Continuous Data Assimilation with Stochastically Noisy Data

Hakima Bessaih,* Eric Olson† and Edriss S. Titi‡

June 5th, 2014

Dedicated to Professor Ciprian Foias on the occasion of his 80^{th} birthday.

Abstract

We analyze the performance of a data-assimilation algorithm based on a linear feedback control when used with observational data that contains measurement errors. Our model problem consists of dynamics governed by the two-dimension incompressible Navier–Stokes equations, observational measurements given by finite volume elements or nodal points of the velocity field and measurement errors which are represented by stochastic noise. Under these assumptions, the data-assimilation algorithm consists of a system of stochastically forced Navier–Stokes equations. The main result of this paper provides explicit conditions on the observation density (resolution) which guarantee explicit asymptotic bounds, as the time tends to infinity, on the error between the approximate solution and the actual solutions which is corresponding to these measurements, in terms of the variance of the noise in the measurements. Specifically, such bounds are given for the limit supremum, as the time tends to infinity, of the expected value of the L^2-norm and of the H^1 Sobolev norm of the difference between the approximating solution and the actual solution. Moreover, results on the average time error in mean are stated.

*University of Wyoming, Department of Mathematics, Dept. 3036, 1000 East University Avenue, Laramie WY 82071, USA; *email:* `bessaih@uwyo.edu`

†Department of Mathematics and Statistics, University of Nevada, Reno, NV 89557, USA. *email:* `ejolson@unr.edu`

‡Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel; *email:* `edriss.titi@weizmann.ac.il`. ALSO: Department of Mathematics and Department of Mechanical and Aerospace Engineering, The University of California, Irvine, CA 92697, USA; *email:* `etiti@math.uci.edu`

Keywords: Determining modes, volumes elements and nodes, continuous data assimilation, downscaling, Navier-Stokes equations, stochastic PDEs, signal synchronization.

Mathematics Subject Classification 2000: Primary 35Q30, 60H15, 60H30; Secondary 93C20, 37C50, 76B75, 34D06.

1 Introduction

Data assimilation is a process by which a time series of observational data for a physical system is used along with the knowledge about the physics which govern the dynamics to obtain an improved estimate of the current state of the system. Applications of data assimilation arise in many fields of geosciences, perhaps most importantly in weather forecasting and hydrology. The classical method of continuous data assimilation, see, *e.g.*, Daley [14], is to insert observational measurements directly into a computer model as the latter is being integrated in time.

A new approach, inspired by ideas from control theory [2], has been proposed in [3] that consists in introducing a feedback control term that forces the approximating solution obtained by data assimilation toward the reference solution that is corresponding to the observations (see also [18, 18] for other applications). This approach admits a general framework of interpolant operators which include operators arising from local volume averages and pointwise nodal measurements. Special attention is given to nodal measurements because they may be likened to the data collected by an array of weather-vane anemometers placed throughout the physical domain. This is unlike previous rigorous work [9, 27, 28, 32, 35] where the observed data is assumed to be the lower Fourier modes of the reference unknown full solution. The method of data assimilation introduced in [3] extends equally to all dissipative dynamical systems and relies on the fact that such dynamical systems possess finite number of determining parameters, such as determining modes, nodes and local volume averages, see, for example, [12, 21, 22, 23, 24, 25, 29, 30, 31] and references therein.

With the exception of Blömker and coauthors [6] for the 3DVAR Gaussian filter, previous theoretical work assumed that the observational measurements are error free. In this paper, we extend the approach of [3] to the case where the observations are contaminated with random errors. This allows us to treat measurement errors for general interpolant observables and, in particular, when the measurements are nodal values with random noise. In this way our analysis may also be seen to extend the work of [6] from Fourier mode measurements to general interpolant observables.

The method of data assimilation studied in this paper can be described mathematically as follows. Let $U(t)$ be a solution trajectory lying on the global attractor of a known dissipative continuous dynamical system and let $u(t)$ be the approximating solution obtained from data assimilation of noisy observational measurements of $U(t)$. Assume that the dynamics of U are governed by an evolution equation of the form

$$\frac{dU}{dt} = F(U) \tag{1}$$

with unknown initial condition $U_0 \in V$ at time t_0. Here $F\colon V \to H$, where V and H are suitable function spaces.

Denote by $\mathcal{O}_h(U(t))$, for $t \geq 0$, the exact observational measurements without error of the true solution U at time t. For two-dimensional physical domains we assume $\mathcal{O}_h\colon V \to \mathbb{R}^D$ to be linear operator where D is of the order $(L/h)^2$ and L is a typical large length scale of the physical domain of interest and h is the observation density or resolution. Denote by $R_h(U(t))$ the interpolation of the observational data, namely,

$$R_h(U(t)) = \mathcal{L}_h \circ \mathcal{O}_h(U(t)),$$

where $\mathcal{L}_h\colon \mathbb{R}^D \to H$ is linear operator. Note that R_h need not be a projection nor does its range need be included in the domain of F. Further assumptions on the general interpolant observable R_h are given in (6) and (7) below.

In the absence of measurement errors the data-assimilation method proposed in [3] would construct the approximating solution u from the interpolant observables $R_h(U(t))$ dynamically as the solution to

$$\frac{du}{dt} = F(u) - \mu(R_h(u) - R_h(U)), \tag{2}$$

with arbitrary initial condition $u(0) = u_0$. Here $\mu > 0$ is a relaxation parameter (nudging), whose value will be determined later, which forces the coarse spatial scales of u, *i.e.*, $R_h(u)$, toward those of the observed data, *i.e.*, $R_h(U)$.

Suppose the exact measurements $\mathcal{O}_h(U(t))$ are subjected to some random errors. Thus, the only observations available for data assimilation are the noisy observations $\tilde{\mathcal{O}}_h(U(t))$ given by

$$\tilde{\mathcal{O}}_h(U(t)) = \mathcal{O}_h(U(t)) + \mathcal{E}(t), \tag{3}$$

where $\mathcal{E}\colon [0,\infty) \to \mathbb{R}^D$ represents the measurement error, for example, due to instrumental errors. That is, in reality the actual interpolated measurements of $U(t)$ contain random errors and are given by

$$\begin{aligned}\tilde{R}_h(U(t)) &= \mathcal{L}_h(\tilde{\mathcal{O}}_h(U(t))) \\ &= \mathcal{L}_h(\mathcal{O}_h(U(t)) + \mathcal{L}_h(\mathcal{E}(t)) = R_h(U(t)) + \xi(t),\end{aligned}$$

where the random vector $\xi(t)$ lies in the range of the interpolant operator R_h.

We will assume that the components of the random errors $\mathcal{E}(t)$ are independent identically distributed of Gaussian type. In particular, the random error $\xi(t)$ will be expressed in terms of a finite-dimensional Wiener process W, white noise in time with an appropriate covariance operator. The precise assumptions will be given in the following sections. We observe that these results could be generalized to other kinds of processes such as a Levy noise.

In this paper we examine the data-assimilation method given by equation (2) when the noise-free interpolant observable $R_h(U(t))$ is replaced by $\tilde{R}_h(U(t))$. In this case, our algorithm for constructing $u(t)$ from the observational measurements $\tilde{\mathcal{O}}_h(U(t))$ is given by the stochastic evolution equation

$$du = (F(u) - \mu R_h(u) + \mu R_h(U))\, dt + \mu\xi dt, \tag{4}$$

with arbitrary initial condition $u(0) = u_0$.

The two-dimensional incompressible Navier–Stokes equations, subject to period boundary conditions, provide a concrete example of a dissipative dynamical system, which we will use as a model problem for our analysis. We find explicit conditions on the observation density or resolution h and relaxation parameter μ which guarantees that the resulting approximate solution $u(t)$ converges, as $t \to \infty$, in some sense, to the exact reference solution $U(t)$ within an error that is determined by the observation density h, the relaxation parameter μ and the variance of the error in the measurements. It is worth mentioning that the application of algorithm (4) to recover solutions to fluid flow problems provides a concrete and justifiable reason for investigating stochastically forced equations such as the Navier-Stokes equations.

In the remainder of this work, the reference solution U will be determined by the two-dimensional incompressible Navier–Stokes equations

$$\begin{cases} \dfrac{\partial U}{\partial t} - \nu\Delta U + (U\cdot\nabla)U = -\nabla p + f \\ \hfill \nabla\cdot U = 0, \end{cases} \tag{5}$$

which describe the motion of an incompressible fluid in $\mathbb{R}^2$. We assume periodic boundary conditions with fundamental domain $\mathcal{D} = [0, L]^2$ and take the initial condition $U(0,x)=U_0(x)$ and the body forcing $f=f(x)$ to be an L-periodic function with zero spatial average. The kinematic viscosity $\nu>0$ is assumed to be known. The unknowns are the velocity vector $U=U(t,x)$ and scalar pressure $p=p(t,x)$. We observe that (5) preserves the L-periodicity and zero spatial average of the initial condition. Thus $\int_{\mathcal{D}} U(t,x)\,dx = 0$, for all $t\geq 0$, provided $\int_{\mathcal{D}} f(x)\,dx = \int_{\mathcal{D}} U_0(x)\,dx = 0$, which we will assume throughout this paper.

