + \|f\|^2))^{2/s}\left(\frac{\nu}{3(2-s)}\right)^{(s-2)/s}\right]\end{aligned}$$

by lemma 3.10, noting that $\|\cdot\| \leq \|\cdot\|_s$ and the fact that $\|u_0\|_s, \|v_0\|_s \leq R$. Thus, upon setting

$$K := \sqrt{\exp\left[\frac{3Cc_s^2}{\nu}(3R^2 + 2\|f\|^2) + sh(Cc_s(R^2 + \|f\|^2))^{2/s}\left(\frac{\nu}{3(2-s)}\right)^{(s-2)/s}\right]},$$

we see that $\|u(h) - v(h)\|_s \leq K\|u_0 - v_0\|_s$, as required. □

3.2.2 3DVar

We proceed to prove analogous results to theorem 3.7 for the Navier–Stokes equations in the case $\alpha = -1$. From proposition 2.1 we have (with $\eta = \frac{\varepsilon}{\delta}$ as before)

$$(\eta^2 I + A)\hat{u}_{n+1} = \eta^2 I\Psi(\hat{u}_n) + Ay_{n+1}.$$

By proposition 3.3,

$$\hat{u}_{n+1} = \eta^2(\eta^2 I + A)^{-1}\Psi(\hat{u}_n) + (\eta^2 I + A)^{-1}A\, y_{n+1}.$$

We define $\Psi_\eta := \eta^2(\eta^2 I + A)^{-1}\Psi$; note that for $u \in H^s$, by proposition 3.3 we have

$$\|\Psi_\eta(u)\|_s \leq \frac{\eta^2}{\eta^2 + 4\pi^2}\|\Psi(u)\|_s.$$

We make the following assumptions:

Assumption 3.12. *Given $s \in [0,1]$, we assume that:*

- *the solution u_n is uniformly bounded in H^s: i.e. $\sup_{n\in\mathbb{Z}^+}\|u_n\|_s \leq M_1$; and*
- *the noise $\bar{\xi}_n$ is uniformly bounded in H^s: i.e. $\sup_{n\in\mathbb{N}}\|\bar{\xi}_n\|_s \leq M_2$.*

The assumption on the solutions is, in fact, physically reasonable. Bounds on the solution of the Navier–Stokes equations in H^1 can be proved using the ideas of absorbing sets; see Robinson (chap. 12, [**13**]) and Temam (chap. III, § 2, [**15**]). We use these assumptions to prove that the estimators $\hat{u}_n$ are also uniformly bounded in H^s.

Lemma 3.13. *Fix $s \in [0,1]$. Let $\hat{u}_n$ be the estimator of the truth u_n at time nh with observation $y_n \in H^s$. Suppose that assumption 3.12 holds. Given initial estimator $\hat{u}_0$, for sufficiently small $\eta > 0$, the estimators $\hat{u}_n$ are uniformly bounded in H^s: that is, there exists $R > 0$ and $\delta > 0$ such that for each $\eta < \delta$, we have $\sup_{n\in\mathbb{N}}\|\hat{u}_n\|_s \leq R$.*

Proof. Set $N = M_1 + M_2 + \|\Psi_\eta(0)\|_s$, and let $R = \|\hat{u}_0\|_s + 2N$. By proposition 3.11, Ψ is Lipschitz with Lipschitz constant $K(R)$ on the H^s-ball of radius R, so choose η small enough so that $\lambda := \frac{\eta^2}{4\pi^2+\eta^2}K(R) < \frac{1}{2}$. We prove by induction that

$$\|\hat{u}_n\|_s \leq \lambda^n \|\hat{u}_0\|_s + N\sum_{j=0}^{n-1}\lambda^j$$

for all $n \in \mathbb{N}$; this yields the desired result since, as $\lambda < \frac{1}{2}$, we have

$$\|\hat{u}_n\|_s \leq \|\hat{u}_0\|_s + \frac{N}{1-\lambda} \leq \|\hat{u}_0\|_s + 2N = R.$$

Assume that $\|\hat{u}_n\|_s \leq R$ and that $\|\hat{u}_n\|_s \leq \lambda^n\|\hat{u}_0\|_s + N\sum_{j=0}^{n-1}\lambda^j$. Then

$$\|\Psi_\eta(\hat{u}_n)\|_s \leq \|\Psi_\eta(\hat{u}_n) - \Psi_\eta(0)\|_s + \|\Psi_\eta(0)\|_s \leq \lambda\|\hat{u}_n\|_s + \|\Psi_\eta(0)\|_s,$$

by assumption on $\|\hat{u}_n\|_s$. Now, since

$$\hat{u}_{n+1} = \Psi_\eta(\hat{u}_n) + (\eta^2 I + A)^{-1}Ay_{n+1} = \Psi_\eta(\hat{u}_n) + (\eta^2 I + A)^{-1}Au_{n+1} + (\eta^2 I + A)^{-1}A\bar{\xi}_{n+1},$$

by proposition 3.3 and corollary 3.4 we have that

$$\begin{aligned}
\|\hat{u}_{n+1}\|_s &\leq \|\Psi_\eta(\hat{u}_n)\|_s + \|(\eta^2 I + A)^{-1}A\|_{\mathcal{L}(H^s,H^s)}\|u_{n+1}\|_s + \|(\eta^2 I + A)^{-1}A\|_{\mathcal{L}(H^s,H^s)}\|\bar{\xi}_{n+1}\|_s \\
&\leq \lambda\|\hat{u}_n\|_s + \|\Psi_\eta(0)\|_s + M_1 + M_2 \\
&\leq \lambda\|\hat{u}_n\|_s + N \\
&\leq \lambda\left[\lambda^n\|\hat{u}_0\|_s + N\sum_{j=0}^{n-1}\lambda^j\right] + N \\
&= \lambda^{n+1}\|\hat{u}_0\|_s + N\sum_{j=0}^{n}\lambda^j.
\end{aligned}$$

As $\lambda < \frac{1}{2}$, we have

$$\|\hat{u}_{n+1}\|_s \leq \|\hat{u}_0\|_s + \frac{N}{1-\lambda} \leq \|\hat{u}_0\|_s + 2N = R.$$

By induction on n, this completes the proof. □

We now have all the theory required to prove the following.

Theorem 3.14. *Fix $s \in [0,1]$, and assume that assumption 3.12 holds. Given initial estimators $\hat{u}_0, \hat{v}_0 \in H^s$, and observations $y_n \in H^s$ for every $n \in \mathbb{N}$, we have that, for all sufficiently small η, there exists a constant $\lambda < 1$ such that*

$$\|\hat{u}_{n+1} - \hat{v}_{n+1}\|_s \leq \lambda^{n+1}\|\hat{u}_0 - \hat{v}_0\|_s$$

for every $n \in \mathbb{Z}^+$.

Proof. By assumption 3.12, $\|u_n\|_s \leq M_1$ and $\|\bar{\xi}_n\|_s \leq M_2$ for all n. Set $R = \max\{\|\hat{u}_0\|_s, \|\hat{v}_0\|_s\} + 2(M_1 + M_2 + \|\Psi_\eta(0)\|_s)$, and choose η small enough so that $\lambda := \frac{\eta^2}{4\pi^2+\eta^2}K(R) < \frac{1}{2}$. By lemma 3.13, $\|\hat{u}_n\|_s, \|\hat{v}_n\|_s \leq R$ for all n, so by proposition 3.11 we have that

$$\|\hat{u}_{n+1} - \hat{v}_{n+1}\|_s \leq \lambda\|\hat{u}_n - \hat{v}_n\|_s,$$

and iterating yields the desired result. □

We now consider the difference of the true solution u_n and the estimator of the solution $\hat{u}_n$ with observation $y_{n+1} \in H^s$. We assume that we have no model uncertainty and proceed to define a measurement of how far the estimator is from the truth. We define the *error* between the estimator to the truth at time n by $e_n := \|\hat{u}_n - u_n\|_s$. To estimate this error, recall that by proposition 2.1 one has

$$\hat{u}_{n+1} = \eta^2(\eta^2 I + A)^{-1}\Psi(\hat{u}_n) + (\eta^2 I + A)^{-1}Ay_{n+1}.$$

where $y_{n+1} = \Psi(u_n) + \bar{\xi}_{n+1}$. As we are assuming we have no model error, we have $u_{n+1} = \Psi(u_n)$, so

$$u_{n+1} = \Psi(u_n) = (\eta^2 I + A)^{-1}[(\eta^2 I + A)\Psi(u_n)] = \eta^2(\eta^2 I + A)^{-1}\Psi(u_n) + (\eta^2 I + A)^{-1}A\Psi(u_n).$$

Subtracting yields

$$\hat{u}_{n+1} - u_{n+1} = \eta^2(\eta^2 I + A)^{-1}[\Psi(\hat{u}_n) - \Psi(u_n)] + (\eta^2 I + A)^{-1}A\bar{\xi}_{n+1}$$

by definition of the observation y_{n+1}. Thus, we obtain the following bound on the error e_n:

Theorem 3.15. *Fix $s \in [0,1]$, and suppose that assumption 3.12 holds, and that $\hat{u}_0 \in H^s$. Given initial estimators $\hat{u}_0, \hat{v}_0 \in H^s$, and observations $y_n \in H^s$ for every $n \in \mathbb{N}$, we have that, for all sufficiently small η, there exists a constant $\lambda < 1$ such that*

$$e_{n+1} \leq \lambda^{n+1}e_0 + \frac{M_2}{1-\lambda}$$

for every $n \in \mathbb{Z}^+$.

Proof. By assumption 3.12, $\|u_n\|_s \leq M_1$ and $\|\bar{\xi}_n\|_s \leq M_2$ for all n. Set $R = \max\{\|\hat{u}_0\|_s, \|\hat{v}_0\|_s\} + 2(M_1 + M_2 + \|\Psi_\eta(0)\|_s)$, and choose η small enough so that $\lambda := \frac{\eta^2}{4\pi^2+\eta^2}K(R) < \frac{1}{2}$. By lemma 3.13, $\|\hat{u}_n\|_s, \|\hat{v}_n\|_s \leq R$ for all n, so

$$e_{n+1} \leq \|\Psi_\eta(\hat{u}_n) - \Psi_\eta(u_n)\|_s + \|\bar{\xi}_{n+1}\|_s \leq \lambda e_n + M_2 \leq \lambda^{n+1}e_0 + M_2\sum_{j=0}^{n}\lambda^j \leq \lambda^{n+1}e_0 + \frac{M_2}{1-\lambda},$$

as required. □

Remark 3.16. *Note carefully the following:*

- *in theorems 3.14 and theorem 3.15, how large η can be depends on the sizes $\|\hat{u}_0\|_s$ and $\|\hat{v}_0\|_s$ of the initial estimators;*
- *in theorem 3.15, if the size M_2 of the noise is small, then the error is, in fact, asymptotically small.*

4 Numerics

In the preceding section we demonstrated the analytical properties of 3DVar and proved theorems regarding the convergence of different initial data and the convergence of the assimilation to the true solution. We now seek to validate and extend our results with numerical simulations. We first introduce the forward numerical model and discuss our choice of initial data, viscosity value and level of observation noise. In section 4.1 we describe our numerical system for the Navier–Stokes equation, and in section 4.2 we describe the behaviour of the numerical assimilator. We go on to demonstrate the implementation of 3DVar for the Stokes equations in section 4.3, and for the Navier–Stokes equations in section 4.4.

4.1 Numerics of forward model

In this section we will describe the properties of the Navier–Stokes equations (NSE) as it is run forward in time. The NSE displays quite diverse behaviour for different choices of parameter values. In particular, for a fixed forcing f we can cause the solution to be either stable, periodic or turbulent by varying the viscosity parameter ν.

We choose the forcing in the Navier–Stokes equations to be

$$f = \nabla^{\perp}(-\cos(5\pi(x_1 + x_2))) = (5\pi \sin(5\pi(x_1 + x_2)), -5\pi \sin(5\pi(x_1 + x_2)))^{\mathrm{T}},$$

where $\nabla^{\perp} = (\partial_2, -\partial_1)^{\mathrm{T}}$ (refer to figure 1).

In 2D turbulence, Kraichnan predicted heuristically that, if energy is injected into the system at a wavenumber k_f, then there is an inertial range for $k < k_f$ in which $E_k \propto k^{-5/3}$, with faster decay expected for $k > k_f$ due to the increasing role of viscosity (see [**11**]). Having chosen the forcing f, we observe that this is, indeed, the case (refer to figure 2).

There is a spike at $|k| = 7$ in figure 2 as $k_f = \sqrt{25 + 25} \approx 7$ is the wavenumber of the forcing term f. Up until this wavenumber E_k is approximately proportional to $k^{-5/3}$ and then decreases more rapidly. The presence of the double cascade also validates our simulation of 2D turbulence.

Our forward solver employs a 4^{th} order Runge–Kutta exponential time differencing scheme [**4**], as implemented in MATLAB by Dr. Kody Law. With that we simulated the NSE for different viscosity values to explore the different behaviour that can occur. We set our initial data to be $\sqrt{C}\beta$ where C is diagonal with

$$C_{kk} = \begin{cases} |k|^{-\frac{5}{3}}, & \text{if } |k| < 7 \\ |k|^{-5}, & \text{if } 7 \leq |k| < 16 \\ 0, & \text{else} \end{cases}$$

and β is a Gaussian random field (refer to figure 3).

Our simulations show that the NSE with a high viscosity ($\nu \geq 0.035$) and the above forcing leads to a steady state (figure 4(a)). For $0.027 \leq \nu \leq 0.035$ a periodic state is obtained (figure 4(c)) and for $\nu \leq 0.016$ we have a turbulent state (figure 4(e)). In the intermediate region ($0.016 \leq \nu \leq 0.027$) a quasiperiodic state is achieved where we observe a mixture of characterisitics from both periodic and turbulent states.

Figures 4(a), 4(c) and 4(e) show the stream function and its Fourier coefficients. The energy profile at each of the three states is shown in figures 4(b), 4(d) and 4(f). Recall that the stream function ζ is related to the velocity field u via the relation $u = \nabla^{\perp}\zeta$.

In the rest of the report we will focus our simulations on steady ($\nu = 0.04$), periodic ($\nu = 0.03$) and turbulent ($\nu = 0.016$) states.

4.2 Sensitivity of assimilator

The observations and the dynamics of the NSE are key components of the assimilation process. The relative weights of these two components, along with the level of observation noise and the complexity of the dynamics (caused by the value of viscosity ν) determines the behaviour of the assimilator.

Recall that

$$\hat{u}_{n+1} = \eta^2(I + \eta^2 A^{\alpha})^{-1} A^{\alpha} \Psi(\hat{u}_n) + (I + \eta^2 A^{\alpha})^{-1} y_{n+1}.$$

Looking at a particular mode k, we see that

$$[\hat{u}_{n+1}]_k = a(|k|)[\Psi(\hat{u}_n)]_k + b(|k|)[y_{n+1}]_k,$$

where we denote $[u]_k$ to be the k^{th} Fourier mode of u and

$$a(|k|) := \frac{\eta^2(2\pi|k|)^{2\alpha}}{1+\eta^2(2\pi|k|)^{2\alpha}}, \quad b(|k|) = 1 - a(|k|).$$

So for each k we have a convex combination of the dynamics and the observations. The choice of α decides how the wavenumber affects the weighting we give to the observations and to the dynamics. Since we took out the zeroth mode to make A invertible, $\min |k| = 1$. Let η_{cr} be defined as the value of η such that $a(|k|) = b(|k|)$ for $|k| = 1$, i.e.

$$\eta_{cr} = \frac{1}{(2\pi)^\alpha}.$$

Figures 5(a), 5(b), 5(c), 5(d) show the effect of α on the cases $\eta < \eta_{cr}$ and $\eta > \eta_{cr}$.

Let $\|e_n\|$ denote the absolute error between the truth u_n and our best estimate $\hat{u}_n$ at time nh. Then e_{n+1} satisfies

$$\|e_{n+1}\| \leq \|a(|k|)(\Psi(u_{n+1}) - \Psi(\hat{u}_n))\| + \|b(|k|)\xi_{n+1}\|$$

where $y_{n+1} = u_{n+1} + \xi_{n+1}$ for $\xi_{n+1} \sim N(0, \Gamma)$.