For notational convenience we will denote $L^1_{\rm per}$ as simply L^1, and similarly $L^2_{\rm per}$ by L^2 and $H^1_{\rm per}(\mathcal{D})$ by H^1. For $\varphi\in L^1$ we define the average

$$\langle\varphi\rangle = \frac{1}{L^2}\int_{\mathcal{D}}\varphi(x)\,dx,$$

and for every $\mathcal{Z}\subseteq L^1$ we denote $\dot{\mathcal{Z}} = \{\,\varphi\in\mathcal{Z} : \langle\varphi\rangle = 0\,\}$.

In the absence of measurement errors the data assimilation method given by (2) for the two-dimensional incompressible Navier-Stokes equations allows the use of two kinds of linear interpolant observables. The first kind are $R_h\colon [\dot H^1]^2\to[\dot L^2]^2$, which satisfy the approximating identity property

$$\|\varphi - R_h(\varphi)\|_{L^2}^2 \leq c_1 h^2 \|\varphi\|_{H^1}^2, \tag{6}$$

for every $\varphi\in[\dot H^1]^2$; and the second kind of interpolant observables are $R_h\colon [\dot H^2]^2\to[\dot L^2]^2$, which satisfy

$$\|\varphi - R_h(\varphi)\|_{L^2}^2 \leq c_1 h^2 \|\varphi\|_{H^1}^2 + c_2 h^4\|\varphi\|_{H^2}^2, \tag{7}$$

for every $\varphi\in[\dot H^2]^2$. In the presence of measurement errors this same data assimilation method becomes the stochastic differential equation (2) and our analysis needs additional regularity assumptions on R_h for interpolants that satisfy (7). In particular, we assume the range of R_h is in $[\dot H^1]^2$ for interpolants which satisfy (7). This does not result in loss of generality because any interpolant operator whose range is in $[\dot L^2]^2$ can be smoothed so its range is in $[\dot H^s]^2$, for any $s>0$.

The orthogonal projection onto the Fourier modes, with wave numbers k such that $|k|\leq 1/h$, is an example of an interpolant operator which satisfies both approximation properties (6) and (7). A physically relevant interpolant which satisfies (6) is given by the volume elements studied in [30] and [31], see also [25]. Suppose the observations of volume elements $\mathcal{O}_h\colon[\dot H^1]^2\to\mathbb{R}^{2N}$ are given by

$$\mathcal{O}_h(\Phi) = (\bar\varphi_1,\bar\varphi_2,\ldots,\bar\varphi_{2N}) \quad\text{where}\quad \begin{bmatrix}\bar\varphi_{2n-1}\\ \bar\varphi_{2n}\end{bmatrix} = \frac{N}{L^2}\int_{Q_n}\Phi(x)\,dx, \tag{8}$$

for $n = 1, 2, \dots, N$, where the domain $\mathcal{D} = [0, L]^2$ has been divided into $N = K^2$ disjoint equal squares Q_n with sides $h = L/K$. Define $R_h = \mathcal{L}_h \circ \mathcal{O}_h$, where $\mathcal{L}_h : \mathbb{R}^{2N} \to [\dot{L}^2(\mathcal{D})]^2$ with $\mathcal{L}_h(\zeta)$ is the L-periodic function given by

$$\mathcal{L}_h(\zeta)(x) = \sum_{n=1}^{N} \begin{bmatrix} \zeta_{2n-1} \\ \zeta_{2n} \end{bmatrix} \Big(\chi_{Q_n}(x) - \frac{h^2}{L^2} \Big) \quad \text{on} \quad \mathcal{D}. \tag{9}$$

As shown in [30] the interpolant R_h satisfies (6), with $c_1 = 1/6$. Note there are many other choices for $\mathcal{L}_h$ that result in interpolant observables based on volume elements which also satisfy (6). For example, the appendix of [3], which will be summarized in section 2.2 below, presents a smoothed choice for $\mathcal{L}_h$ which results in an $R_h\colon [\dot{H}^1]^2 \to [\dot{H}^2]^2$ which also satisfies (6). In addition, volume elements generalize to any domain $\mathcal{D}$ on which the Bramble–Hilbert inequality holds. An elementary discussion of this inequality in the context of finite element methods appears in Brenner and Scott [7], see also [11, 29, 42].

An interpolant observable $R_h\colon [\dot{H}^2]^2 \to [\dot{H}^2]^2$ which satisfies (7) is obtained, following the ideas of [31], when the observational measurements are given by nodal measurements of the velocity. This corresponds to the data collected from an array of weather-vane anemometers placed throughout the physical domain. Suppose the observations of nodes $\mathcal{O}_h\colon [\dot{H}^2]^2 \to \mathbb{R}^{2N}$ are given by

$$\mathcal{O}_h(\Phi) = (\varphi_1, \varphi_2, \dots, \varphi_{2N}) \quad \text{where} \quad \begin{bmatrix} \varphi_{2n-1} \\ \varphi_{2n} \end{bmatrix} = \Phi(x_n), \tag{10}$$

and $x_n \in Q_n$, for $n = 1, 2, \dots, N$. Here Q_n are, as above, $N = K^2$ disjoint squares with sides $h = L/K$ such that $\mathcal{D} = \cup_{n=1}^{N} Q_n$. Setting $R_h = \mathcal{L}_h \circ \mathcal{O}_h$, where $\mathcal{L}_h$ is the smoothed version of (9) given in section 2.2, results in an interpolant which satisfies (7).

The rest of this paper is organized as follows: section 2 describes the functional setting for our analysis, gives the stochastic setting for our measurement errors and recalls the *a-priori* estimates for classical solutions of the two-dimensional incompressible Navier–Stokes equations that we shall use to obtain our bounds. Section 3 shows the stochastic data assimilation algorithm given by (4) is well-posed. Section 4 states and proves our main results. We end with a few concluding remarks.

2 Preliminaries

The preliminaries have been divided into three subsections. Subsection 2.1 sets up notation and the functional setting we will use in our analysis. Subsection 2.2 gives the stochastic setting for our measurement errors and summarizes the specific details from [3] on the general interpolant observables needed for our analysis. Subsection 2.3 recalls the theory and *a-priori* estimates for classical solutions of the two-dimensional incompressible Navier–Stokes equations needed for our work.

2.1 The Functional Setting

We describe the functional setting which will be used to study the Navier-Stokes equations. We refer to [13, 38, 39, 41] for the main results. Denote by $\mathcal{V}$ all divergence-free $\mathbb{R}^2$ valued L-periodic trigonometric polynomials with zero spatial averages. Let H and V be the closures of $\mathcal{V}$ in $[L^2]^2$ and $[H^1]^2$, respectively. Note that H and V are separable Hilbert spaces with the inner products and norms inherited from $[L^2]^2$ and $[H^1]^2$, respectively. In particular,

$$|u|_H^2 = \langle u, u\rangle, \qquad \text{where} \quad \langle u, v\rangle = \int_{\mathcal{D}} \big(u(x)\cdot v(x)\big)\,dx,$$

and, thanks to the Poincaré inequality (11),

$$\|u\|_V^2 = (\!(u,u)\!), \quad \text{where} \quad (\!(u,v)\!) = \int_{\mathcal{D}} \big(\nabla u(x) : \nabla v(x)\big)\,dx.$$

Denote by H' and V' the dual spaces of H and V, respectively. If we identify H with H', then we have the Gelfand triple $V \subset H \subset V'$ with continuous, compact and dense injections. We denote the dual pairing between $\varphi \in V'$ and $\psi \in V$ by $\langle \varphi, \psi\rangle_{V',V}$. When $\varphi \in H$, we have $\langle \varphi, \psi\rangle_{V',V} = \langle \varphi, \psi\rangle$.

Let Π be the Leray–Helmholtz projector from $[\dot{L}^2]^2$ onto H. The Stokes operator A is given by

$$Au = -\Pi\Delta u \quad \text{for every} \quad u \in D(A) = [\dot{H}^2]^2 \cap V.$$

Note that A is a closed, positive, unbounded self-adjoint operator in H with inverse A^{-1} which is a self-adjoint compact operator on H. By the spectral theorem there exists a sequence $\{\lambda_j\}_{j=1}^{\infty}$ of eigenvalues of the Stokes operator, with $0 < \lambda_1 \le \lambda_2 \le \cdots$, with corresponding eigenvectors $e_j \in D(A)$ such that the set $\{e_j : j \in \mathbb{N}\}$ forms an orthonormal basis in H. Moreover, we have $\lambda_j \sim \lambda_1 j$, as $j \to \infty$, where $\lambda_1 = (2\pi/L)^2$ (*cf.* [13]).

For $\alpha > 0$ we will denote the α-th power of the operator A by A^α and its domain by $D(A^\alpha)$. We have $\|u\|_{D(A^\alpha)}^2 = \sum_{j=1}^{\infty} \lambda_j^{2\alpha} |\langle u, e_j\rangle|^2$. Moreover, it follows that $V = D(A^{1/2})$ with $\|\varphi\|_V = |A^{1/2}\varphi|_H$, for every $\varphi \in V$, and $D(A^{\alpha_1})$ is compactly embedded in $D(A^{\alpha_2})$, for $\alpha_1 > \alpha_2$. Finally, let $D(A^{-\alpha})$ denote the dual of $D(A^\alpha)$.

We have the following Poincaré inequalities:

$$|u|_H^2 \le \lambda_1^{-1} \|u\|_V^2 \quad \text{for} \quad u \in V \tag{11}$$

and

$$\|u\|_V^2 \le \lambda_1^{-1} |u|_{D(A)}^2 \quad \text{for} \quad u \in D(A). \tag{12}$$

Let $b(\cdot,\cdot,\cdot)\colon V \times V \times V \to \mathbb{R}$ be the continuous trilinear form defined as

$$b(u, v, z) = \int_{\mathcal{D}} \big((u(x) \cdot \nabla) v(x)\big) \cdot z(x)\, dx.$$

It is well known that there exists a continuous bilinear operator $B(\cdot,\cdot)\colon V \times V \to V'$ such that $\langle B(u, v), z\rangle_{V',V} = b(u, v, z)$, for all $z \in V$.

Lemma 2.1. *(cf. [13, 38, 39, 41]) Let $u, v, z \in V$. Then*

$$\langle B(u, v), z\rangle_{V',V} = -\langle B(u, z), v\rangle_{V',V} \quad \text{and} \quad \langle B(u, v), v\rangle_{V',V} = 0. \tag{13}$$

Furthermore,

$$|\langle B(u, v), z\rangle_{V',V}| \le \|u\|_{L^4} \|v\|_V \|z\|_{L^4}. \tag{14}$$

Moreover, one can apply the two-dimensional Ladyzhenskaya interpolation inequality (cf. [13])

$$\|u\|_{L^4}^2 \le C_{\mathrm{L}} |u|_H \|u\|_V, \tag{15}$$

to the right-hand side of (14) to obtain

$$|\langle B(u, v), z\rangle_{V',V}| \le C_{\mathrm{L}} |u|_H^{1/2} \|u\|_V^{1/2} \|v\|_V |z|_H^{1/2} \|z\|_V^{1/2}, \tag{16}$$

for functions in V.

We will also make use of the Brézis–Gallouet inequality [8], which may be stated as

$$\|v\|_\infty \le C_{\mathrm{B}} \|v\|_V \left\{1 + \log \frac{|Av|_H^2}{\lambda_1 \|v\|_V^2}\right\}, \tag{17}$$

for functions in $D(A)$.

Lemma 2.2. *(cf. [13, 40]) In the case of periodic boundary conditions the bilinear term has the additional orthogonality property*

$$\langle B(v, v), Av\rangle = 0, \tag{18}$$

for every $v \in D(A)$. In addition, one has

$$\langle B(u, v), Av\rangle + \langle B(v, u), Av\rangle = -\langle B(v, v), Au\rangle, \tag{19}$$

for every $u, v \in D(A)$.

Applying the Leray-Helmholtz projector Π to (5) one obtains the equivalent functional evolution equation

$$\frac{dU}{dt} + \nu AU + B(U,U) = f, \tag{20}$$

with initial condition $U(0) = U_0$, where we assume that $f \in H$ and $U_0 \in V$. Similarly the data-assimilation equation (4) becomes

$$du + (\nu Au + B(u,u))dt = \big(f - \mu \Pi R_h(u - U)\big)dt + \mu dW, \tag{21}$$

where $dW(t) = \Pi\xi(t)dt$ is the noise term.

2.2 The Noise Term

In this section we describe the error term $\mathcal{E}\colon [0,\infty) \to R^{2N}$ that gives rise to the noisy observations $\tilde{\mathcal{O}}_h$ in equation (3) in terms of Brownian motions. We then use the definition $\tilde{R}_h = \mathcal{L}_h \circ \tilde{\mathcal{O}}_h$ to obtain dW in (21).

Following Da Prato and Zabczyk [15] fix a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ on which is defined is a sequence of independent one-dimensional Brownian motions $b_d(t)$, for $d = 1, 2, \ldots, D$, relative to the filtration $(\mathcal{F}_t)$ such that

$$\mathbb{E}(b_d(t)) = 0 \quad \text{and} \quad \mathbb{E}(b_d(t)^2) = t\sigma^2/2 \quad \text{for} \quad t \geq 0.$$

For convenience we shall assume the filtration is complete and right continuous. The measurement errors may now be described by

$$\mathcal{E}(t)dt = (db_1(t), db_2(t), \ldots, db_D(t)). \tag{22}$$

Note that σ is a dimensional constant whose units of measurement must be chosen so the units of measurement for $\mathcal{O}_h(U(t))$ are the same as $\mathcal{E}$. Given a quantity z let $[z]$ represent the units used to measure z. Then $[\mathcal{O}_h] = [\mathcal{E}]$ implies $[\sigma^2] = [\mathcal{O}_h]^2[t]$. In particular, if our observations are velocities as in (8) or (10), we then have $[\mathcal{O}_h] = [L]/[t]$ so that $[\sigma^2] = [L]^2/[t]$.

Writing the linear operator $\mathcal{L}_h\colon \mathbb{R}^D \to [\dot{H}^\alpha]^2$, for $\alpha \geq 0$, as

$$\mathcal{L}_h(\zeta)(\cdot) = \sum_{d=1}^{D} \zeta_d \ell_d(\cdot), \quad \text{where} \quad \zeta \in \mathbb{R}^D \quad \text{and} \quad \ell_d \in [\dot{H}^\alpha]^2, \tag{23}$$

it follows that the noise term in (21) is the Wiener process

$$W(t) = \sum_{d=1}^{D} b_d(t)\gamma_d, \quad \text{where} \quad \gamma_d = \Pi\ell_d. \tag{24}$$

We do not assume γ_d are orthogonal or even linearly independent.

When $\alpha \geq 0$ our assumptions dictate that W is a $[\dot{H}^\alpha]^2$-valued Q-Brownian motion, where $\mathbb{E}(W(t)) = 0$. Following [15] pages 26–27, we have

$$tQ = \mathrm{Cov}(W(t)) = \mathbb{E}\Big(\sum_{d,p=1}^{D} b_d(t)\gamma_d \otimes b_p(t)\gamma_p\Big).$$

Note that Q is a nonnegative and symmetric linear operator with finite trace. In particular, we have

$$\begin{aligned}
\mathrm{Tr}\left[\mathrm{Cov}(W(t))\right] &= \sum_{j=1}^{\infty} \left\langle \mathrm{Cov}(W(t))e_j, e_j\right\rangle \\
&= \sum_{j=1}^{\infty} \mathbb{E}\left(\sum_{d,p=1}^{D} \langle b_d(t)\gamma_d, e_j\rangle \langle b_p(t)\gamma_p, e_j\rangle\right) \\
&= \sum_{j=1}^{\infty} \left(\sum_{d,p=1}^{D} \mathbb{E}\left(b_d(t)b_p(t)\right) \langle \gamma_d, e_j\rangle \langle \gamma_p, e_j\rangle\right) \\
&= t\frac{\sigma^2}{2} \sum_{j=1}^{\infty}\sum_{d=1}^{D} |\langle \gamma_d, e_j\rangle|^2 = t\frac{\sigma^2}{2}\sum_{d=1}^{D} |\gamma_d|_H^2.
\end{aligned}$$

Therefore,

$$\mathrm{Tr}[Q] = \frac{\sigma^2}{2}\sum_{d=1}^{D} |\gamma_d|_H^2 < \infty. \tag{25}$$

We next turn our attention to the specific interpolant observable based on volume elements given by (8) and (9). In this case, setting

$$\ell_{2n-1}(x) = \begin{bmatrix} \chi_{Q_n}(x) - h^2/L^2 \\ 0 \end{bmatrix} \quad \text{and} \quad \ell_{2n-1}(x) = \begin{bmatrix} 0 \\ \chi_{Q_n}(x) - h^2/L^2 \end{bmatrix}, \tag{26}$$

yield, for $n = 1, 2, \ldots, N$, the $D = 2N$ functions needed in (23) and we obtain

Proposition 2.3. *Let $W(t)$ be defined as in* (24)*, where ℓ_d are given by* (26)*, for $d = 1, 2, \ldots, 2N$. Then W is a $[\dot{L}^2]^2$-valued Q-Brownian motion with covariance operator Q that satisfies* $\mathrm{Tr}[Q] \leq \sigma^2 L^2$.

Proof. The calculation

$$\begin{aligned}\mathrm{Tr}[Q] &= \frac{\sigma^2}{2}\sum_{d=1}^{2N}|\gamma_d|_H^2 \le \frac{\sigma^2}{2}\sum_{d=1}^{2N}|\ell_d|_{L^2}^2 = \sigma^2\sum_{n=1}^{N}\int_{\mathcal{D}}\Big|\chi_{Q_n}(x)-\frac{h^2}{L^2}\Big|^2 dx\\ &= \sigma^2\sum_{n=1}^{N}\int_{\mathcal{D}}\Big\{\Big(1-\frac{2h^2}{L^2}\Big)\chi_{Q_n}(x)+\frac{h^4}{L^4}\Big\}dx \le \sigma^2(L^2-h^2)\le\sigma^2L^2\end{aligned}$$

immediately yields the result. □

We now recall the construction of the smooth interpolant observables used for nodal measurements that were constructed in the appendix of [3], and which satisfy (7). Then we derive the estimates needed in our analysis of (21) for the terms resulting from the Itô formula.