In the case $\alpha = 1$, the weighting of the dynamics, $a(|k|)$, tends to 1 as $k \to \infty$ while the weighting of the observations tends to 0. The intuition is that for η sufficiently small, the assimilator should put more weight on the observations for small Fourier modes and then switch to rely on the dynamics on high Fourier modes. In contrast for $\alpha = -1$, the behaviour of $a(|k|)$ and $b(|k|)$ are exchanged and the assimilation is largely influenced by the observations for η sufficiently small.

We gain some insight on how sensitive the assimilator is in reponse to the weight ratio η and observation noise level by running some initial numerics. The observation noise level is the average deviation of the observations from the true solution. We choose to base the initial numerics on $\eta = 1$ at a particular wavenumber: here we choose $(7, 7)$ (bearing no relation to the forcing at $(5, 5)$). We will double the level of noise from 0.1 to 0.2 and examine how the assimilation is affected for $\alpha = -1$ and 1. We will discuss further what we mean by the level of noise at the end of this section.

The true trajectory in each of the three regimes are shown in figure 6. Note that the scale on the graph in the steady and periodic case is around the order of 10^{-7} and the chosen level of observation noise is of order 10^{-1}. Hence we will not be able to see these characteristics in the initial assimilation of the steady and periodic regimes.

We observe from figures 7 and 8 that the main difference between the two cases is that the assimilation is more influenced by the observations when $\alpha = -1$, in agreement with figure 5(c); whereas when $\alpha = 1$ the assimilation pays little attention to the observations, in agreement with figure 5(b). We also note that for $\alpha = 1$ in the steady and periodic states, doubling the noise level does not have a significant effect on the assimilation, which converges to the true trajectory quickly. However in the turbulent regime, we see that the assimilation goes out of step with the true trajectory, but the distance between the two curves is no larger than 0.1 even after the level of noise is doubled.

When $\alpha = -1$ the observations play a bigger role in the assimilation process. The assimilation in the steady and periodic states do not lock onto the truth as in the case $\alpha = 1$. Rather the assimilation

route forms an envelope surrounding the true trajectory with width around the same order as the level of noise. Doubling the noise level doubles the width of the envelope. We also observe similar behaviour in the turbulent state, where the assimilation forms an envelope surrounding the true trajectory, as opposed to its counterpart in the case $\alpha = 1$.

To summarise, from the initial numerics we observe that the assimilation behaves well and locks on immediately in the steady and periodic regime when $\alpha = 1$. But for $\alpha = -1$ we see the variation in the assimilation doubles when the level of noise is doubled. Nonetheless the deviation is at most as high as the level of observation noise. For the turbulent regime we note similar behaviour when $\alpha = -1$. But when $\alpha = 1$ the assimilation, although unable to fully lock on, deviates at most 0.1 from the true solution.

To help us compare the different regimes fairly we keep the noise constant for the all three states and define the observation noise in the following way. Let u denote a true turbulent trajectory and $u_n = u(nh)$ denote its value at timestep n. Let $u^K = [u_j]_{j=1}^K$ where $u_j = \langle u, \psi_j \rangle$ and ψ_j is the basis function from the analysis section. We drop the K and consider the finite-dimensional setting in the following discussion. We define the average of the truth over time to be

$$\mu = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau u(t)\, \mathrm{d}t \approx \frac{1}{Nh} \sum_{j=1}^N u_j h = \frac{1}{N} \sum_{j=1}^N u_j.$$

Let the covariance matrix Σ be defined as

$$\Sigma = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau (u - \mu)(u - \mu)^{\mathrm{T}}\, \mathrm{d}t \approx \frac{1}{N} \sum_{j=1}^N (u_j - \mu)(u_j - \mu)^{\mathrm{T}}$$

where $(u_j - \mu)^{\mathrm{T}}$ denotes the transpose of $(u_j - \mu)$. Let $\vec{\sigma}^2$ be the diagonal of Σ, this gives us a sense of what the fluctuations are. We let $\sigma = \max \vec{\sigma}$, the maximum variation in the dynamics over all modes. From this we set $\varepsilon = 0.1 \times \sigma$ where $\Gamma = \varepsilon^2 I$, and the observations are defined as $y_n = u_n + \bar{\xi}_n$, where $\bar{\xi}_n \sim N(0, \Gamma)$. Note that this would violate the finite variance assumption if it was in infinite dimensions, but numerically we are working in finite dimensions. Hence with the above setting, the assumption on the noise in the analysis section is satisfied. In the subsequent sections, we will demonstrate how the numerics reflect the main results proved in the analysis section.

4.3 Numerics for Stokes equation

Here we demonstrate the implementation of 3DVar for the Stokes equations [**9**]. The analysis section covered both $\alpha = 1$ and -1 cases, but we will just discuss the $\alpha = 1$ case briefly before discussing numerical results in more detail for the NSE. Unlike the Navier–Stokes equations this is a linear equation and we always obtain a steady state regardless of the value of the viscosity. Recall from the analysis section that h is the time between observations. In the simulation we take h to be 0.16, $\alpha = 1$ and $\nu = 0.016$, a value which would have caused turbulent dynamics for the NSE. We observe from figure 9 that the estimators easily lock onto the true trajectory for different initial conditions (which verifies theorem 3.6).

4.4 Numerics for Navier–Stokes equations

In this section we apply the 3DVar algorithm to the NSE and show that the numerical results agree with the results proved in the analysis section for $\alpha = -1$. We will then show that similar results can

be expected for the case $\alpha = 1$. We begin by showing we obtain the predicted behaviour in the worst case scenario of turbulent dynamics. If 3DVar performs well in this situation then it is reasonable to suppose it will perform even better in the cases of periodic and stable dynamics.

4.4.1 Turbulent dynamics, case $\alpha = -1$

Throughout this section we take the viscosity value $\nu = 0.016$, which gives light turbulence. We have confirmed that our implementation of 3DVar works well for high levels of turbulence but parameter values have to be chosen more carefully.

Theorem 3.14 says that given sufficiently small values of η, the estimators arising from two different initial conditions given the same observations will get arbitrarily close. The numerics confirm this. For $\eta = 25$ and timestep $h = 0.032$ the estimators are initially far apart, but as the number of iterations increases they get closer together and closer to the truth (figures 12, 13 and 14). A smaller value of η, such as $\eta = 2$, results in quicker convergence. However for small h in the $\alpha = -1$ case, even if we take a large value of η, such as $\eta = 150$, the estimators still converge (figures 15, 16, 17). This is because for large wavenumbers we are giving all the weight to the observations and therefore cannot get too far away from the truth. Eventually this forces the lower wavenumbers to be close to the truth and we get convergence. This is not necessarily true if h is taken so large as to miss the underlying dynamics, as in figure 18.

Theorem 3.15 says that for sufficiently small η we expect the error $e_n^2 = \|\hat{u}_n - u_n\|_{L^2}^2$ to be less than some small threshold eventually; this small threshold is governed by the size of the noise (see remark 3.16). For $\eta = 25$ we see e_n^2 decreases to just below 10^{-4} (figure 19). Even for moderate η, the $\alpha = -1$ case weights the observation noise quite highly and hence prevents e_n^2 decreasing much below 10^{-4}. Therefore taking $\alpha = -1$ does not offer much benefit over just taking the observation values as our estimator. We will see later that $\alpha = 1$ is the better choice in practice.

4.4.2 Turbulent dynamics, case $\alpha = 1$

We observe that the numerics suggest similar results should hold for the case $\alpha = 1$. Again we use the viscosity value $\nu = 0.016$ to get light turbulence. If we take $\eta = 0.05$ and $h = 0.032$ then the estimators are initially far apart, but as the number of iterations increases they get closer together and closer to the truth (figures 20, 21 and 22). As before smaller values of η, such as $\eta = 0.015$, result in quicker convergence (figure 23). However, unlike in the case $\alpha = -1$, if we take a value of η which is too large, for example $\eta = 0.5$, then the estimators will not necessarily converge (figures 24, 25 and 26). Hence, from our simulations, we expect an analogue to theorem 3.14 for the case $\alpha = 1$, which would say that the estimators arising from different initial conditions will become arbitrarily close, so long as η is chosen sufficiently small.

We also expect an analogue of theorem 3.15 for the case $\alpha = 1$, which would say that for sufficiently small η the error e_n^2 will be less than some small threshold eventually. We can see this behaviour for $\eta = 0.05$ in figure 27. Since we are taking $\alpha = 1$ we give the model a lot of weight for large wavenumbers. The downside of this is that, even for small h, if η is too large then the observations are unable to pull the estimators close to the truth, as we saw in figure 24. However the estimators do lock on for well chosen values of η and h. In addition, when we get convergence the error e_n^2 tends to a much smaller value of around 7×10^{-5} compared to 10^{-4} for the $\alpha = -1$ case. Hence $\alpha = 1$ gives us

almost two orders of magnitude less error than just taking the observations, which shows that using 3DVar is worthwhile.

As we are not guaranteed convergence even for small h, we have to be more careful with the choice of parameter values in the $\alpha = 1$ case. If we increase h then we need to take η smaller to still get convergence of the estimator to the truth. For example, if we increase h from 0.032 to 0.16 then for $\eta = 0.05$ we no longer get convergence (figure 29). However we do get convergence if we decrease η to 0.01 (figure 30).

4.4.3 Periodic and stable dynamics

Now that we have verified that the implementation works for turbulent dynamics we will look at the nicer situation of periodic dynamics ($\nu = 0.032$). In this case we can get away with much larger values of η and still have the estimators lock onto the true solution. For example, with turbulence, for $\eta = 0.5$ in the $\alpha = 1$ case we did not get successful assimilation. However in the periodic case the assimilation works, as can be seen in figures 31, 32 and 33.

For the case of stable dynamics the assimilation is easier still. So long as we give some weight to the observations the estimators will eventually lock onto the truth. This is shown in figures 34 and 35. Here we take the even larger values $\eta = 1$ and $h = 0.2$ (i.e. we give very high weight to the model and have infrequent observations) but we still get convergence.

We do not include any plots for the $\alpha = -1$ case for stable or periodic dynamics as we have already demonstrated it to be effective in the turbulent case.

5 Conclusions and Further Research

In this report we have established the following two analytic results for 3DVar:

- that estimators starting from two different initial conditions become asymptotically equal as $n \to \infty$, at a geometric rate; and
- that estimators come within a bounded distance of the true solution, as long as the observational noise is uniformly bounded;

in the following cases:

- for the Stokes equations, in the cases $\alpha = 1$ and $\alpha = -1$;
- for the Navier–Stokes equations, in the case $\alpha = -1$.

Then, using numerical simulations, we have confirmed these results, and we predict that similar results should hold for the Navier–Stokes equations in the case $\alpha = 1$. While we made some progress in establishing analogous results to theorems 3.14 and 3.15 for $\alpha = 1$, unfortunately we were not able to complete the proofs; this is thus a natural avenue for further research. Beyond that, an immediate generalisation would be to that of general $\alpha \in \mathbb{R}$.

Besides this, there are two obvious ways in which we could generalise the results we have proved. Firstly, it would be interesting to look at the case when the observation operator $\mathcal{H}$ is not the identity operator. Secondly, it would be very useful to consider *model error*: instead of setting

$$u_{n+1} := \Psi(u_n), \tag{5.1}$$

we let $\{\xi_n\}_{n\in\mathbb{Z}^+}$ be a sequence of independent and identically distributed X-valued random variables, and define $\{u_n\}_{n\in\mathbb{Z}^+}$, $u_n \in X$ by

$$u_{n+1} := \Psi(u_n) + \xi_n \tag{5.2}$$

for $n \in \mathbb{Z}^+$. Both of these are significant generalisations, worthy of considerable further study.

In closing, we thank Prof. Andrew Stuart, Dr Kody Law, and Dr Masoumeh Dashti for their help, guidance and inspiration during this project.

References

1. A. Apte, C. K. R. T. Jones, A. M. Stuart, and J. Voss. Data assimilation: mathematical and statistical perspectives. *Internat. J. Numer. Methods Fluids*, 56(8):1033–1046, 2008.

2. S. L. Cotter, M. Dashti, J. C. Robinson, and A. M. Stuart. Bayesian inverse problems for functions and applications to fluid mechanics. *Inverse Problems*, 25(11):115008, 43, 2009.

3. P. Courtier, E. Andersson, W. Heckley, D. Vasiljevic, M. Hamrud, A. Hollingsworth, F. Rabier, M. Fisher, and J. Pailleux. The ECMWF implementation of three-dimensional variational assimilation (3d-Var). I: Formulation. *Quart. J. R. Met. Soc.*, 124(550):1783–1807, 1998.

4. S. M. Cox and P. C. Matthews. Exponential time differencing for stiff systems. *J. Comput. Phys.*, 176(2):430–455, 2002.

5. K. Hayden, E. Olson, and E. S. Titi. Discrete Data Assimilation in the Lorenz and 2D Navier–Stokes equations. 2010. arXiv preprint 1010.6105v1.

6. M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE constraints*, volume 23 of *Mathematical Modelling: Theory and Applications*. Springer, New York, 2009.

7. L. Hörmander. *The analysis of linear partial differential operators. I*, volume 256 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1983.

8. K. Ide, P. Courtier, M. Ghil, and A. C. Lorenc. Unified notation for data assimilation: operational, sequential and variational. *J. Met. Soc. Japan*, 75(1B):181–189, 1997.

9. A. C. Lorenc. Analysis methods for numerical weather prediction. *Quart. J. R. Met. Soc.*, 112 (474):1177–1194, 2000.

10. A. C. Lorenc, S. P. Ballard, R. S. Bell, N. B. Ingleby, P. L. F. Andrews, D. M. Barker, J. R. Bray, A. M. Clayton, T. Dalby, D. Li, T. J. Payne, and F. W. Saunders. The Met. Office global three-dimensional variational data assimilation scheme. *Quart. J. R. Met. Soc.*, 126(570):2991–3012, 2000.

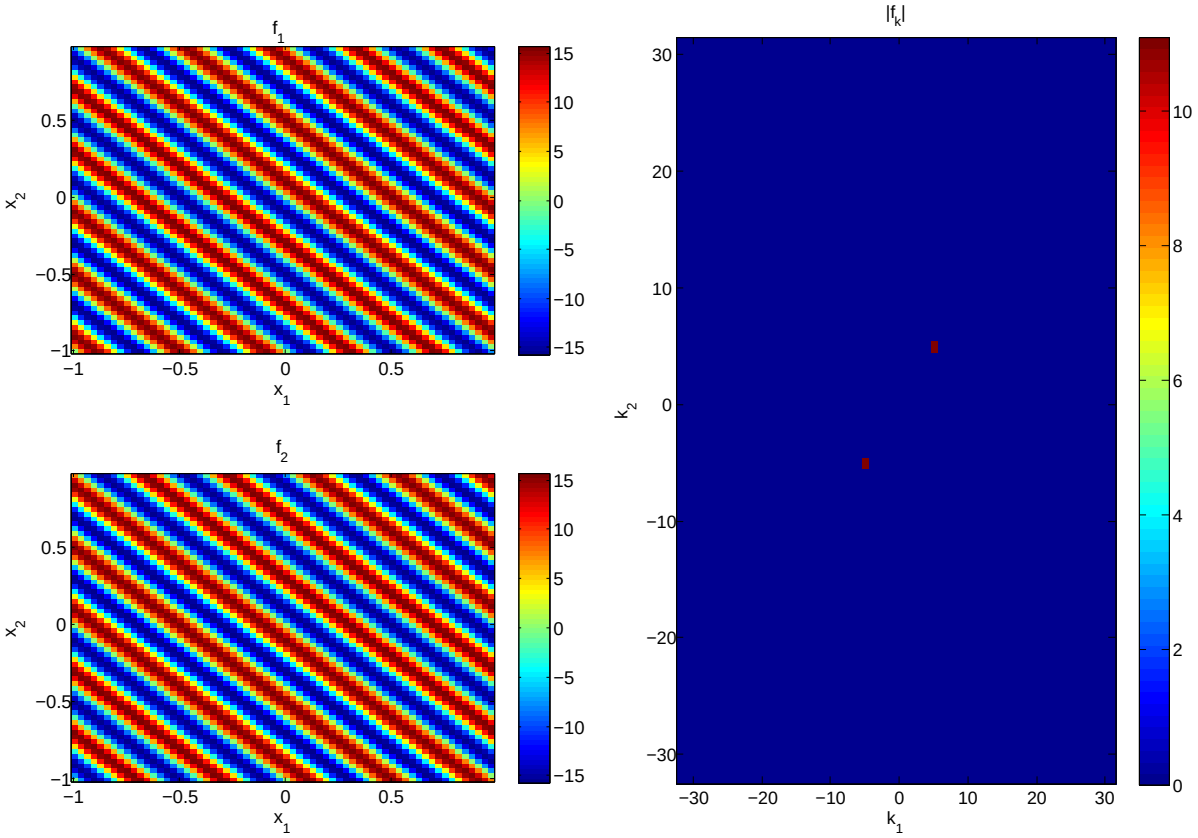
11. A. S. Monin and A. M. Yaglom. *Statistical fluid mechanics: mechanics of turbulence. Vols. I and II*. Dover Publications Inc., Mineola, NY, 2007. Translated from the 1965 Russian original.

12. D. F. Parrish and J. C. Derber. The national meteorological centers spectral statistical-interpolation analysis system. *Monthly Weather Review*, 120(8):1747–1763, 1992.

13. J. C. Robinson. *Infinite-dimensional dynamical systems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2001.

14. A. M. Stuart. Inverse problems: a Bayesian perspective. *Acta Numer.*, 19:451–559, 2010.

15. R. Temam. *Infinite-dimensional dynamical systems in mechanics and physics*, volume 68 of *Applied Mathematical Sciences*. Springer-Verlag, New York, second edition, 1997.



(a) The first and second components f_1 and f_2 of the forcing f in real space.

(b) Fourier coefficients of the forcing.

Figure 1: The forcing f

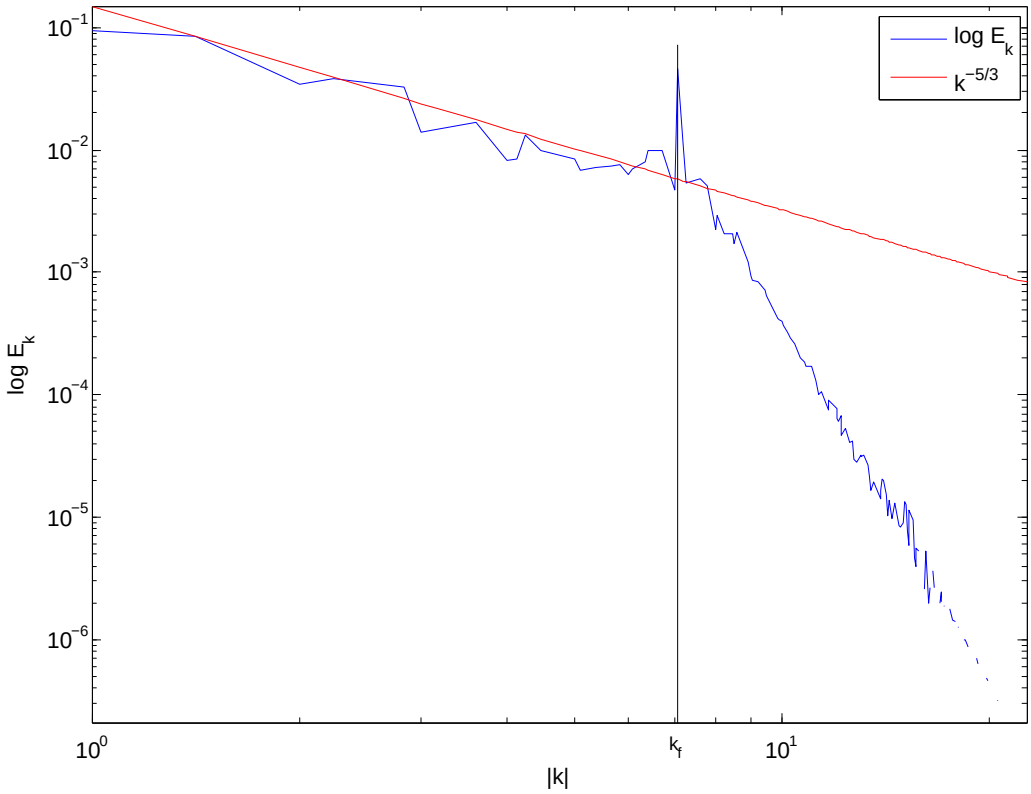


Figure 2: The energy E_k against wavenumber k for $\nu = 0.012$. We observe a Kolmogorov cascade.

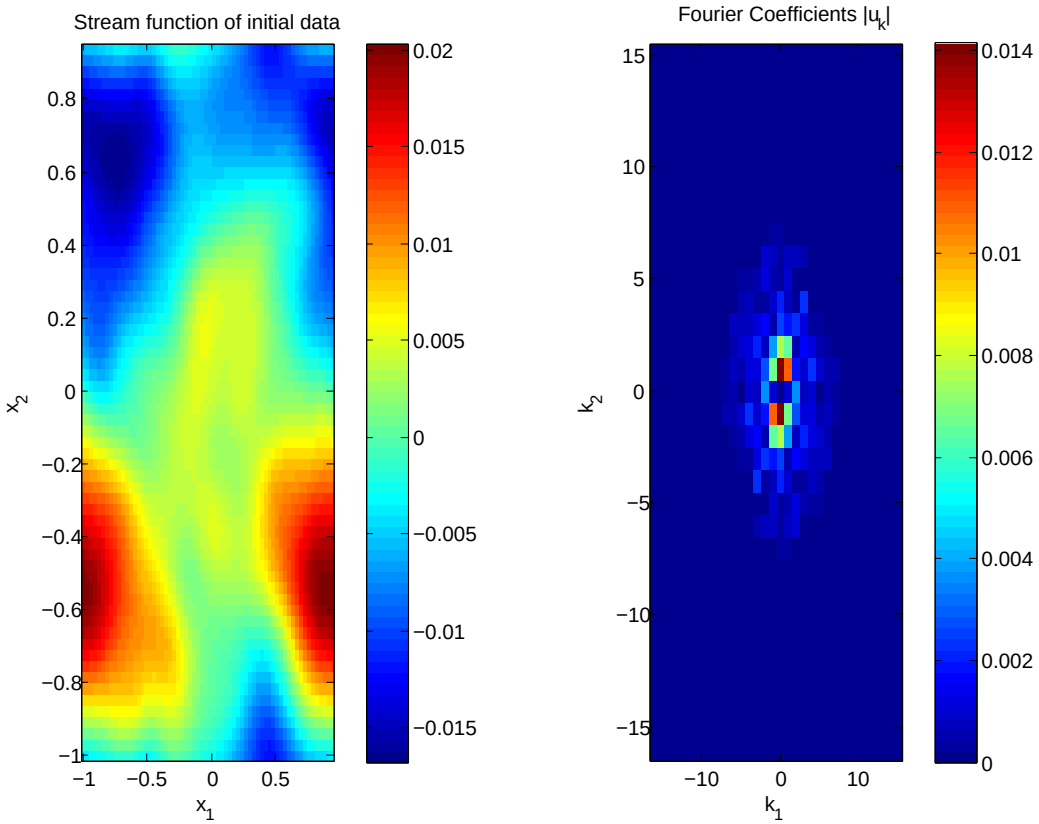
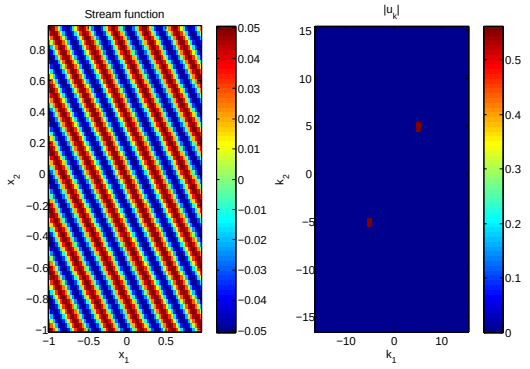
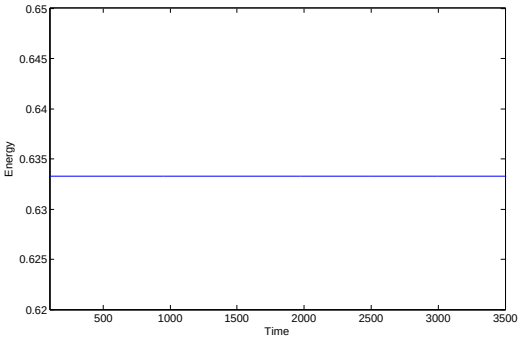


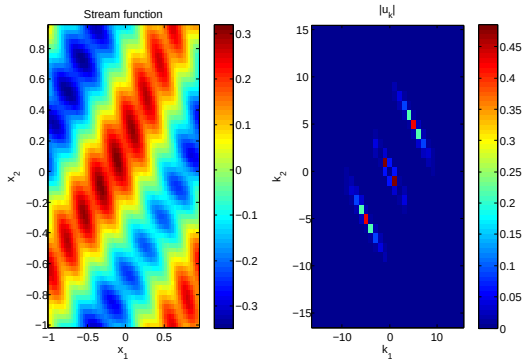
Figure 3: Initial data in real space and Fourier space.



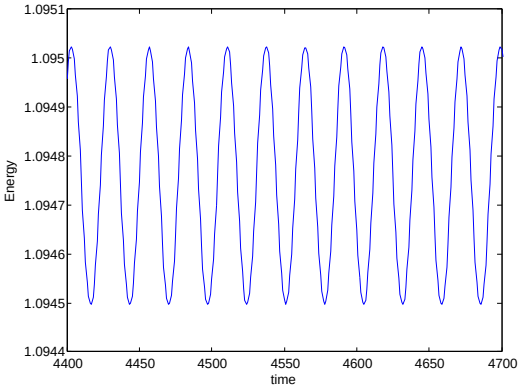
(a) Steady state with $\nu = 0.04$.



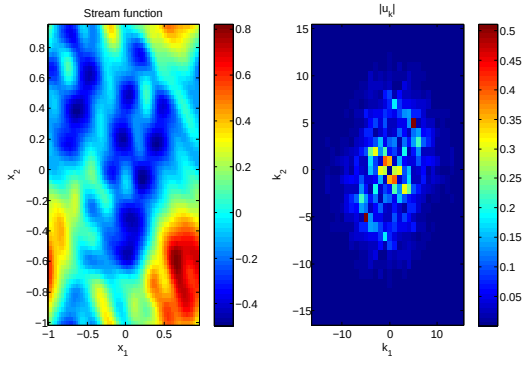
(b) Energy plot of steady state.



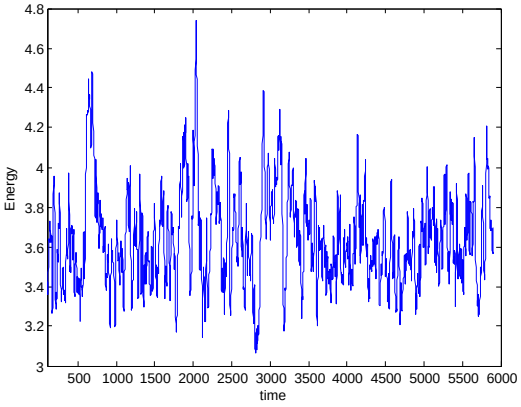
(c) Periodic state with $\nu = 0.03$.



(d) Energy plot of periodic state.

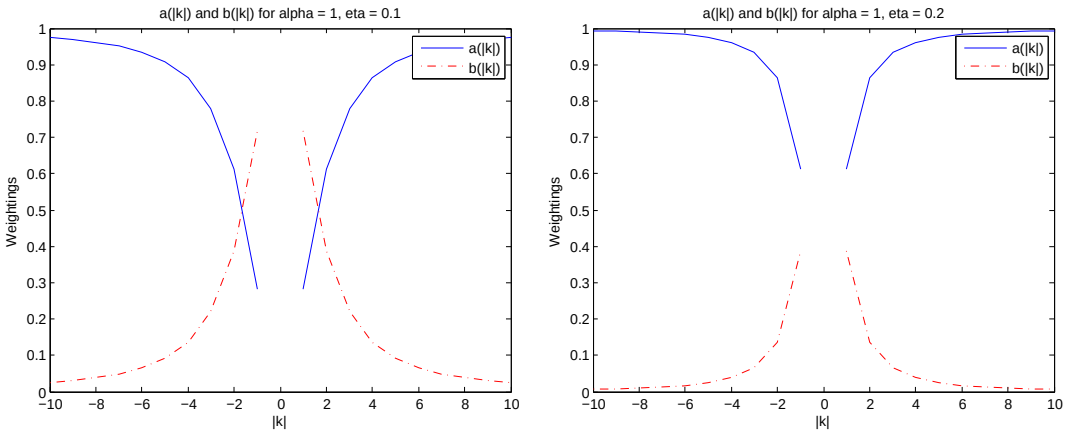


(e) Turbulent state with $\nu = 0.016$.



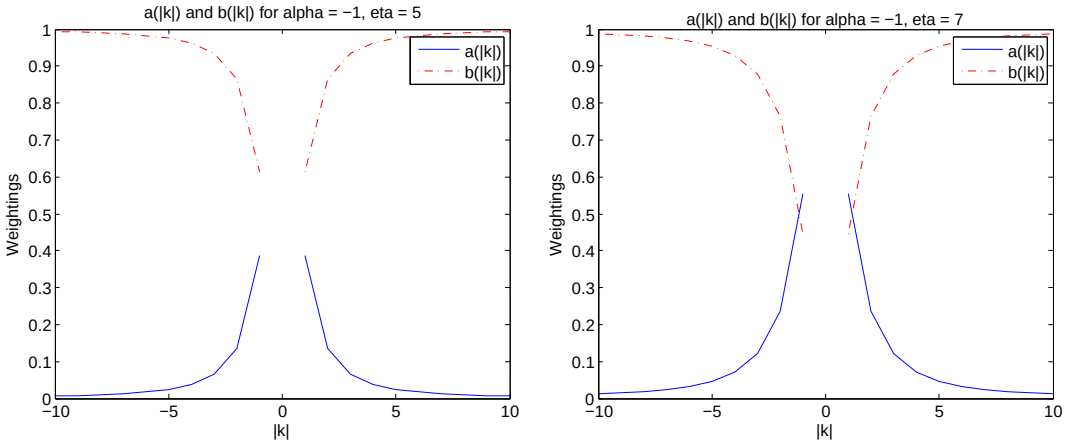
(f) Energy plot of turbulent state.