Let Q_n, for $n = 1, 2, \ldots, N$, be the $N = K^2$ squares with sides $h = L/K$ described in the introduction such that $\mathcal{D} = \cup_{i=1}^N Q_n$. In particular, we set $\mathcal{J} = \{\,1,\ldots,K\,\}^2$ and for $(i,j)\in\mathcal{J}$ define

$$Q_n = [(i-1)h, ih)\times[(j-1)h, jh), \tag{27}$$

where $n = i + (j-1)K$. Further define

$$\psi_n(x) = \sum_{k\in\mathbb{Z}^2}\chi_{Q_n}(x+kL), \quad \text{for} \quad x\in\mathbb{R}^2, \tag{28}$$

as the L-periodicized characteristic function of Q_n. Note that $\psi_n \in L^2$, and moreover, that $\langle\psi_n^2\rangle = \langle\psi_n\rangle = h^2/L^2$.

To obtain a smoother interpolant let

$$\tilde{\psi}_n(x) = (\rho_{h/10} * \psi_n)(x)$$

be the mollified version of ψ_n, where $\rho_\epsilon(x) = \epsilon^{-2}\rho(x/\epsilon)$, and

$$\rho(z) = \begin{cases} K_0\exp\Big(\dfrac{1}{1-|z|^2}\Big) & |z|<1\\ 0 & |z|\ge 1\end{cases}$$

with

$$(K_0)^{-1} = \int_{|z|<1}\exp\Big(\frac{1}{1-|z|^2}\Big)\,dz.$$

Now setting

$$\ell_{2n-1} = \begin{bmatrix}\tilde{\psi}_n - \langle\tilde{\psi}_n\rangle\\ 0\end{bmatrix} \quad \text{and} \quad \ell_{2n} = \begin{bmatrix}0\\ \tilde{\psi}_n - \langle\tilde{\psi}_n\rangle\end{bmatrix}, \tag{29}$$

for $n = 1, 2, \ldots, N$, yields the $D = 2N$ functions needed in (23) for the definition of $\mathcal{L}_h$. As shown in the appendix of [3], if the observations are given by volume elements, then the resulting interpolant satisfies (6); if the observations are given by nodal points, then the resulting interpolant satisfies (7).

We finish this section with some explicit estimates on the trace of the covariance operator Q, for the Wiener process W given in (24) for the choice of ℓ_n given in (29). Before that, we state two propositions which we shall make use of in the proof as well as in other parts of this paper. Detailed proofs of these propositions appear in the appendix of [3], where the functions $\tilde{\psi}_n$ have been introduced along with their associated interpolant observables.

Proposition 2.4. *Let*

$$\mathcal{U}_n = \{\, x + y : x \in Q_n \text{ and } |y| < \epsilon \,\}, \quad \text{for} \quad n = 1, 2, \ldots N.$$

Then $\{\, \tilde{\psi}_n : n = 1, 2, \ldots, N \,\}$ *is a smooth partition of unity satisfying*

(i) $0 \leq \tilde{\psi}_n(x) \leq 1$ *and* $\operatorname{supp}(\tilde{\psi}_n) \subseteq \mathcal{U}_n + \big(L\mathbb{Z}\big)^2$,

(ii) $\tilde{\psi}_n(x) = 1, \quad \text{for all} \quad x \in \Big(\mathcal{C}_n + \big(L\mathbb{Z}\big)^2\Big)$ *and*

$$\textstyle\sum_{n=1}^{N} \tilde{\psi}_n(x) = 1, \quad \text{for all } x \in \mathbb{R}^2,$$

(iii) $\langle \tilde{\psi}_n \rangle = \big(h/L\big)^2$ *and* $\frac{4}{5}h \leq \big\|\tilde{\psi}_n\big\|_{L^2(D)} \leq \frac{6}{5}h$,

(iv) $\operatorname{supp}(\nabla \tilde{\psi}_n) \subseteq \big(\mathcal{U}_n \setminus \mathcal{C}_n\big) + L\mathbb{Z}^2$,

(v) $|\nabla \tilde{\psi}_n(x)| \leq ch^{-1}$ *and* $|\partial^2 \tilde{\psi}_n(x)/\partial x_i \partial x_j| \leq ch^{-2}, \quad \text{for all } x \in \mathbb{R}^2$,

(vi) $\big\|\nabla \tilde{\psi}_n\big\|_{L^2(\mathcal{D})} \leq c.$

Proposition 2.5. *Let* $\mathcal{K} = \{1 - K, -1, 0, 1, -1 + K\}^2$. *The functions* $\tilde{\psi}_n$ *are nearly orthogonal in the following sense: Suppose* $\alpha, \beta \in \mathcal{J}$ *are such that* $n = \alpha_1 + (\alpha_2 - 1)K$ *and* $m = \beta_1 + (\beta_2 - 1)K$. *Then*

(i) $\displaystyle\int_{\mathcal{D}} \tilde{\psi}_n(x)\tilde{\psi}_m(x)\, dx = 0, \quad \text{for } \beta - \alpha \notin \mathcal{K},$

(ii) $\displaystyle\int_{\mathcal{D}} \big(\nabla \tilde{\psi}_n(x)\big) \cdot \big(\nabla \tilde{\psi}_m(x)\big)\, dx = 0, \quad \text{for } \beta - \alpha \notin \mathcal{K}.$

(iii) $\displaystyle\Big|\int_{\mathcal{D}} \tilde{\psi}_n(x)\tilde{\psi}_m(x)\, dx\Big| \leq (h + 2\epsilon)^2 = \frac{36}{25}h^2, \quad \text{for } \beta - \alpha \in \mathcal{K}.$

(iv) $\left| \int_{\mathcal{D}} \big(\nabla\tilde{\psi}_n(x)\big) \cdot \big(\nabla\tilde{\psi}_m(x)\big)\, dx \right| \leq c, \quad$ *for* $\beta - \alpha \in \mathcal{K}$.

Let us emphasize that the constant c, appearing in Proposition 2.4 parts (v) and (vi), is independent of h. We are now ready to prove the following proposition on the trace of the covariance operator Q for Wiener process W.

Proposition 2.6. *Let $W(t)$ be defined as in* (24) *for the choice of ℓ_n given by equations* (29)*. Then W is a $[\dot{H}^1]^2$-valued Q-Brownian motion with covariance operator Q that satisfies*

$$\mathrm{Tr}[Q] \leq \frac{36}{25}\sigma^2 L^2 \tag{30}$$

and

$$\mathrm{Tr}[A^{1/2} Q A^{1/2}] \leq c\sigma^2 \frac{L^2}{h^2}. \tag{31}$$

Proof. Since $\rho \in C^\infty(\mathbb{R}^2)$ then the range of $\mathcal{L}_h$ is in $[\dot{H}^\alpha]^2$, for every $\alpha \geq 0$, and in particular for $\alpha = 1$. Therefore, W is an $[\dot{H}^1]^2$-valued Q-Brownian motion. From (25), Proposition 2.4 part (iii) and Proposition 2.5 part (iii) we estimate

$$\begin{aligned} \mathrm{Tr}[Q] &= \frac{\sigma^2}{2}\sum_{d=1}^{2N} |\gamma_d|_H^2 \leq \frac{\sigma^2}{2}\sum_{d=1}^{2N} \|\ell_d\|_{L^2}^2 = \sigma^2 \sum_{n=1}^{N} \|\tilde{\psi}_n - \langle\tilde{\psi}_n\rangle\|_{L^2}^2 \\ &= \sigma^2 \sum_{n=1}^{N} L^2\big(\langle\tilde{\psi}_n^2\rangle - \langle\tilde{\psi}_n\rangle^2\big) \leq \sigma^2 N L^2\Big(\frac{36h^2}{25L^2} - \frac{h^4}{L^4}\Big) \leq \frac{36}{25}\sigma^2 L^2. \end{aligned}$$

Since in the periodic case we have $\|\Pi\varphi\|_V \leq \|\nabla\varphi\|_{L^2}$ for every $\varphi \in \dot{H}^1$, then Similarly estimate

$$\begin{aligned} \mathrm{Tr}[A^{1/2} Q A^{1/2}] &= \frac{\sigma^2}{2}\sum_{d=1}^{2N} \|\gamma_d\|_V^2 \leq \sigma^2 \sum_{n=1}^{N} |\nabla\tilde{\psi}_n|_{L^2}^2 \\ &\leq c\sigma^2 N = c\sigma^2 \frac{L^2}{h^2}, \end{aligned}$$

where Proposition 2.4 part (vi) has been used in the final inequality. $\square$

2.3 The Deterministic Navier-Stokes Equations

The deterministic two-dimensional incompressible Navier-Stokes equations, subject to periodic boundary conditions, are well-posed and possess a compact finite-dimensional global attractor. Specifically, the following result can be found in [13], [21], [38] and [39].

Theorem 2.7. *Let $U_0 \in V$ and $f \in H$. Then* (20) *has a unique strong solution that satisfies*

$$U \in C([0,T];V) \cap L^2([0,T];D(A)), \quad \text{for any} \quad T > 0.$$

Moreover, the solution U depends continuously on U_0 in the V norm.

Let us denote by G the Grashof number

$$G = \frac{|f|_H}{\nu^2 \lambda_1}, \tag{32}$$

which is a dimensionless physical parameter. We now give bounds on solutions U of (20) that will be used in our later analysis. With the exception of inequality (35) these estimates appear in the references listed above. The improved estimate in (35) is given in [20].

Theorem 2.8. *Let $T > 0$, and let G be the Grashof number given by* (32)*. There exists a time t_0, which depends on U_0, such that for all $t \geq t_0$ we have*

$$|U(t)|_H^2 \leq 2\nu^2 G^2 \quad \text{and} \quad \int_t^{t+T} \|U(\tau)\|_V^2 d\tau \leq 2(1 + T\nu\lambda_1)\nu G^2. \tag{33}$$

Furthermore, we also have

$$\|U(t)\|_V^2 \leq 2\nu^2 \lambda_1 G^2 \quad \text{and} \quad \int_t^{t+T} |AU(\tau)|_H^2 d\tau \leq 2(1 + T\nu\lambda_1)\nu\lambda_1 G^2. \tag{34}$$

Moreover,

$$|AU(t)|_H^2 \leq c\nu^2 \lambda_1^2 (1 + G)^4. \tag{35}$$

3 The Data Assimilation Algorithm

Let U be the strong solution of (20) given by Theorem 2.7, and let R_h be an interpolation operator satisfying either (6) or (7). Suppose the only knowledge we have about U is from the noisy observational measurements $R_h(U(t)) + \xi(t)$, that have been continuously recorded for times $t \in [0,T]$. Our goal in this section is to show that the data assimilation algorithm given by equations (21) for computing the approximating solution u are well posed.

The proof combines the well-posedness results for the noise-free data data assimilation equations (2), appearing in [3], with techniques from [17]. Similar results can be found in [15] for stochastically forced partial differential equations. Namely, we have the following two theorems.

Theorem 3.1. *Suppose U is the strong solution of* (20) *given by Theorem 2.7, where $U_0 \in V$ and $f \in H$. Moreover, assume $R_h \colon [\dot{H}^1]^2 \to [\dot{L}^2]^2$ satisfies* (6) *and that $2\mu c_1 h^2 \leq \nu$. Then for any $u_0 \in H$ and $T > 0$, there exists a unique stochastic process solution u of equation* (21) *in the following sense: $\mathbb{P}$-a.s.*

$$u \in C([0,T]; H) \cap L^2([0,T]; V)$$

and

$$\begin{aligned} \langle u(t), \varphi\rangle + \int_0^t \langle u(\tau), A\varphi\rangle d\tau - \int_0^t \langle B(u(\tau), \varphi), u(\tau)\rangle d\tau = \langle u_0, \varphi\rangle \\ + \int_0^t \langle f(\tau), \varphi\rangle d\tau - \mu \int_0^t \langle R_h(u(\tau) - U(\tau)), \varphi\rangle d\tau + \langle W(t), \varphi\rangle \end{aligned} \tag{36}$$

for all $t \in [0,T]$ and for all $\varphi \in D(A)$. Moreover,

$$\mathbb{E}\left(\sup_{0\leq t\leq T} |u(t)|_H^2 + \nu \int_0^T \|u(t)\|_V^2 dt\right) < \infty. \tag{37}$$

Theorem 3.2. *Suppose U is the strong solution of* (20) *given by Theorem 2.7, where $U_0 \in V$ and $f \in H$. Moreover, assume $R_h \colon [\dot{H}^2]^2 \to [\dot{H}^1]^2$ satisfies* (7) *and that $2\mu h^2 \max(c_1, \sqrt{c_2}) \leq \nu$. Then for any $u_0 \in V$ and $T > 0$, the stochastic process solution of equation* (21), *given in the previous theorem is such that $\mathbb{P}$-a.s.*

$$u \in C([0,T]; V) \cap L^2([0,T]; D(A))$$

and

$$\mathbb{E}\left(\sup_{0\leq t\leq T} \|u(t)\|_V^2 + \nu \int_0^T |Au(t)|_H^2 dt\right) < \infty. \tag{38}$$

Proof of Theorem 3.1. The proof of this theorem is based on a pathwise argument. We proceed along the lines of [17] in which a similar proof appears except without the function U and the additional linear term. Consider the auxiliary process z which satisfies

$$dz + \nu Az dt = \mu dW, \qquad z(0) = 0. \tag{39}$$

It is known, see [15], that

$$z(t) = \mu \int_0^t e^{-\nu A(t-\tau)} dW(\tau)$$

is a stationary $D(A^{1/2})$-valued ergodic solution to (39) with continuous trajectories. In particular, we have

$$\mathbb{E}\,\|z(t)\|^2_{D(A^{1/2})} \leq \frac{\mu^2\sigma^2}{2\nu}\,\mathrm{Tr}[Q].$$

This estimate may be obtained by writing

$$z = \sum_{j=1}^{\infty} z_j e_j \quad \text{and} \quad W = \sum_{j=1}^{\infty} W_j e_j = \sum_{j=1}^{\infty} \Big(\sum_{d=1}^{D} \gamma_{d,j} b_d \Big) e_j$$

where $z_j(t) = \langle z(t), e_j \rangle$ and $\gamma_{d,j} = \langle \gamma_d, e_j \rangle$.

Then,

$$z_j(t) = \mu \int_0^t e^{-\nu\lambda_j(t-\tau)} dW_j(\tau) = \mu \sum_{d=1}^{D} \gamma_{d,j} \int_0^t e^{-\nu\lambda_j(t-\tau)} db_d(\tau).$$

Using the independence of the b_d's and the Itô isometry, it follows that

$$\begin{aligned}
\mathbb{E}\,|z_j(t)|^2 &= \mu^2 \sum_{d=1}^{D} \gamma_{d,j}^2 \,\mathbb{E} \left| \int_0^t e^{-\nu\lambda_j(t-\tau)} db_d(\tau) \right|^2 \\
&= \frac{\mu^2\sigma^2}{2} \sum_{d=1}^{D} \gamma_{d,j}^2 \int_0^t e^{-2\nu\lambda_j(t-\tau)} d\tau \leq \frac{\mu^2\sigma^2}{4\nu\lambda_j} \sum_{d=1}^{D} \gamma_{d,j}^2.
\end{aligned}$$

Therefore, provided $2\alpha \leq 1$ we have

$$\begin{aligned}
\mathbb{E}\,\|z\|^2_{D(A^\alpha)} &= \sum_{j=1}^{\infty} \lambda_j^{2\alpha}\,\mathbb{E}\,|z_k|^2 \\
&\leq \frac{\mu^2\sigma^2}{4\nu} \sum_{d=1}^{D} \sum_{j=1}^{\infty} \frac{1}{\lambda_j^{1-2\alpha}} \gamma_{d,j}^2 \leq \frac{\mu^2}{2\nu\lambda_1^{1-2\alpha}}\,\mathrm{Tr}[Q].
\end{aligned}$$

Now, using the change of variable $\tilde{u} = u - z$, we find that $\tilde{u}$ is solution of the (random) differential equation

$$\frac{d}{dt}\tilde{u} + \nu A\tilde{u} + B(\tilde{u} + z, \tilde{u} + z) + \mu\Pi\, R_h(\tilde{u} + z) = \tilde{f}, \tag{40}$$

where $\tilde{f} = f + \mu\Pi\, R_h(U)$ and $\tilde{u}(0) = \tilde{u}_0 = u_0$.

Theorem 2.7 implies that $U \in C([0,T]; V)$. Hence, using (6) and the Poincaré inequality

$$|\Pi\, R_h(U)|_H \leq \|U - R_h(U)\|_{L^2} + |U|_H \leq \big(\sqrt{c_1}h + \lambda_1^{-1/2}\big)\|U\|_V$$

which implies that $\Pi\, R_h(U) \in C([0,T];H)$. Therefore $\tilde{f} \in C([0,T];H)$.

For every $\omega \in \Omega$, there exists a unique weak solution $\tilde{u}$ of equation (40) and it depends continuously, in $C([0,T];H) \cap L^2(0,T;V)$ norms, for any given $T > 0$, on the initial condition $\tilde{u}_0 = u_0$ in H. A full rigorous proof of this statement is very long, but at the same time it is very classical. Similar proofs are detailed, for instance in [17] for the stochastically forced Navier–Stokes equations and in [13] or [41] in the case of the classical Navier-Stokes equations (*i.e.*, when $z = 0$). The rigorous proof is based on the Galerkin approximation procedure and then passing to the limit using the appropriate compactness theorems. We state the necessary *a-priori* estimates here.