Figure 4: Simulated states of the NSE



(a) $\alpha = 1$ and $\eta < \frac{1}{2\pi}$

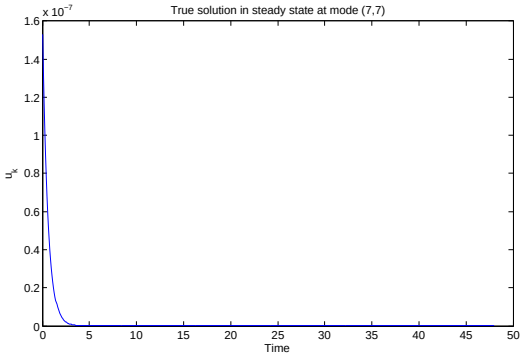
(b) $\alpha = 1$ and $\eta > \frac{1}{2\pi}$



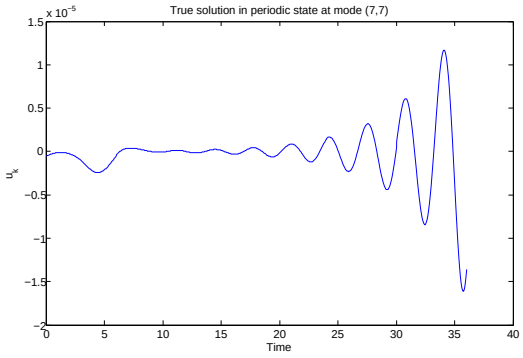
(c) $\alpha = -1$ and $\eta < 2\pi$

(d) $\alpha = -1$ and $\eta > 2\pi$

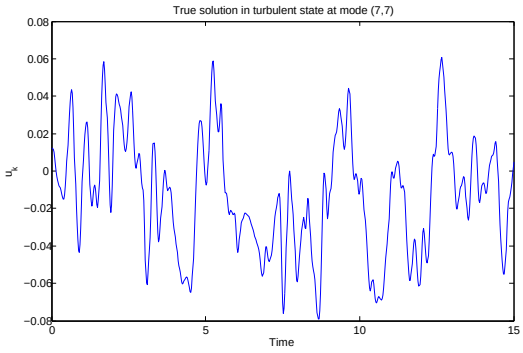
Figure 5: Effects of α on η



(a) True solution in steady state

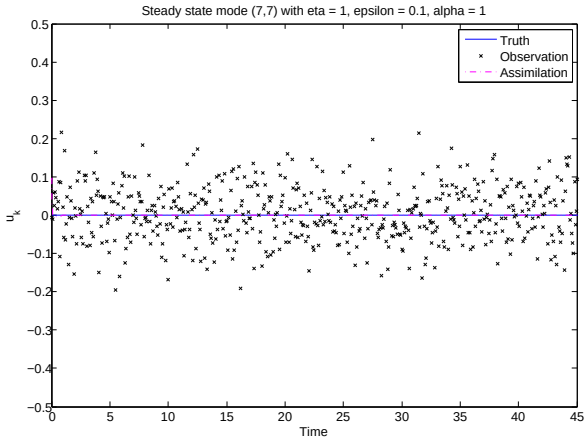


(b) True solution in periodic state

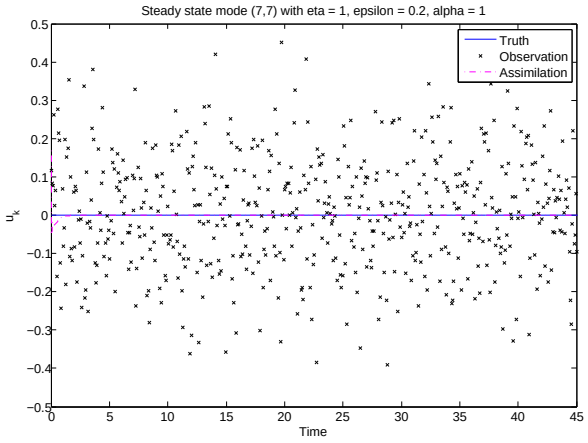


(c) True solution in turbulent state

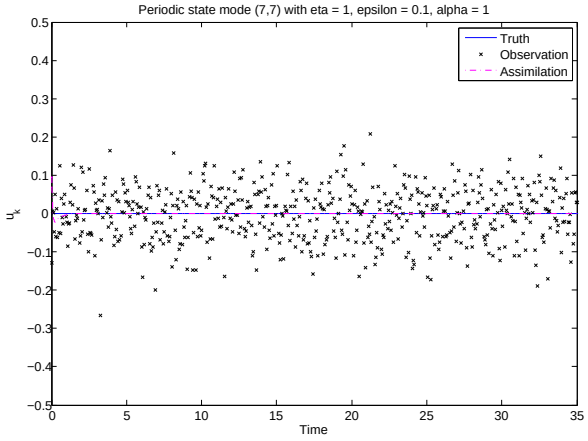
Figure 6: The true solution in each of the three regimes



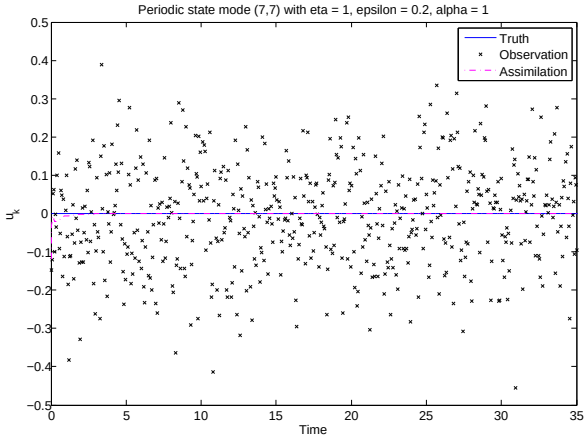
(a) Steady state with $\epsilon = 0.1$



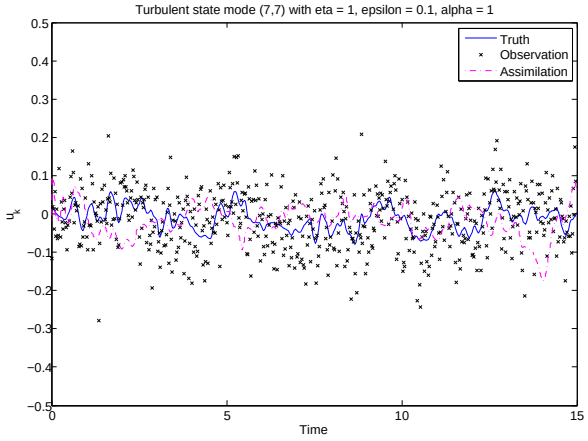
(b) Steady state with $\epsilon = 0.2$



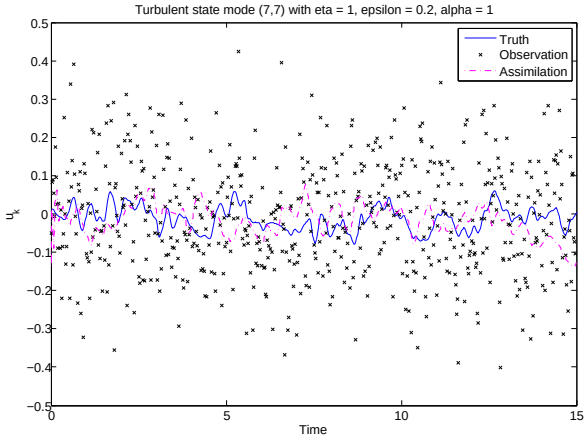
(c) Periodic state with $\epsilon = 0.1$



(d) Periodic state with $\epsilon = 0.2$

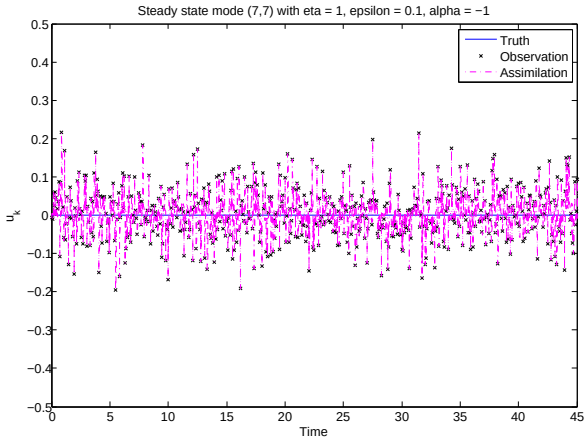


(e) Turbulent state with $\epsilon = 0.1$

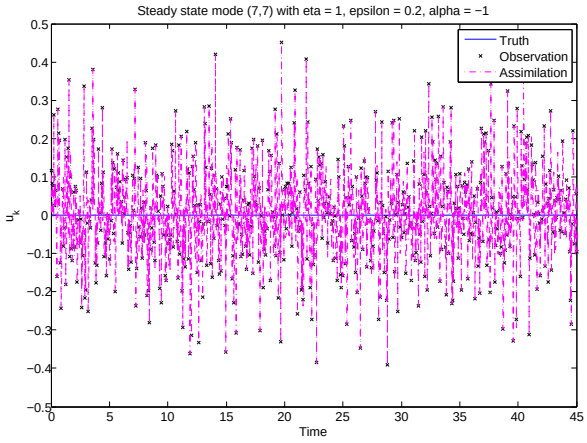


(f) Turbulent state with $\epsilon = 0.2$

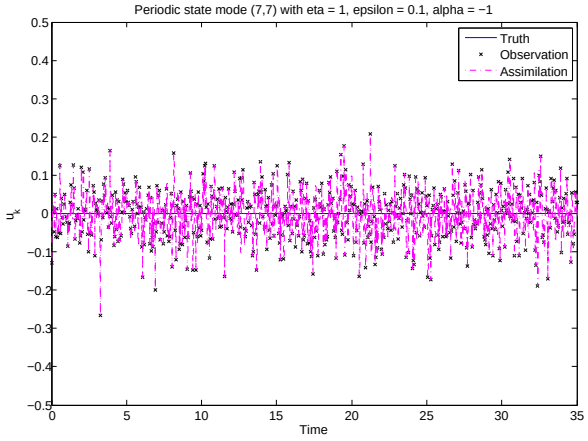
Figure 7: Initial assimilation for $\alpha = 1$



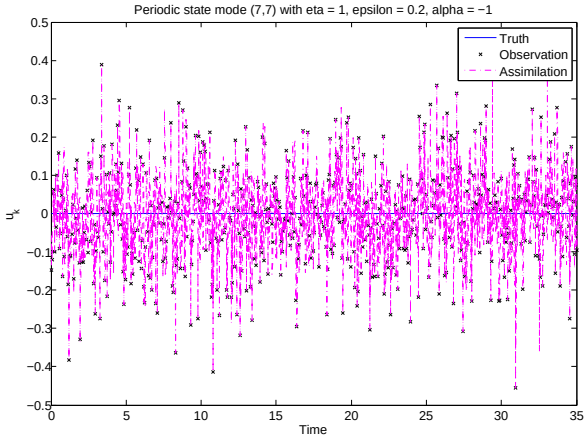
(a) Steady state with $\epsilon = 0.1$



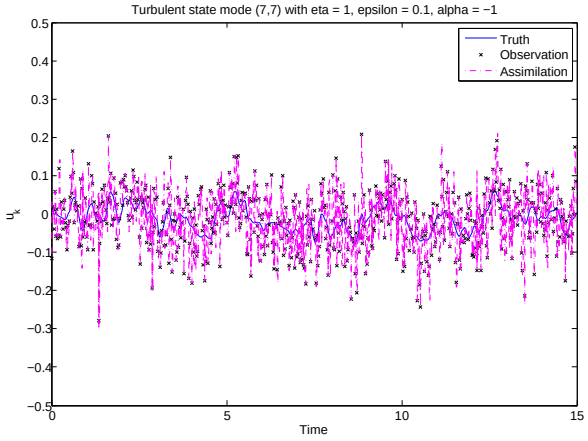
(b) Steady state with $\epsilon = 0.2$



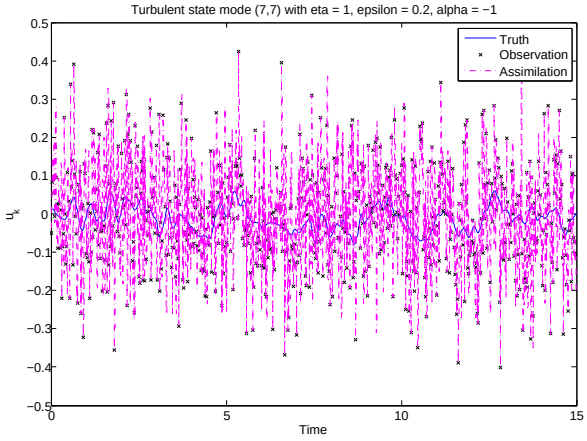
(c) Periodic state with $\epsilon = 0.1$



(d) Periodic state with $\epsilon = 0.2$



(e) Turbulent state with $\epsilon = 0.1$



(f) Turbulent state with $\epsilon = 0.2$

Figure 8: Initial assimilation for $\alpha = -1$

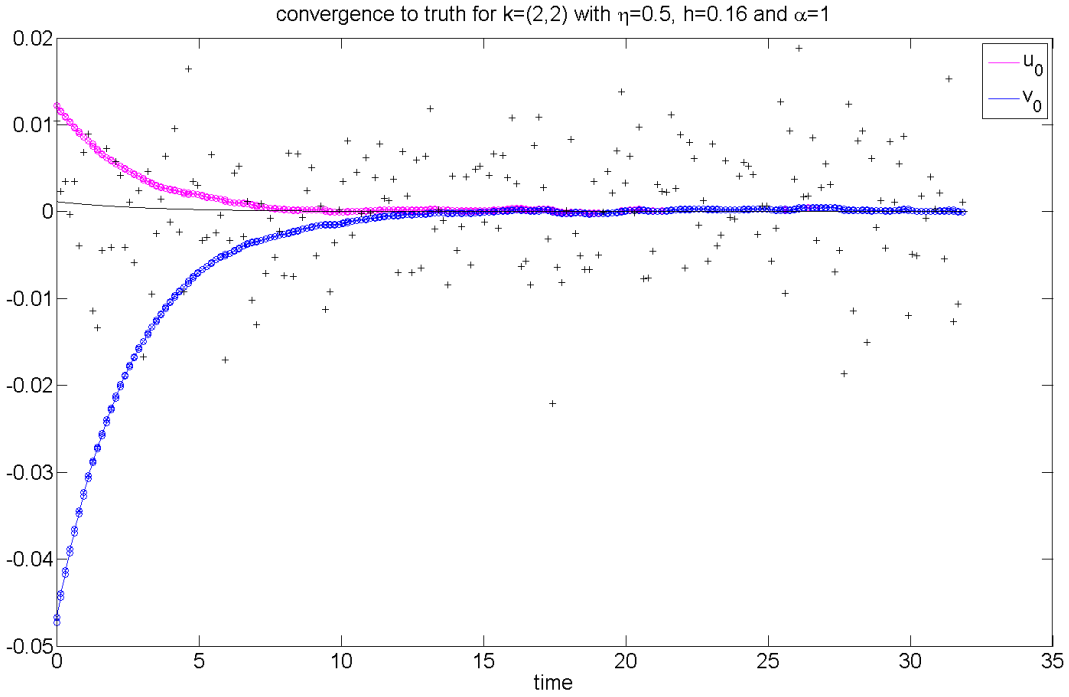


Figure 9: Estimators for two different initial conditions. The circles with a cross inside give the estimators $\hat{u}_n$ and the empty circles gives $\Psi(\hat{u}_n)$, the estimators evolved forward one time step. The plus signs give the observations y_{n+1} that are being assimilated and the black line is the truth. We see that the estimators rapidly converge together.

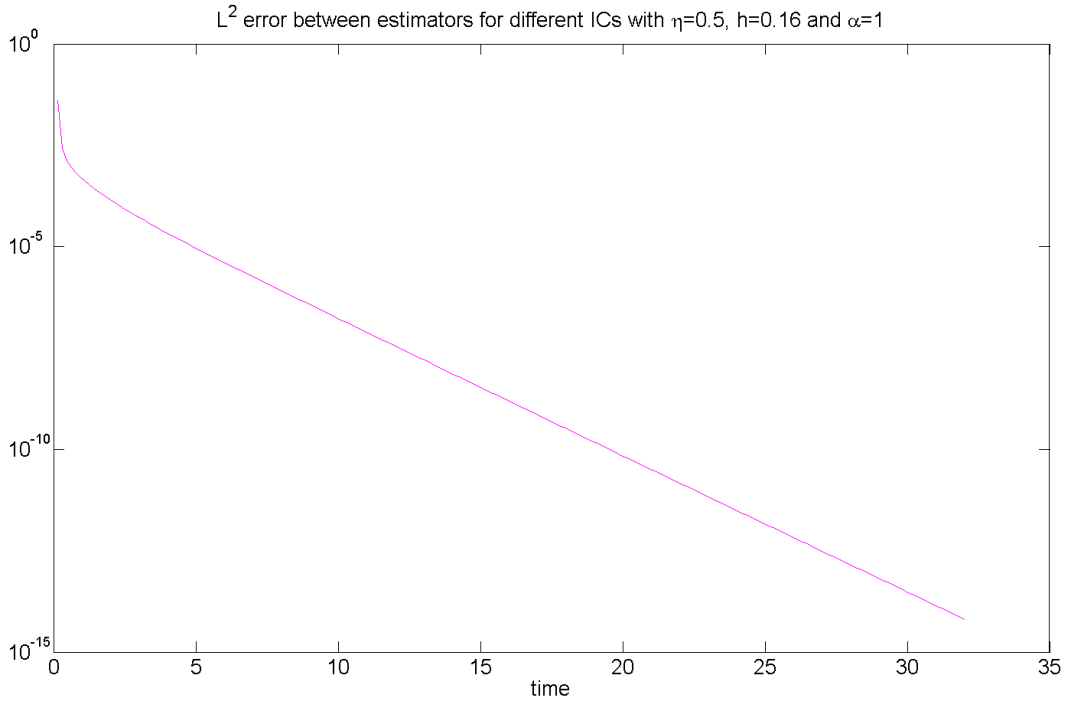


Figure 10: The square of the L^2 difference between the estimators for two different initial conditions for Stokes equation. The difference decreases rapidly.

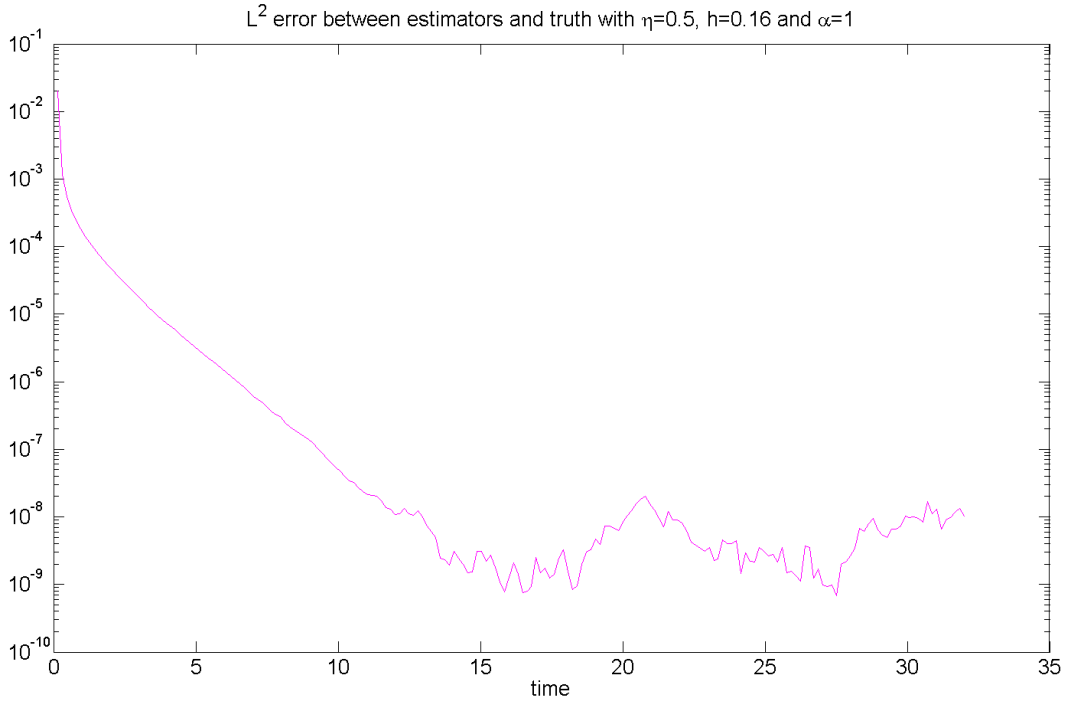


Figure 11: A plot of $e_n^2 := \|\hat{u}_n - u_n\|_{L^2}^2$ against time for Stokes equation. The stable solution allows the error to decrease exponentially to a very small value of around 10^{-9}.

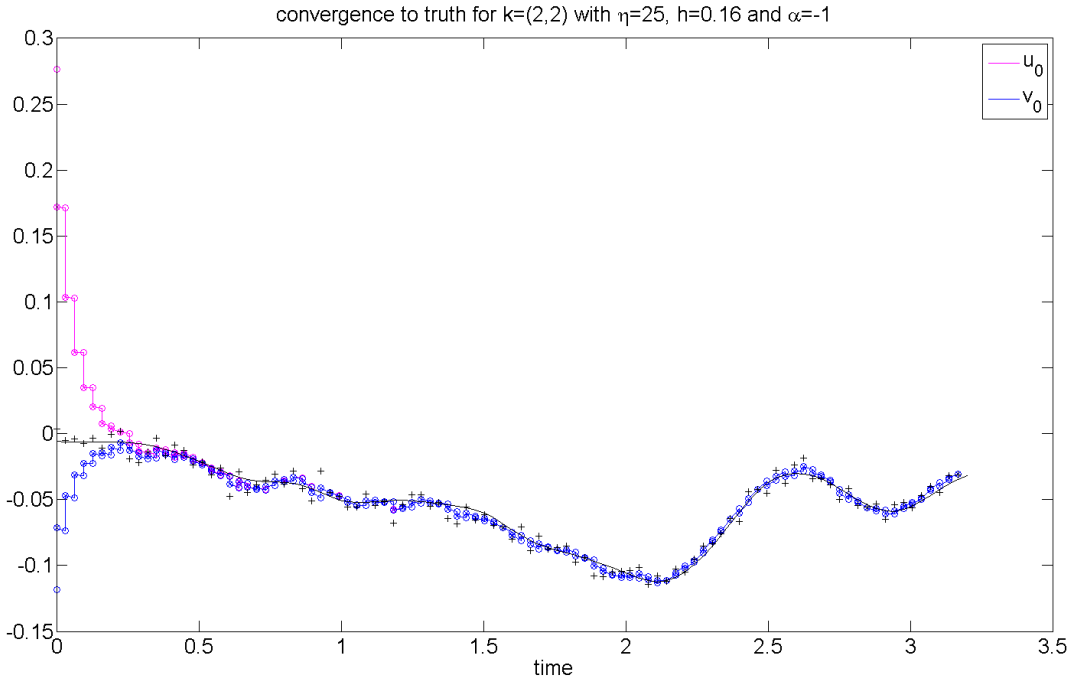


Figure 12: Estimators for the same observations but with two different initial conditions. The circles with a cross inside give the estimators $\hat{u}_n$ and the empty circles give $\Psi(\hat{u_n})$, the estimators evolved forward one time step. The plus signs give the observations y_{n+1} that are being assimilated and the black line is the truth. We can see that the estimators rapidly converge together and become almost indistinguishable.

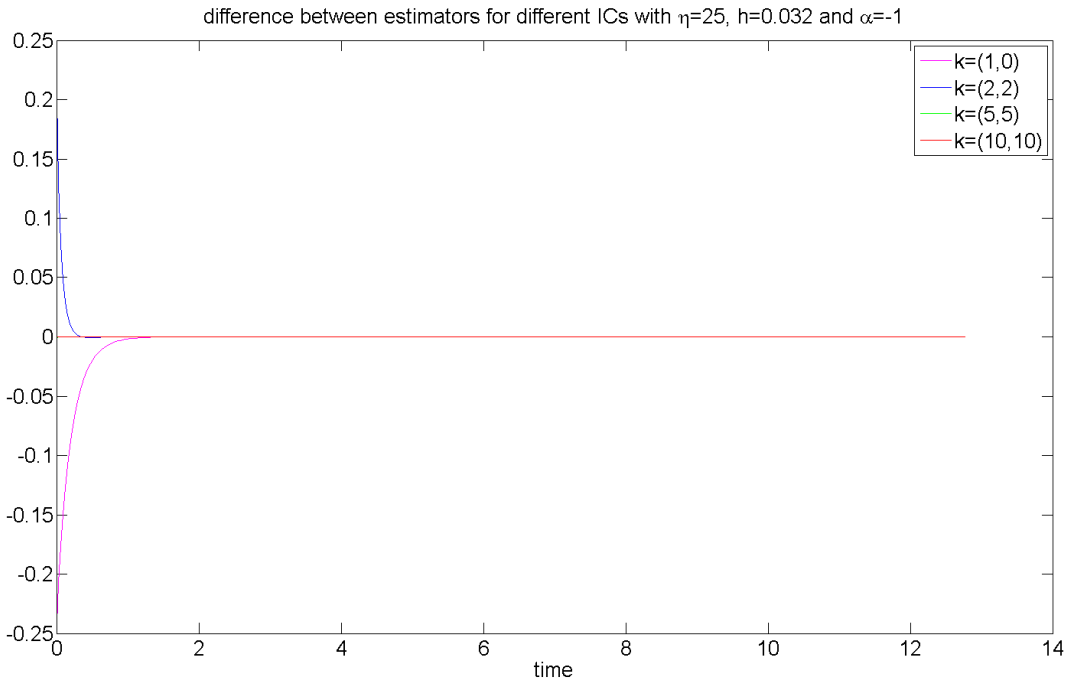


Figure 13: The difference between the estimators for the two initial conditions in figure 12 for a selection of different wavenumbers. For all wavenumbers the difference becomes almost zero very quickly.

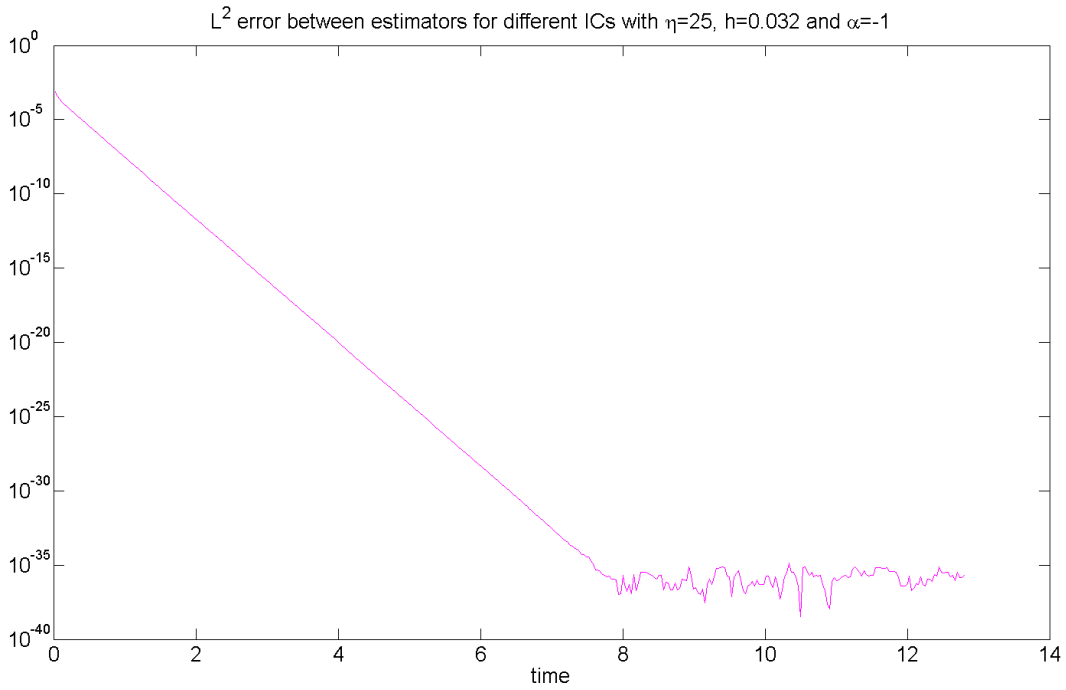


Figure 14: The square of the L^2 difference between the estimators for two different initial conditions. The difference decreases exponentially until limited by machine precision.

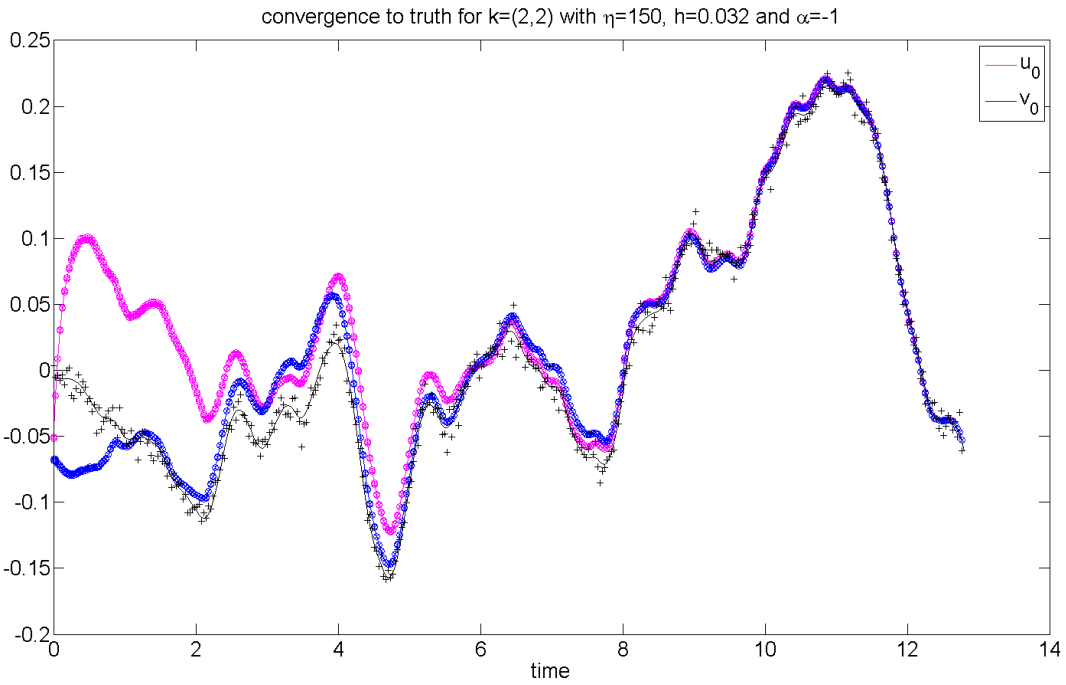


Figure 15: The same plot as figure 12. Even for this large value of η the estimators converge, though it takes a little longer than in figure 12.

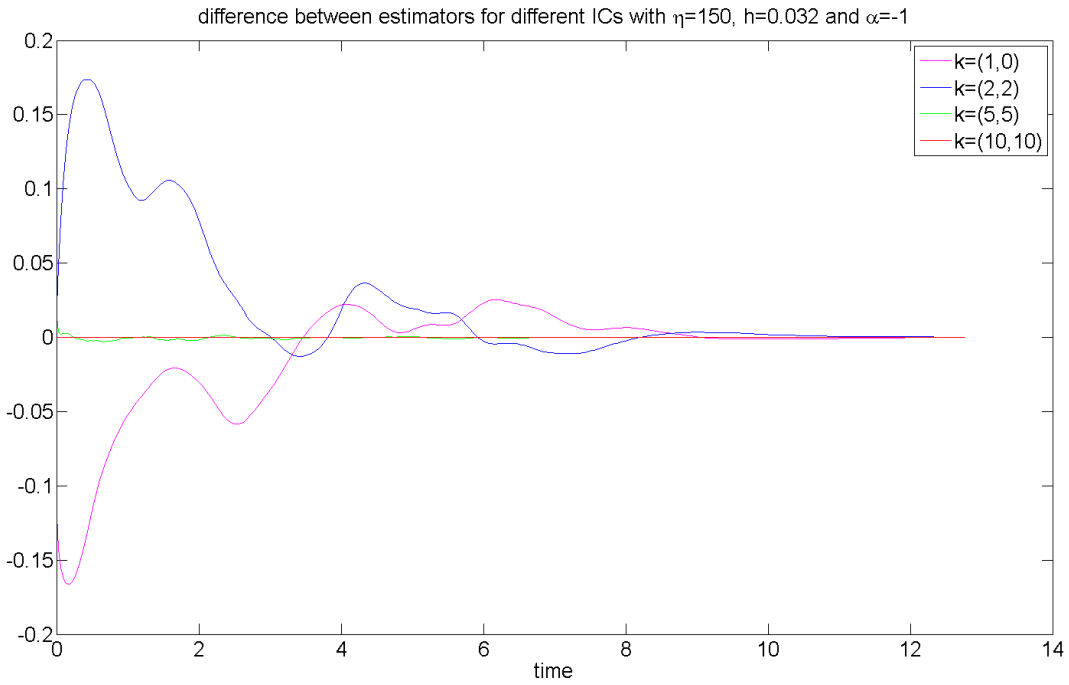


Figure 16: The difference between the estimators for two different initial conditions for a selection of different wavenumbers. The differences eventually decrease to zero.