Take the inner product of equation (40) by $\tilde{u}$

$$\left\langle \frac{d\tilde{u}}{dt}, \tilde{u} \right\rangle + \nu \langle A\tilde{u}, \tilde{u}\rangle = -\langle B(\tilde{u}+z, \tilde{u}+z), \tilde{u}\rangle - \mu \langle \Pi\, R_h(\tilde{u}+z), \tilde{u}\rangle + \langle \tilde{f}, \tilde{u}\rangle.$$

Using Lemma 2.2 and Young's inequality, we get that

$$\begin{aligned} |\langle B(\tilde{u}+z, \tilde{u}+z), \tilde{u}\rangle| &= |\langle B(\tilde{u}, \tilde{u}), z\rangle + \langle B(z, \tilde{u}), z\rangle| \\ &\le C_{\mathrm{L}} \left(\|\tilde{u}\|_{L^4}^2 \|z\|_V + \|\tilde{u}\|_V \|z\|_{L^4}^2 \right) \\ &\le \frac{\nu}{2}\|\tilde{u}\|_V^2 + \frac{C_{\mathrm{L}}^2}{\nu} \left(|\tilde{u}|_H^2 \|z\|_V^2 + |z|_H^2 \|z\|_V^2 \right). \end{aligned} \tag{41}$$

For the other term we apply the Cauchy-Schwarz, Young's and Poincaré inequalities along with the approximation property (6) to obtain

$$\begin{aligned} -\mu \langle \Pi\, R_h(\tilde{u}+z), \tilde{u}\rangle &= -\mu \langle\, R_h(z), \tilde{u}\rangle - \mu \langle\, R_h(\tilde{u}), \tilde{u}\rangle \\ &\le \mu |z - R_h(z)|_H |\tilde{u}|_H + \mu |z|_H |\tilde{u}|_H \\ &\qquad + \mu \langle \tilde{u} -\, R_h(\tilde{u}), \tilde{u}\rangle - \mu |\tilde{u}|_H^2 \\ &\le \frac{\mu}{2}|\tilde{u}|_H^2 + \frac{\mu}{2}\left(|z|_H^2 + |z - R_h(z)|_H^2 \right) \\ &\qquad + \frac{\mu}{2}|\tilde{u} -\, R_h(\tilde{u})|_H^2 + \frac{\mu}{2}|\tilde{u}|_H^2 - \mu |\tilde{u}|_H^2 \\ &\le \frac{\mu}{2}\left(\lambda_1^{-1} + c_1 h^2 \right) \|z\|_V^2 + \frac{c_1 h^2 \mu}{2} \|\tilde{u}\|_V^2. \end{aligned}$$

Also,

$$\langle \tilde{f}, \tilde{u}\rangle \le |\tilde{f}|_H |\tilde{u}|_H \le \lambda_1^{-1/2} |\tilde{f}|_H \|\tilde{u}\|_V \le \frac{1}{\nu\lambda_1}|\tilde{f}|_H^2 + \frac{\nu}{4}\|\tilde{u}\|_V^2.$$

Hence, since we chose h and μ such that $\nu \ge 2c_1 h^2 \mu$, we get that

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}|\tilde{u}|_H^2 + \frac{\nu}{2}\|\tilde{u}\|_V^2 &\le \frac{\mu}{2}\left(\lambda_1^{-1} + c_1 h^2 \right) \|z\|_V^2 \\ &\quad + \frac{C_{\mathrm{L}}^2}{\nu}\left(|\tilde{u}|_H^2 \|z\|_V^2 + |z|_H^2 \|z\|_V^2 \right) + \frac{1}{\nu\lambda_1}|\tilde{f}|_H^2. \end{aligned}$$

From the above we have

$$\frac{d}{dt}|\tilde{u}|_H^2 + \nu\|\tilde{u}\|_V^2 \le \mu\big(\lambda_1^{-1} + c_1 h^2\big)\|z\|_V^2 + \frac{2}{\nu\lambda_1}|\tilde{f}|_H^2 + \frac{2C_{\mathrm{L}}^2}{\nu}\|z\|_V^2|\tilde{u}|_H^2 + \frac{2C_{\mathrm{L}}^2}{\nu}|z|_H^2\|z\|_V^2.$$

Since $\tilde{f}$ and z are in $C([0,T];V)$, then by Gronwall's Lemma and the previous estimates on z and $\tilde{f}$ to get that

$$\sup_{t\in[0,T]}|\tilde{u}(t)|_H^2 \le C, \quad \text{and} \quad \int_0^T \|\tilde{u}(\tau)\|_V^2 d\tau \le C.$$

Since, $u = \tilde{u} + z$, we deduce from the properties of z that $\mathbb{P}$-a.s.

$$u \in C([0,T];H)\bigcap L^2([0,T];V).$$

The rigorous justification to the fact that the process u is adapted follows from the limiting procedure of adapted processes (via the Galerkin approximation).

Let us sketch the proof of (37). As usual, these calculations are performed in a first step on the Galerkin approximation and then in a second step the estimates for the solution u are obtained by a limiting procedure. But for simplicity, we only sketch them for u. For similar estimates, see [5, 10]

Using Itô formula to $|u(t)|_H^2$, where $u(t)$ is solution of (21), we get that

$$d|u(t)|_H^2 = 2\langle u(t), du(t)\rangle + \mu^2\,\mathrm{Tr}[Q].$$

Then, using assumption (13) and integrating over $(0,t)$, we get

$$\begin{aligned}\sup_{0\le t\le T}|u(t)|_H^2 + 2\nu\int_0^T \|u(\tau)\|_V^2 d\tau &= |u_0|_H^2 + \int_0^T \langle f, u(\tau)\rangle d\tau \\ &- 2\mu\int_0^T \langle R_h(u(\tau) - U(\tau)), u(\tau)\rangle d\tau \\ &+ 2\mu\sup_{0\le t\le T}\int_0^t \langle u(\tau), dW(\tau)\rangle + \mu^2 T\,\mathrm{Tr}[Q]. \end{aligned} \tag{42}$$

Using the Burkholder-Gundy-Davis inequality (cf. [15]) on the martingale term in the right-hand side of (42)

$$\begin{aligned} 2\mu\,\mathbb{E}\left(\sup_{0\le t\le T}\int_0^t \langle u(\tau), dW(\tau)\rangle\right) &\le 2\mu\sqrt{\mathrm{Tr}[Q]}\,\mathbb{E}\sqrt{\int_0^t |u(\tau)|_H^2 d\tau} \\ &\le 2\mu\,\mathbb{E}\sup_{0\le t\le T}|u(t)|_H\sqrt{T\,\mathrm{Tr}[Q]} \\ &\le \frac{1}{2}\,\mathbb{E}\sup_{0\le t\le T}|u(t)|_H^2 + 2\mu^2 T\,\mathrm{Tr}[Q]. \end{aligned}$$

On the other hand, using the approximation (6) and the Poincaré inequality, we get that

$$
\begin{aligned}
&-2\mu\langle R_h(u-U),u\rangle \\
&\qquad\qquad \le 2\mu|u-R_h(u)|_{L^2}|u|_H - 2\mu|u|_H^2 + 2\mu|R_h(U)|_{L^2}|u|_H \\
&\qquad\qquad \le 2\mu|u-R_h(u)|_{L^2}^2 - \mu|u|_H^2 + 2\mu|R_h(U)|_{L^2}^2 \\
&\qquad\qquad \le 2\mu c_1 h^2\|u\|_V^2 - \mu|u|_H^2 + 2\mu\left(|U-R_h(U)|_{L^2}^2 + |U|_H^2\right) \\
&\qquad\qquad \le 2\mu c_1 h^2\|u\|^2 - \mu|u|_H^2 + 2\mu(c_2h^2+\lambda_1^{-1})\|U\|_V^2.
\end{aligned}
$$

Since $2\mu c_1 h^2 \le \nu$, combining the previous estimates with the Gronwall lemma one obtains

$$
\mathbb{E}\left(\sup_{0\le t\le T}|u(t)|_H^2\right) \le C\left(|u_0|_H, T, \mu, \nu, h, \lambda_1, \mathrm{Tr}[Q], \|U\|_V\right).
$$

Using again this estimate in (42) finishes the proof. $\square$

Proof of Theorem 3.2. Similar arguments as used in Theorem 3.1 may be used to prove this theorem. $\square$

4 Main Results

Our goal is to prove that u, the solution of (21), approximates the true solution U of (20), when $t\to\infty$, to within some tolerance depending on the error in the observations. Upon setting $v = U - u$ this is equivalent to showing that v is small. Note that from (20) and (21) the evolution of v is governed by the equation

$$
\begin{aligned}
dv + \big[\nu Av + B(U,v) + B(v,U) - B(v,v)\big]\,dt& \\
&= -\mu\Pi\, R_h(v)dt + \mu\Pi\, dW,
\end{aligned}
\tag{43}
$$

with $v_0 \in V$ is chosen arbitrary. We now look for conditions on h and μ such that the feedback term, R_h, on the right-hand side of this equation, which is stabilizing the coarse scales, together with the viscous term, which is stabilizing the fine scales, controls the growth of v, which is due to the unstable nature of this kind of nonlinear dynamical system.

Section 4.1 studies interpolant observables which satisfy (6) and, in particular, those come from finite volume elements. Section 4.2 studies interpolant observables which satisfy (7) and, in particular, those which come from nodal observations.

4.1 Observations of Volume Elements

This section first proves a general theorem on interpolant observables that satisfy (6). This result is then applied to obtain explicit estimates when the observational measurements arise, for example, from volume elements. The same result holds for other kinds of observables that satisfy (6), such as Fourier modes and the interpolants investigated in [29].