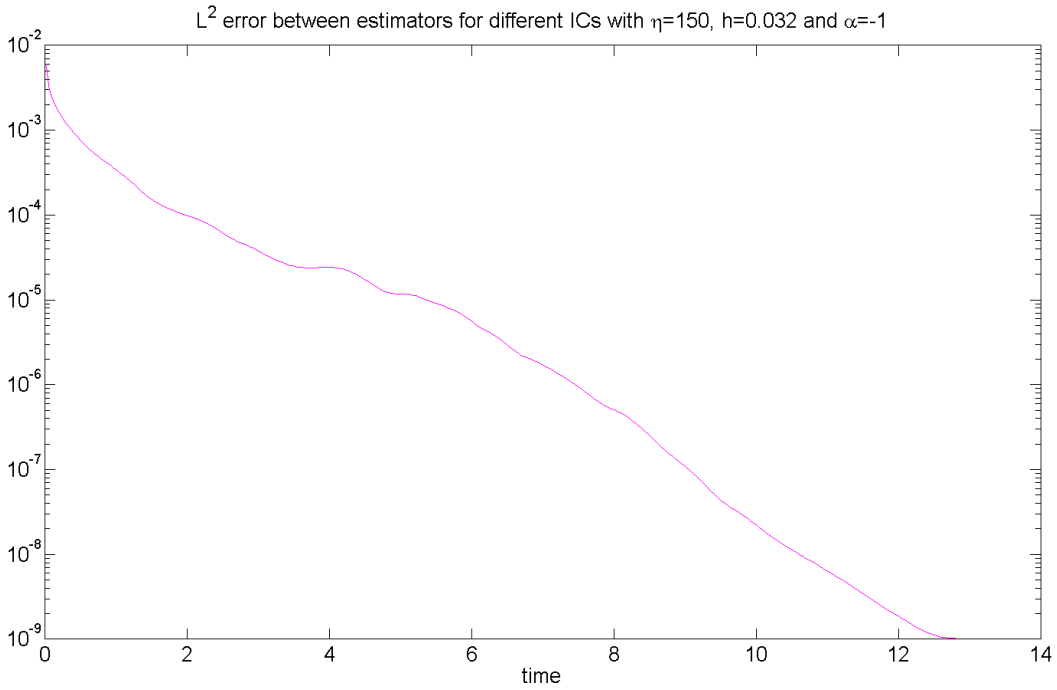


Figure 17: The square of the L^2 difference between the estimators for the two initial conditions. Despite the large η the difference between estimators still decreases exponentially.

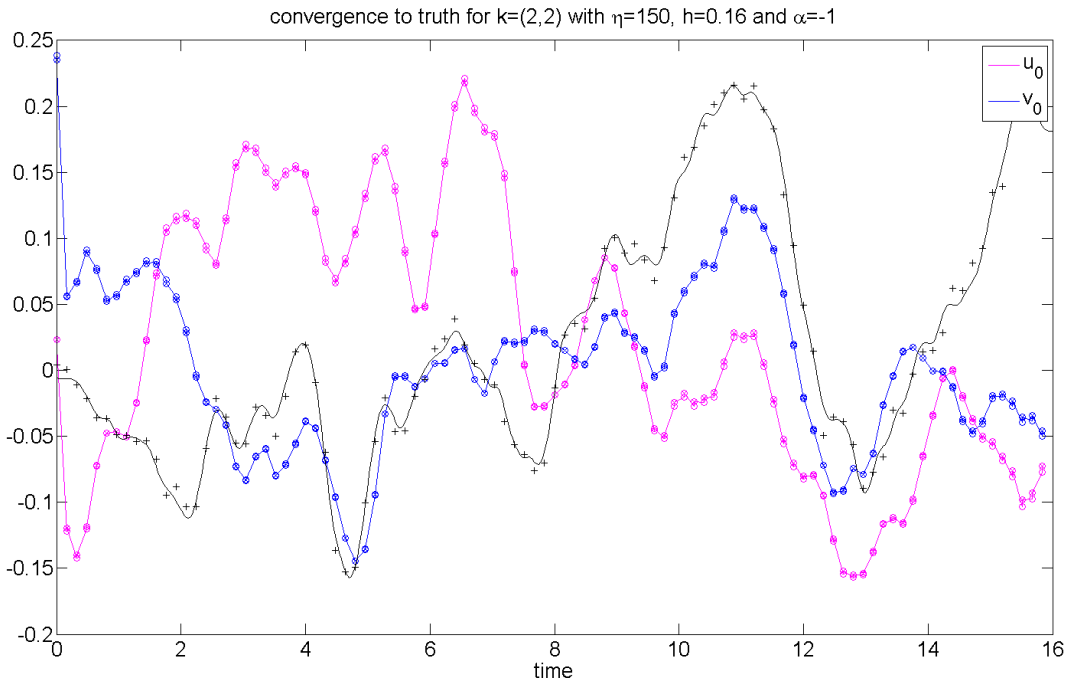


Figure 18: The same plot as in figure 12. The observations are too infrequent to allow the estimators to lock on to the truth, unlike in figure 15.

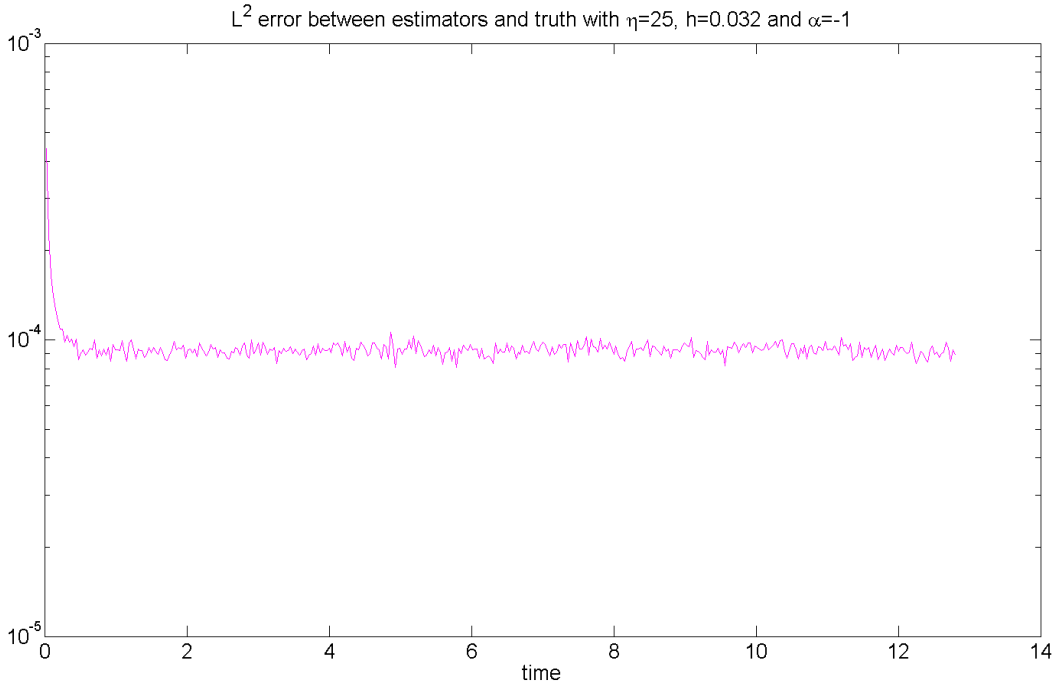


Figure 19: A plot of $e_n^2 := \|\hat{u}_n - u_n\|_{L^2}^2$ against time. The error decreases rapidly to around 10^{-4}.

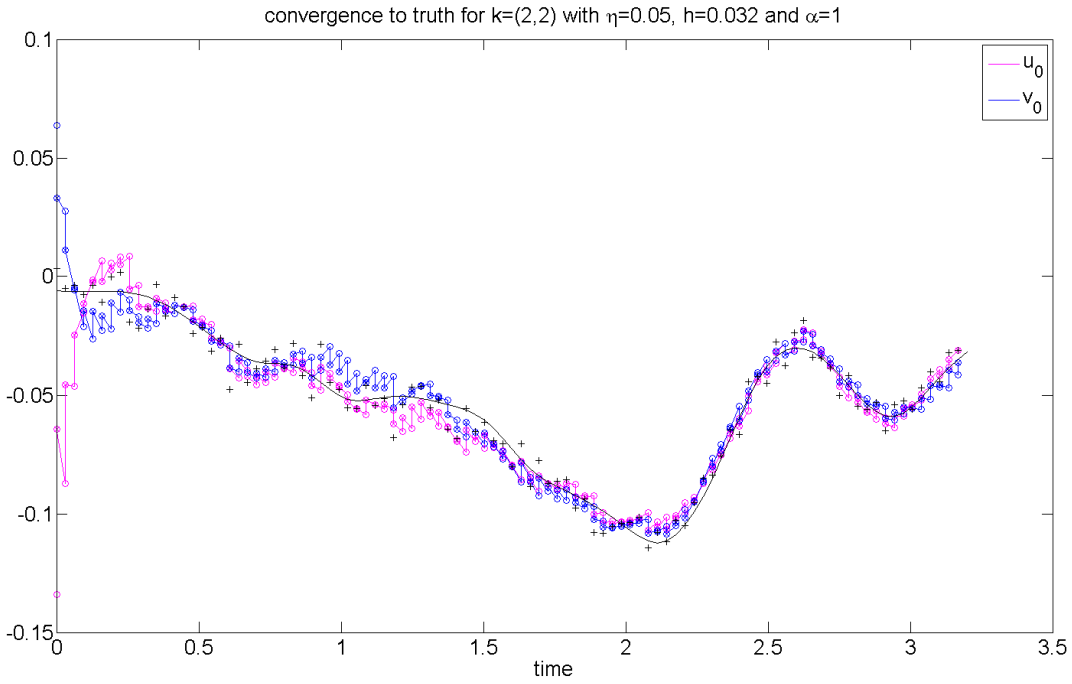


Figure 20: The same plot as in figure 12. The estimators rapidly become close together and close to the truth.

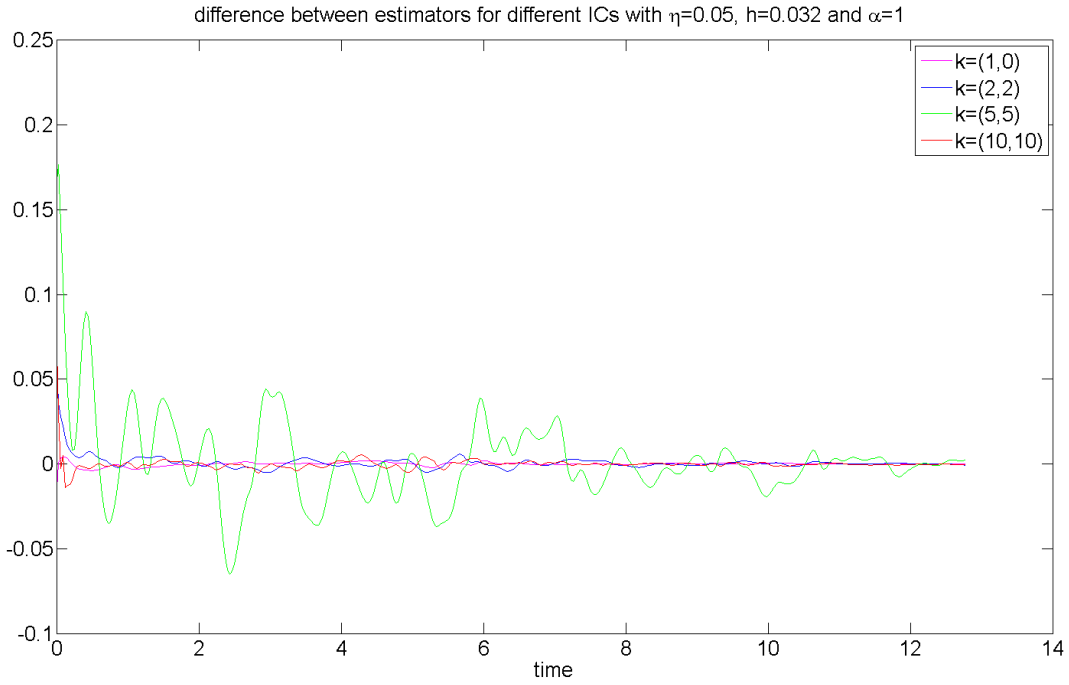


Figure 21: The difference between the estimators for two different initial conditions for a selection of different wavenumbers. We can see that the difference between the estimators becomes small very quickly.

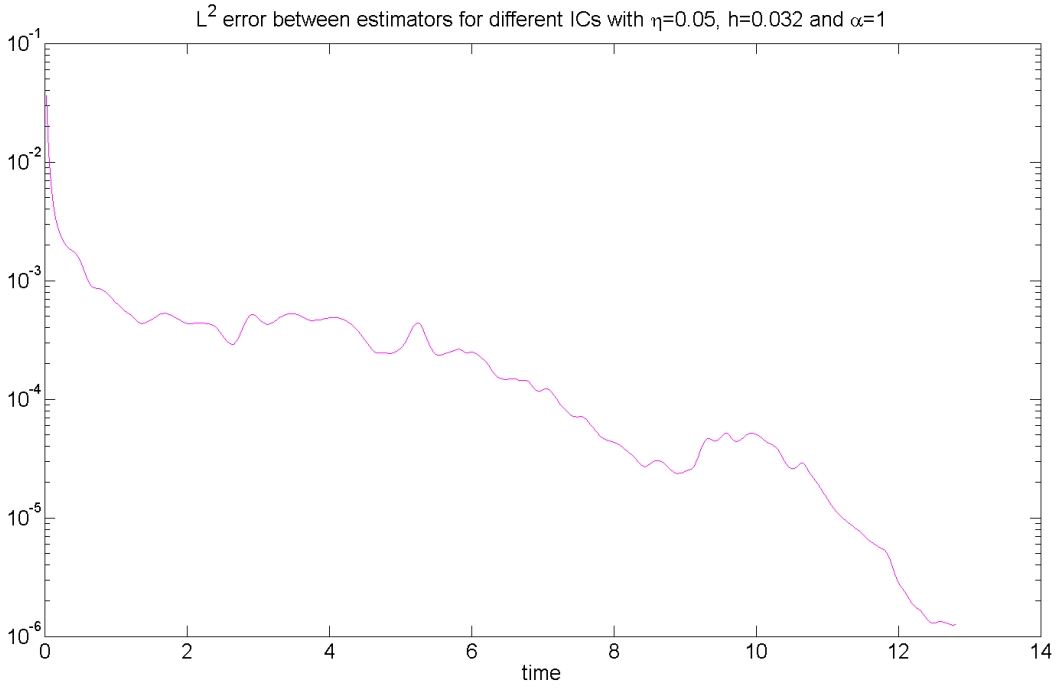


Figure 22: The square of the L^2 difference between the estimators for the two initial conditions for a selection of different wavenumbers. The difference decreases approximately exponentially and would go on decreasing if memory constraints allowed us to assimilate for longer.

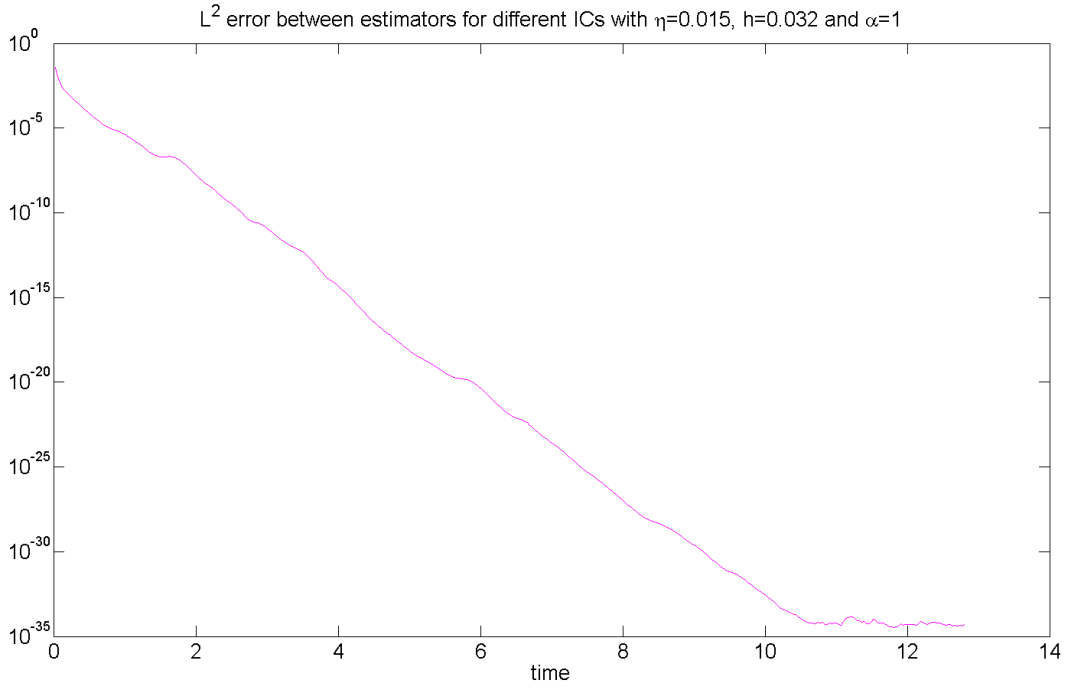


Figure 23: The square of the L^2 difference between the estimators for the two initial conditions. The value of η is smaller than in figure 22 so the difference decreases more rapidly until limited by machine precision.

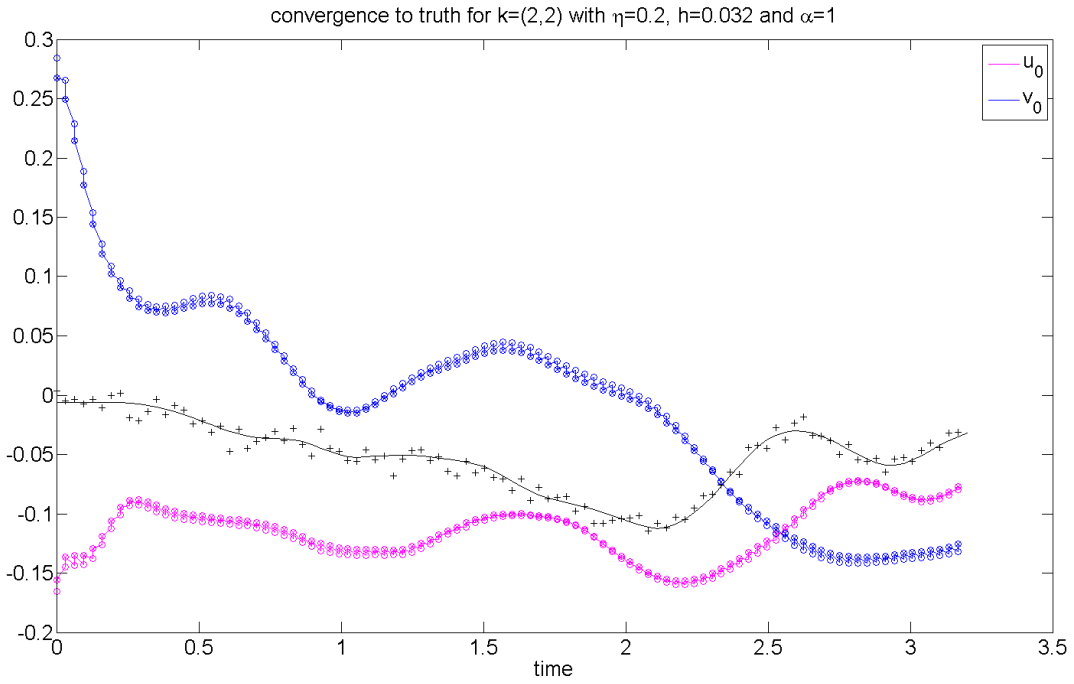


Figure 24: The same plot as figure 20. In this case η was taken too large and the difference between the estimators for the two different initial conditions does not decrease.

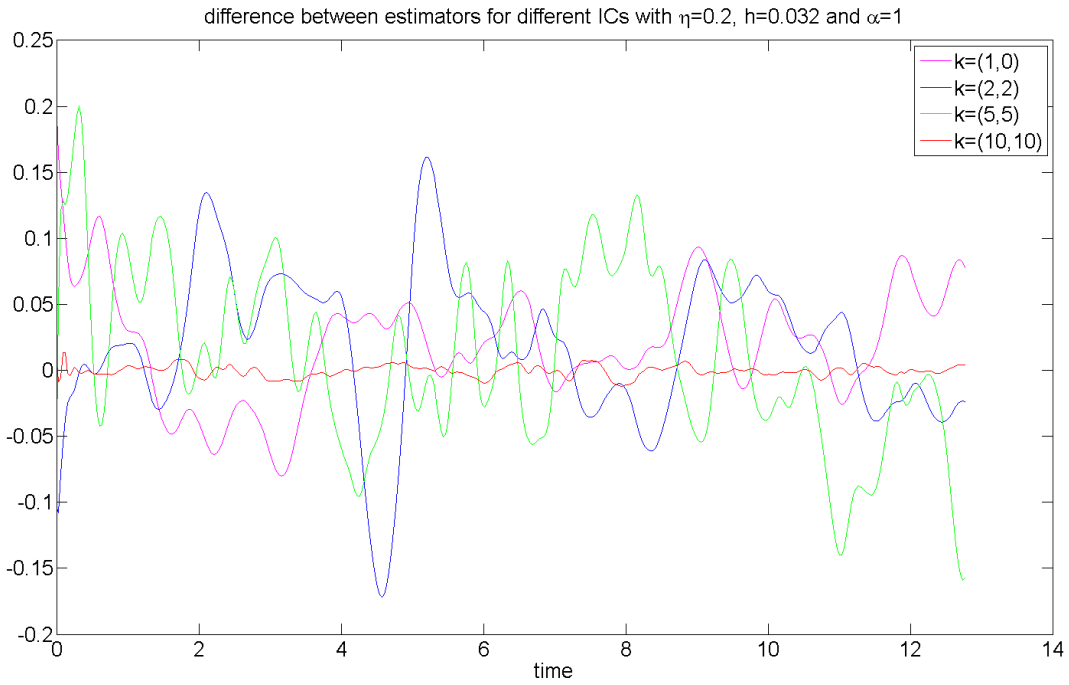


Figure 25: The difference between the estimators for two different initial conditions for a selection of different wavenumbers. Here η is too large so unlike in figure 21 the differences do not decrease.

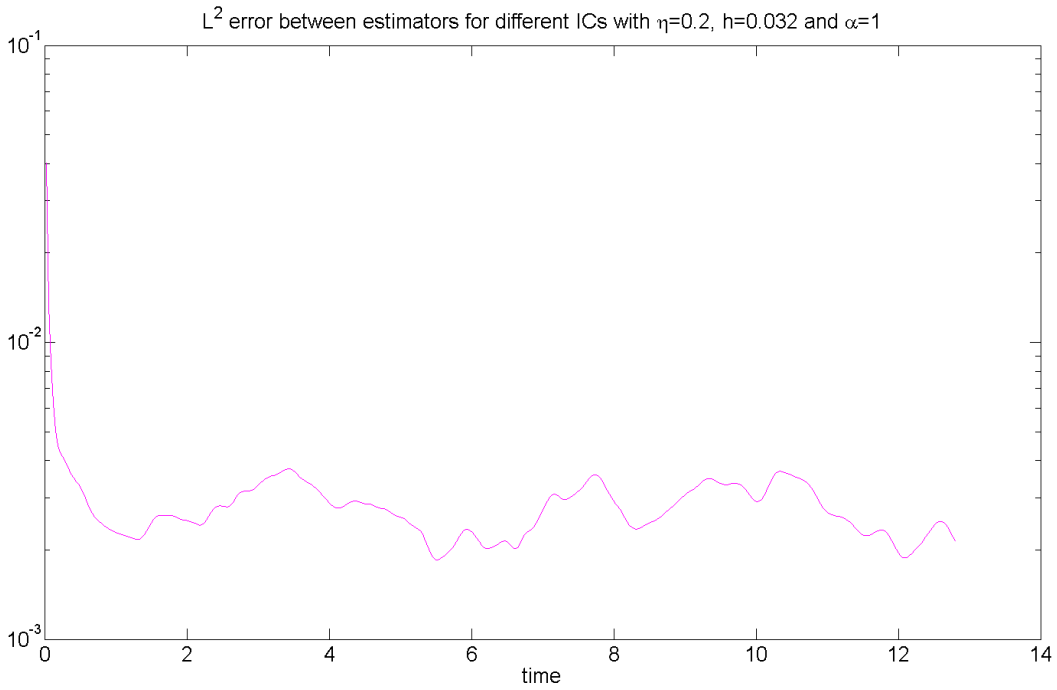


Figure 26: The square of the L^2 difference between the estimators for two different initial conditions does not decrease to an arbitrarily small value like in figure 22 as η is too large.

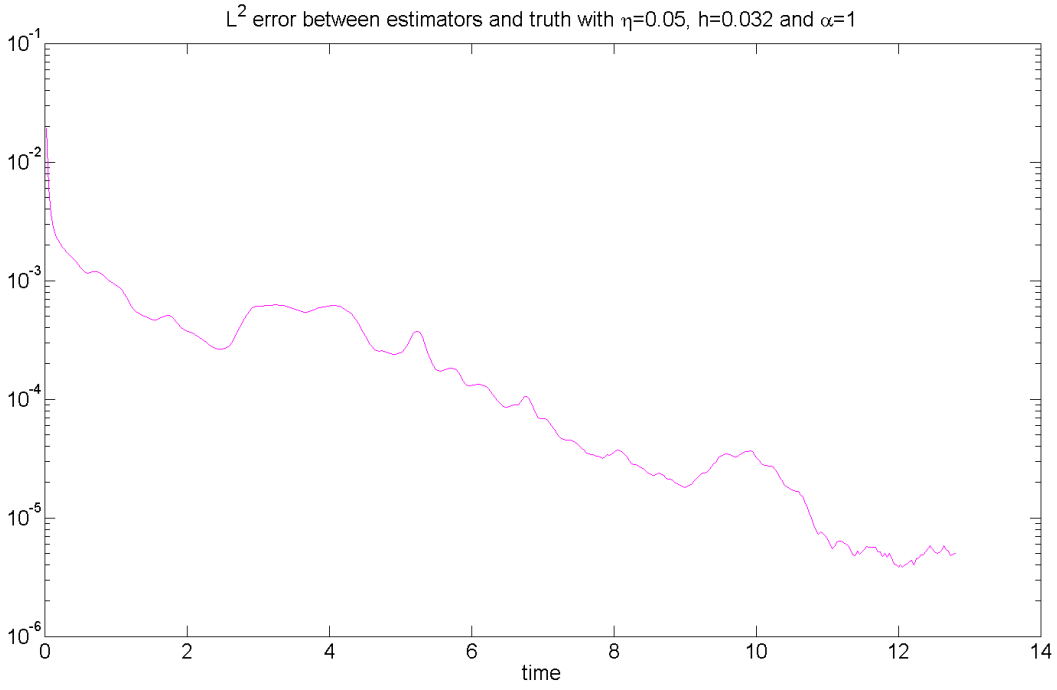


Figure 27: A plot of $e_n^2 := \|\hat{u}_n - u_n\|_{L^2}^2$ against time. The error decreases fairly quickly to the small value of around 10^{-5}.

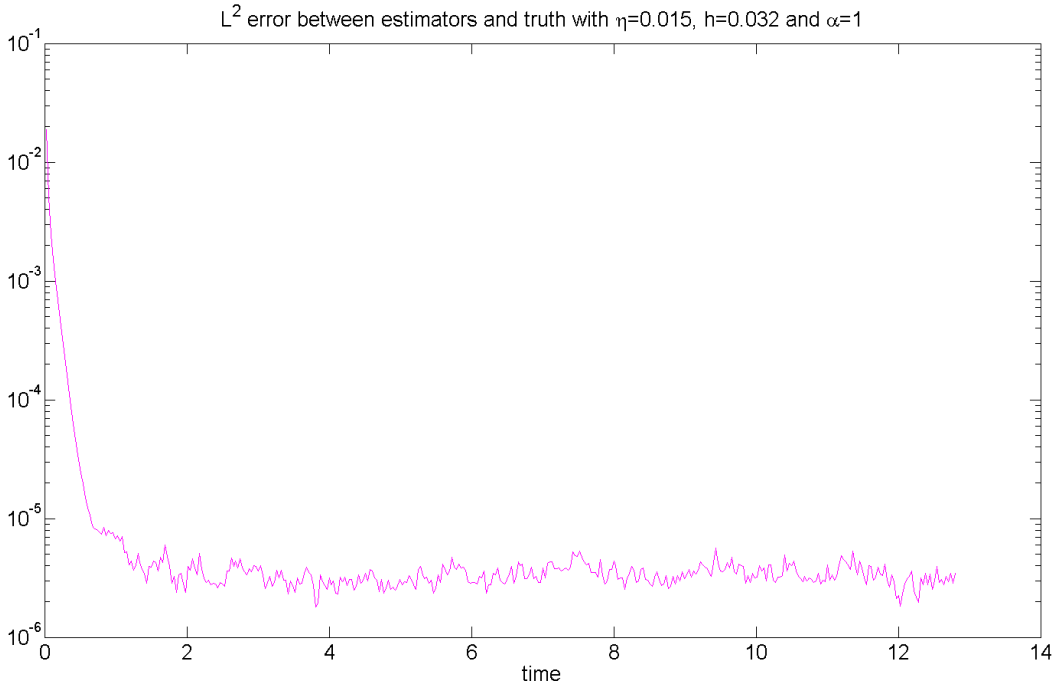


Figure 28: A plot of $e_n^2 := \|\hat{u}_n - u_n\|_{L^2}^2$ against time. As η is smaller than in figure 27 the error decreases more rapidly to around 7×10^{-5}. It can not go arbitrarily small due to noise in the observations.

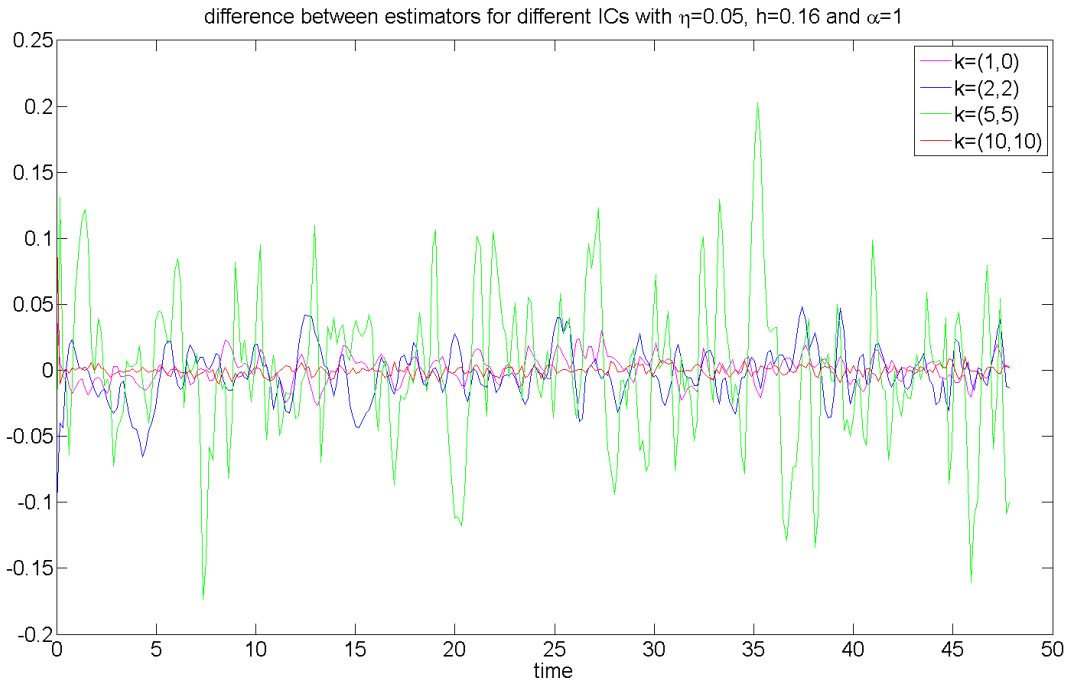


Figure 29: The difference between the estimators for two different initial conditions for a selection of different wavenumbers. As h is 5 times larger than in figure 21 the estimators no longer converge.