Theorem 4.1. *Assume that U is a strong solution of* (20)*, that R_h satisfies assumption* (6) *and that W is a $[\dot{L}^2]^2$-valued Q-Brownian motion. Assume that μ is large enough, and h is small enough such that*

$$\frac{1}{h^2} \geq \frac{2c_1\mu}{\nu} \geq 8c_1 C_{\mathrm{L}}^2 \lambda_1 G^2.$$

where c_1, C_L are respectively given in (6) *and* (15)*. Then the solution u of* (21) *given by Theorem 3.1 satisfies*

$$\limsup_{t\to\infty} \mathbb{E}(|u(t) - U(t)|_H^2) \leq \mu \operatorname{Tr}[Q].$$

Moreover,

$$\limsup_{t\to\infty} \frac{\nu}{T}\int_t^{t+T} \mathbb{E}(\|u(\tau) - U(t)\|_V^2) d\tau \leq \Big(\frac{1}{T} + \mu\Big)\mu \operatorname{Tr}[Q]. \tag{44}$$

Proof. In this proof we focus on the time interval $[t_0, \infty)$, where t_0 is given in Theorem 2.8. Using the Itô formula on $|v(t)|_H^2$ we obtain

$$d|v|_H^2 = 2\langle v, dv\rangle + \mu^2 \operatorname{Tr}[Q]\, dt.$$

Substituting for dv and applying the orthogonality property (13) yields

$$\begin{aligned} d|v|_H^2 + 2\nu\|v\|_V^2 dt = &- 2\langle v, B(v,U)\rangle dt - 2\mu\langle v, R_h v\rangle dt \\ &+ 2\mu\langle v, dW\rangle + \mu^2 \operatorname{Tr}[Q]\, dt. \end{aligned} \tag{45}$$

Estimate the first two terms of the right-hand side as follows: using inequality (16) and Young's inequality

$$-2\langle v, B(v,U)\rangle \leq \nu\|v\|_V^2 + \frac{C_{\mathrm{L}}^2}{\nu}|v|_H^2\|U\|_V^2. \tag{46}$$

Using Young's inequality, the interpolation inequality (6) and the assumption that $2c_1\mu h^2 \leq \nu$ we obtain

$$\begin{aligned} -2\mu\langle v, R_h v\rangle &= 2\mu\langle v, v - R_h(v)\rangle - 2\mu|v|_H^2 \\ &\leq 2\mu|v - R_h(v)|_{L^2}^2 - \frac{3\mu}{2}|v|_H^2 \leq \nu\|v\|_V^2 - \frac{3\mu}{2}|v|_H^2. \end{aligned} \tag{47}$$

Therefore,

$$d|v|_H^2 + \Big(\frac{3\mu}{2} - \frac{C_{\mathrm{L}}^2}{\nu}\|U\|_V^2\Big)|v|_H^2\,dt \le 2\mu\langle v, dW\rangle + \mu^2\,\mathrm{Tr}[Q]\,dt.$$

Since $t \ge t_0$ then inequality (34) and the hypothesis $2C_{\mathrm{L}}^2\nu\lambda_1 G^2 \le \mu/2$ yield

$$\frac{C_{\mathrm{L}}^2}{\nu}\|U(t)\|_V^2 \le 2C_{\mathrm{L}}^2\nu\lambda_1 G^2 \le \frac{\mu}{2},$$

which implies

$$d|v|_H^2 + \mu|v|_H^2\,dt \le 2\mu\langle v, dW\rangle + \mu^2\,\mathrm{Tr}[Q]\,dt, \tag{48}$$

for all $t \ge t_0$. Now integrating over (t_0, t), then taking the expected value and using Gronwall's lemma, we obtain

$$\mathbb{E}(|v(t)|_H^2) \le \mathbb{E}(|v(t_0)|_H^2)e^{-\mu(t-t_0)} + \mu\,\mathrm{Tr}[Q], \quad \text{for all} \quad t \ge t_0.$$

Thus, it follows that

$$\limsup_{t\to\infty}\mathbb{E}(|v(t)|_H^2) \le \mu\,\mathrm{Tr}[Q].$$

To obtain (44), we estimate the terms in (46) and (47) using Young's inequality in a slightly different way. In particular,

$$-2\langle v, B(u, U)\rangle \le \frac{\nu}{2}\|v\|_V^2 + \frac{2C_L^2}{\nu}|v|_H^2\|U\|_V^2$$

and

$$-2\mu\langle v, R_h v\rangle \le \frac{\nu}{2}\|v\|_V^2 - \mu|v|_H^2.$$

Therefore,

$$d|v|_H^2 + \nu\|v\|_V^2 dt \le 2\mu\langle v, dW\rangle + \mu^2\,\mathrm{Tr}[Q]dt.$$

Taking expected values and integrating from t to $t+T$ yields

$$\mathbb{E}(|v(t+T)|_H^2) + \nu\int_t^{t+T}\mathbb{E}(\|v(\tau)\|_V^2)d\tau \le \mathbb{E}(|v(t)|_H^2) + \mu^2 T\,\mathrm{Tr}[Q].$$

Therefore

$$\limsup_{t\to\infty}\nu\int_t^{t+T}\mathbb{E}(\|v(\tau)\|_V^2)d\tau \le \mu\,\mathrm{Tr}[Q] + \mu^2 T\,\mathrm{Tr}[Q],$$

from which (44) immediately follows. $\square$

Theorem 4.1 applies to observations of the volume elements given by (8). We now state and prove a corollary on finite volume elements that gives explicit estimates on the how well u approximates U over time.

Corollary 4.2. *Suppose that the observational measurements are given by finite volume elements* (8) *plus a noise term of the form* (22)*, where each* b_d *is an independent one-dimensional Brownian motion with variance* $\sigma^2/2$*. Interpolate these noisy observations using* (23) *where* ℓ_d *are given by* (26)*. Let* $\mu = 4C_{\mathrm{L}}^2\nu\lambda_1 G^2$ *and choose* $N = K^2$ *large enough such that*

$$h = L/K \le \sqrt{\nu/(2c_1\mu)}.$$

Then the solution u *to* (21) *satisfies*

$$\limsup_{t\to\infty} \mathbb{E}(|u(t) - U(t)|_H^2) \le \kappa_1\nu G^2\sigma^2$$

and

$$\limsup_{t\to\infty} \frac{\nu}{T}\int_t^{t+T} \mathbb{E}(\|u(\tau) - U(\tau)\|_V^2)d\tau \le \Big(\frac{1}{T} + \frac{\kappa_1\nu G^2}{L^2}\Big)\kappa_1\nu G^2\sigma^2$$

where $\kappa_1 = 16\pi^2 C_{\mathrm{L}}^2$ *is an absolute constant.*

Proof. By Proposition 2.3 we obtain

$$\mu\,\mathrm{Tr}[Q] \le \mu\sigma^2 L^2 \le 4C_{\mathrm{L}}^2\nu\lambda_1 G^2\sigma^2 L^2 = \kappa_1\nu G^2\sigma^2.$$

where $\kappa_1 = 16\pi^2 C_{\mathrm{L}}^2$. Similarly

$$\Big(\frac{\mu}{T} + \mu^2\Big)\,\mathrm{Tr}[Q] \le \Big(\frac{\kappa_1\nu G^2}{T} + \frac{\kappa_1^2\nu^2 G^4}{L^2}\Big)\sigma^2.$$

Since the choices of μ and h given in the corollary satisfy the hypothesis of Theorem 4.1, the corollary is thus proved. □

We remark that the upper bound on the error in the approximating solution given by Corollary 4.2 is independent of h. In particular, as we take the observation density finer and finer there is no improvement in the quality of our approximation. This is not surprising, since increasing the resolution of the observations did not lead to any decrease in the size of the measurement errors present in the interpolant observables $\tilde{R}_h$ given by (23). We remedy this defect with

Corollary 4.3. *Suppose that the observational measurements are given by finite volume elements* (8) *plus a noise term of the form* (22) *where each* b_d *is an independent one-dimensional Brownian motion with variance* $\sigma^2/2$*. Let* μ *be as in Corollary 4.2 and* $\epsilon \in (0,1)$*. Then, there exists an interpolant observable based on volume elements with observation density* h *such that*

$$\frac{\kappa_2 G^2}{L^2} \le \frac{\epsilon}{h^2} \le \frac{\max(\epsilon, 16\kappa_2 G^2)}{L^2} \tag{49}$$

where $\kappa_2 = 32\pi^2 c_1 C_{\mathrm{L}}^2$ *is an absolute constant and for which*

$$\limsup_{t\to\infty} \mathbb{E}(|u(t) - U(t)|_H^2) \le \mu\sigma^2 L^2 \epsilon$$

and

$$\limsup_{t\to\infty} \frac{\nu}{T}\int_t^{t+T} \mathbb{E}(\|u(\tau) - U(\tau)\|_V^2)d\tau \le \Big(\frac{1}{T} + \mu\Big)\mu\sigma^2 L^2 \epsilon. \tag{50}$$

Proof. If $\sqrt{\nu/(2c_1\mu)} \ge L$ then we may take $h = L$ in Theorem 4.1. In this case U is a steady state, and consequently no observational data is needed to accurately recover U. Otherwise, let $M = K_2^2$ where $K_2 \ge 2$ is the unique integer such that

$$h' = L/K_2 \le \sqrt{\nu/(2c_1\mu)} < L/(K_2 - 1).$$

Let Q'_m be squares with sides of length h' where $m = 1, 2, \ldots, M$ defined in a similar way as (27). Choose $h = h'/q$ where q is the unique integer satisfying

$$q^2 \ge \epsilon^{-1} > (q-1)^2. \tag{51}$$

With these choices of K_2 and q we have

$$\sqrt{2c_1\mu/\nu} \le 1/h' = K_2/L \le 2(K_2 - 1)/L < 2\sqrt{2c_1\mu/\nu}$$

and

$$\epsilon^{-1/2} \le q = (q-1) + 1 \le \epsilon^{-1/2} + 1 \le 2\epsilon^{-1/2}.$$

Therefore

$$\sqrt{\nu\epsilon/(32c_1\mu)} \le h = h'/q \le \sqrt{\nu\epsilon/(2c_1\mu)}$$

from which (49) follows.

Let Q_n be the squares with sides of length h where $n = 1, 2, \ldots, N$ and $N = L^2/h^2 = q^2 M$. The smaller squares Q_n fit inside the larger squares Q'_m and each larger square is the union over q^2 of the smaller squares.