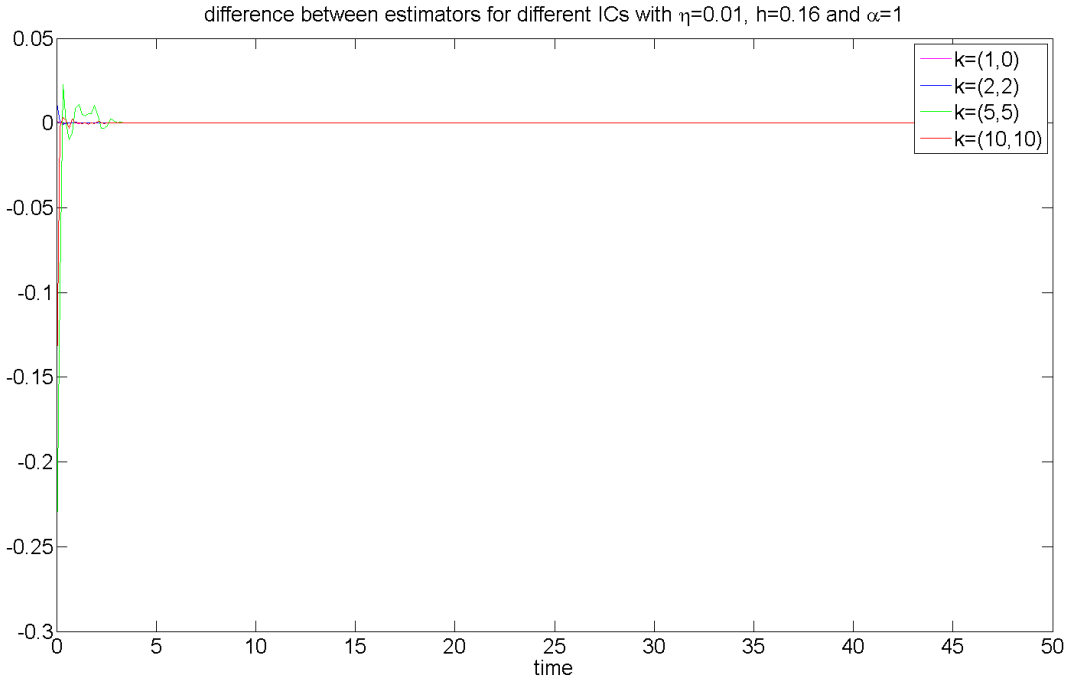


Figure 30: The difference between the estimators for two different initial conditions for a selection of different wavenumbers. Here h is 5 times larger than in figure 21, however we have taken η to be smaller so the estimators still converge.

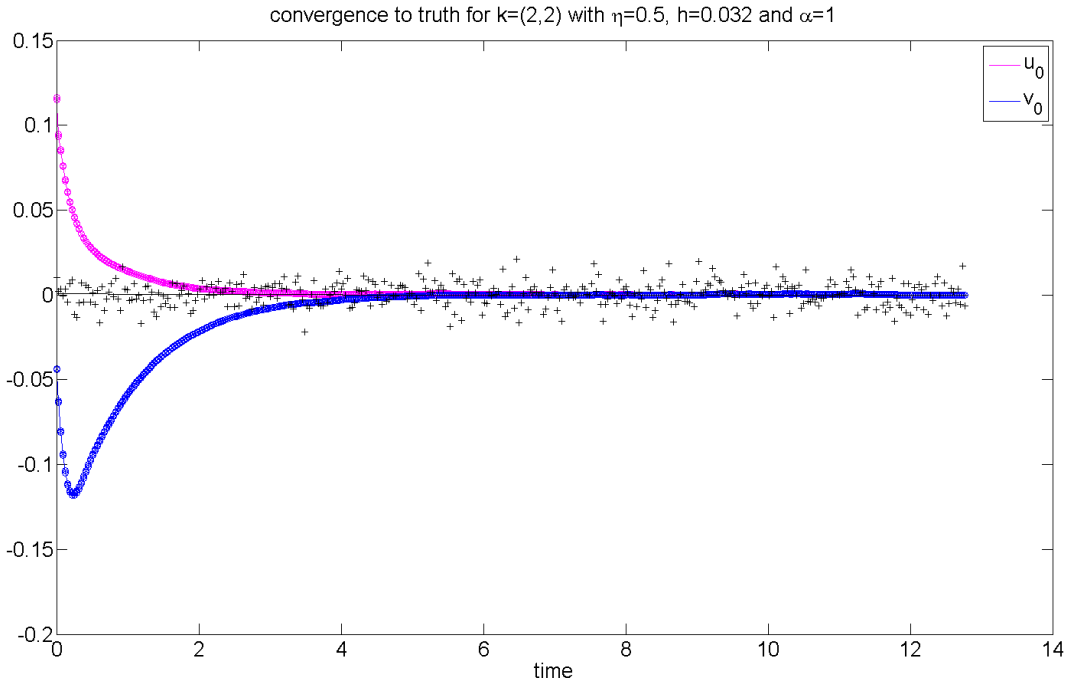


Figure 31: The same plot as figure 12 but for periodic dynamics. Even with the large value of $\eta = 0.5$ the estimators still lock on eventually.

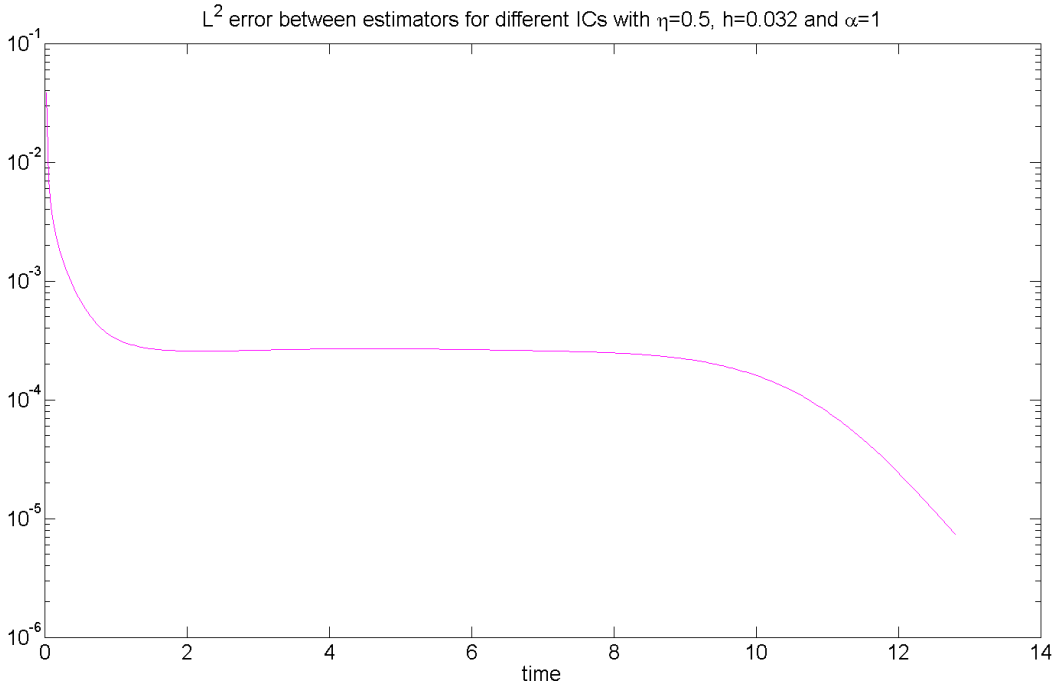


Figure 32: The square of the L^2 difference between the estimators for two different initial conditions for periodic dynamics. The difference decrease rapidly.

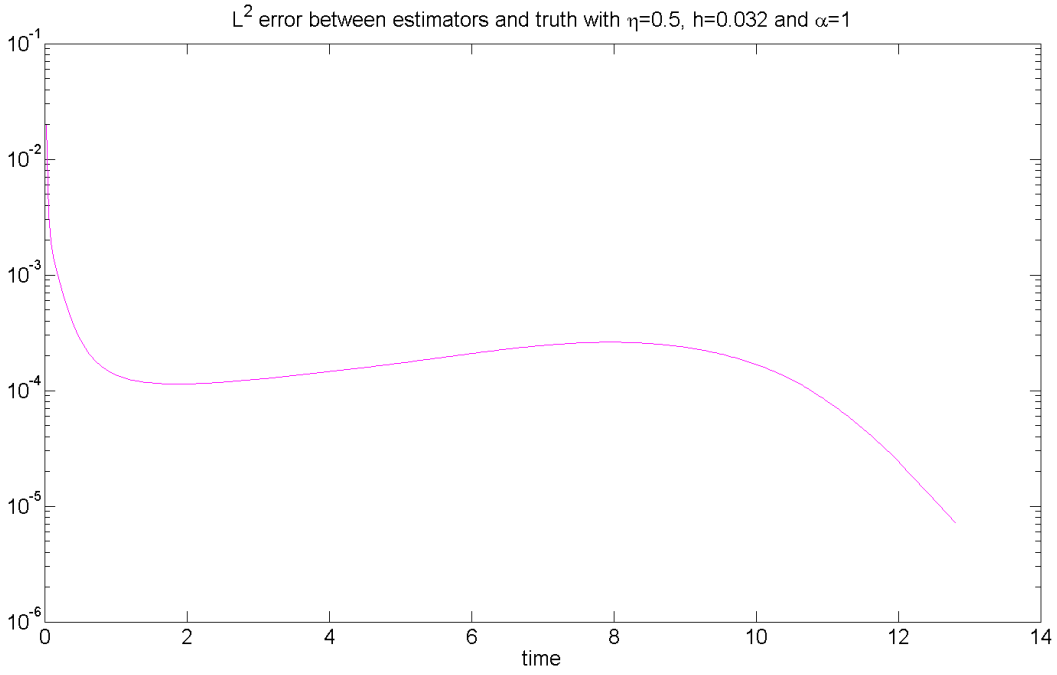


Figure 33: A plot of $e_n^2 := \|\hat{u}_n - u_n\|_{L^2}^2$ against time for periodic dynamics. The error decreases quickly to below 10^{-5}.

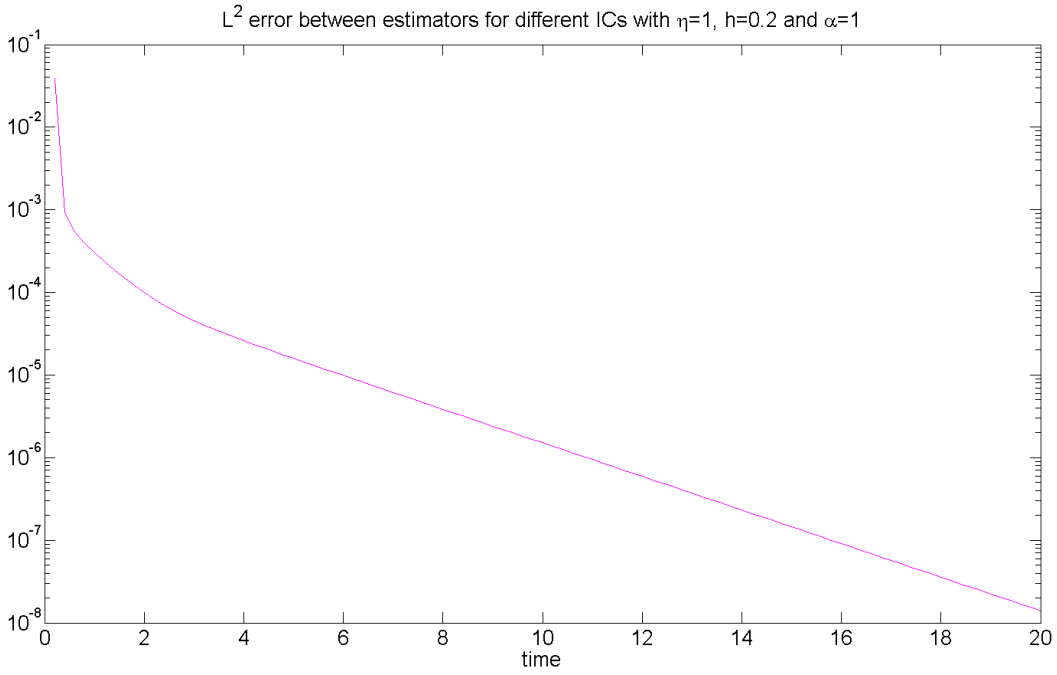


Figure 34: The square of the L^2 difference between the estimators for two different initial conditions for stable dynamics. The difference decreases rapidly despite the large value of η.

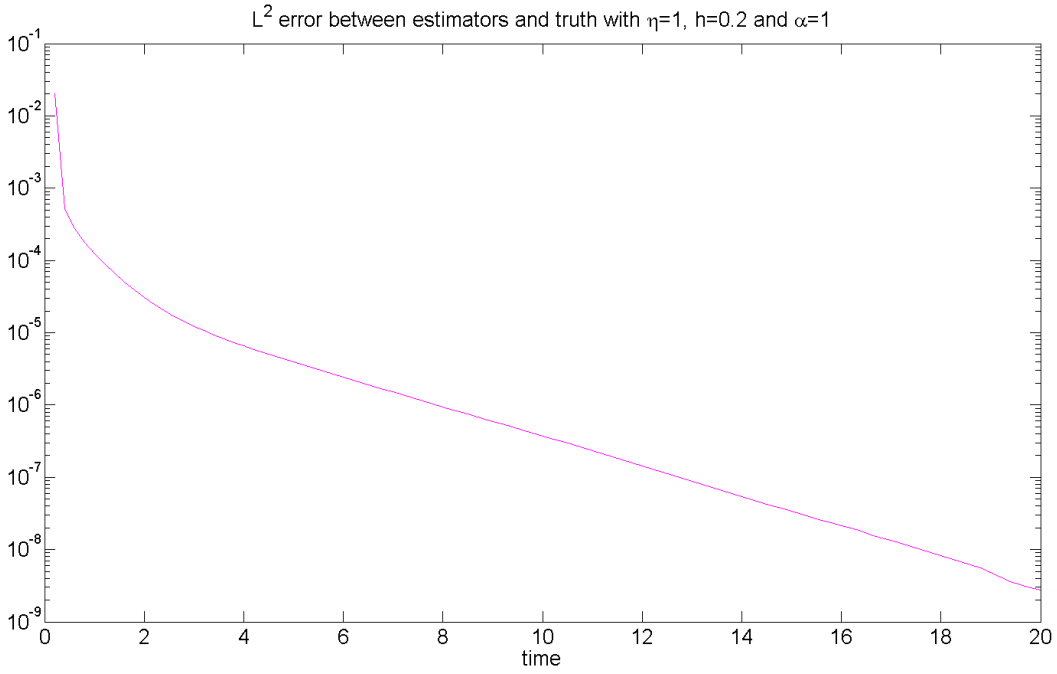


Figure 35: A plot of $e_n^2 := \|\hat{u}_n - u_n\|_{L^2}^2$ against time with stable dynamics. The error decreases quickly to below 10^{-8} despite the large value of η.