Define the averaging operator $\mathcal{A}\colon \mathbb{R}^{2N} \to \mathbb{R}^{2M}$ by

$$\begin{bmatrix} \mathcal{A}(\varphi)_{2m-1} \\ \mathcal{A}(\varphi)_{2m} \end{bmatrix} = \frac{1}{q^2} \sum_{Q_n \subseteq Q'_m} \begin{bmatrix} \varphi_{2n-1} \\ \varphi_{2n} \end{bmatrix},$$

for $m = 1, 2, \ldots, M$. We note that $\mathcal{O}_{h'} = \mathcal{A} \circ \mathcal{O}_h$, where $\mathcal{O}_h$ are the noise-free observations of volume elements given in (8), for the Q_n and $\mathcal{O}_{h'}$ are the analogous observations for Q'_m.

Let $\tilde{\mathcal{O}}_h(U(t))$ be the noisy observations defined by (3), where $\mathcal{E}(t)$ is given by (22). It follows that $\mathcal{A} \circ \tilde{\mathcal{O}}_h(U(t)) = \mathcal{O}_{h'}(U(t)) + \mathcal{F}(t)$, where

$$\mathcal{F}(t)dt = (d\beta_1(t), d\beta_2(t), \ldots, d\beta_{2M}(t))$$

and β_k are one-dimensional independent Brownian motions such that

$$\mathbb{E}(\beta_k(t)) = 0 \quad \text{and} \quad \mathbb{E}(\beta_k(t)^2) = t\frac{\sigma^2}{2q^2},$$

for $k = 1, 2, \ldots, 2M$. Therefore, by taking averages of volume elements, also called spatial oversampling, we have reduced the variance in the noise term of the measurements. In particular, from (51) the noise term is now equivalent to a $[\dot{L}^2]^2$-valued Q'-Brownian motion with $\mathrm{Tr}[Q'] \le \sigma^2 L^2/q^2 \le \sigma^2 L^2 \epsilon$. We now define the interpolant observable

$$\mathcal{R}_{h'} = \mathcal{L}_{h'} \circ \mathcal{A} \circ \mathcal{O}_h.$$

Since $\mathcal{R}_{h'}$ satisfies (6) with the same constants as before, then applying Theorem 4.1 now completes the proof. □

4.2 Observations of Nodal Values

This section first proves a general theorem on interpolant observables which satisfy (7). This result is then applied to obtain explicit estimates when the observational measurements arise from nodal measurements.

Our proof follows the general strategy of the non-stochastic case treated in [35] with modifications as was done in the proof of Theorem 4.1 above to account for the stochastic terms which arise from the stochastic errors. In particular, we shall make use of the following inequality which can be found in [35].

Lemma 4.4. *Let $\varphi(r) = r - \eta(1 + \log r)$ where $\eta > 0$. Then*

$$\min\{\varphi(r) : r \ge 1\} \ge -\eta \log \eta.$$

We commence with the proof of

Theorem 4.5. *Assume that U is a strong solution of (20), that R_h satisfies assumption (7) and that W is a $[\dot{H}^1]^2$-valued Q-Brownian motion. Assume that μ is large enough and that h is small enough such that*

$$\mu \geq 2\nu\lambda_1 GJ, \quad \textit{and} \quad \nu \geq 2c_3 h^2 \mu$$

where $c_3 = \max(c_1, \sqrt{c_2})$ and $J = 2C_{\mathrm{B}}(2 + \log 2C_{\mathrm{B}} c^{1/2})(1 + \log(1 + G))$. Then

$$\limsup_{t\to\infty} \mathbb{E}(\|u - U\|_V^2) \leq 4\mu \exp(\nu\lambda_1 G^2 J^2/\mu) \operatorname{Tr}[A^{1/2} Q A^{1/2}]$$

and

$$\begin{aligned} &\limsup_{t\to\infty} \frac{\nu}{T} \int_t^{t+T} \mathbb{E}(|Au(\tau) - AU(\tau)|_H^2) d\tau \\ &\quad \leq \left\{ 8 \exp(\nu\lambda_1 G^2 J^2/\mu) \left\{ \frac{\mu}{T} + 4J^2 \Big(\frac{1}{T} + \nu\lambda_1 \Big) \nu\lambda_1 G^2 \right\} + \mu^2 \right\} \\ &\qquad \times 2 \operatorname{Tr}[A^{1/2} Q A^{1/2}]. \end{aligned}$$

Proof. We focus on the interval $[t_0, \infty)$, where t_0 is given in Theorem 2.8. Using the Itô formula on $\|v(t)\|_V$ we obtain

$$d\|v\|_V^2 = 2(\!(v, dv)\!) + \mu^2 \operatorname{Tr}[A^{1/2} Q A^{1/2}] dt.$$

For notational convenience we shall write $\Sigma = \operatorname{Tr}[A^{1/2} Q A^{1/2}]$ throughout the rest of this proof. Substituting for dv and applying (18) and (19) yields

$$\begin{aligned} d\|v\|_V^2 + 2\nu |Av|_H^2 dt = &2\langle AU, B(v, v)\rangle dt - 2\mu \langle Av, R_h(v)\rangle dt \\ &+ 2\mu \langle Av, dW\rangle + \mu^2 \Sigma dt. \end{aligned}$$

The Brézis–Gallouet inequality (17) implies

$$\begin{aligned} |\langle AU, B(v, v)\rangle| &\leq \|v\|_\infty \|v\|_V |AU|_H \\ &\leq C_{\mathrm{B}} \|v\|_V^2 \left\{ 1 + \log \frac{|Av|_H^2}{\lambda_1 \|v\|_V^2} \right\} |AU|_H \end{aligned}$$

and the assumption $2\mu \max(c_1, \sqrt{c_2}) h^2 \leq \nu$ along with (7) and Young's inequality implies

$$\begin{aligned} -2\mu \langle Av, R_h(v)\rangle &= 2\mu \langle Av, v - R_h(v)\rangle - 2\mu \|v\|_V^2 \\ &\leq \frac{2\mu^2}{\nu} |v - R_h(v)|_{L^2}^2 + \frac{\nu}{2} |Av|_H^2 - 2\mu \|v\|_V^2 \\ &\leq \nu |Av|_H^2 - \mu \|v\|_V^2. \end{aligned}$$

Therefore

$$d\|v\|_V^2 + \left(\nu\lambda_1 \frac{|Av|_H^2}{\lambda_1\|v\|_V^2} - 2C_{\mathrm{B}}|AU|_H\left\{1 + \log\frac{|Av|_H^2}{\lambda_1\|v\|_V^2}\right\} + \mu\right)\|v\|_V^2 dt$$
$$\leq 2\mu\langle Av, dW\rangle + \mu^2\Sigma dt.$$

Now setting

$$\eta = \frac{2C_{\mathrm{B}}|AU|_H}{\nu\lambda_1} \quad \text{and} \quad r = \frac{|Av|_H^2}{\lambda_1\|v\|_V^2}$$

in Lemma 4.4, and noting that $r \geq 1$, we obtain that

$$d\|v\|_V^2 + \left(\mu - 2C_{\mathrm{B}}|AU|_H \log\frac{2C_{\mathrm{B}}|AU|_H}{\nu\lambda_1}\right)\|v\|_V^2 dt$$
$$\leq 2\mu\langle Av, dW\rangle + \mu^2\Sigma dt.$$

Since $t \geq t_0$, then by (35) we estimate

$$2C_{\mathrm{B}}\log\frac{2C_{\mathrm{B}}|AU|_H}{\nu\lambda_1} \leq 2C_{\mathrm{B}}\log 2C_{\mathrm{B}}c^{1/2}(1+G)^2 \leq J.$$

Consequently,

$$d\|v\|_V^2 + \left\{\mu - J|AU|_H\right\}\|v\|_V^2 dt \leq 2\mu\langle Av, dW\rangle + \mu^2\Sigma dt.$$

Applying Young's inequality we get

$$d\|v\|_V^2 + \frac{1}{2}\left\{\mu - \frac{J^2}{\mu}|AU|_H^2\right\}\|v\|_V^2 dt \leq 2\mu\langle Av, dW\rangle + \mu^2\Sigma dt.$$

Take t_0 as in Theorem 2.8 and define

$$\alpha(t) = \frac{1}{2}\left\{\mu - \frac{J^2}{\mu}|AU(t)|_H^2\right\} \quad \text{and} \quad \Psi(t) = \int_{t_0}^t \alpha(s)ds.$$

Now, integrating and taking expected value yields

$$\mathbb{E}(\|v(t)\|_V^2) \leq \mathbb{E}(\|v(t_0)\|_V^2)e^{-\Psi(t)} + \mu^2\Sigma\int_{t_0}^t e^{-\Psi(t)+\Psi(\tau)}d\tau.$$

Since $\mu > 2\nu\lambda_1 GJ$ then by the estimate (34) we obtain

$$-\Psi(t) + \Psi(\tau) \leq -\frac{\mu}{2}(t-\tau) + \frac{J^2}{\mu}(1 + (t-\tau)\nu\lambda_1)\nu\lambda_1 G^2$$
$$\leq -\frac{\mu}{4}(t-\tau) + \frac{\nu\lambda_1}{\mu}G^2J^2.$$

Therefore $-\Psi(t) \to -\infty$ and

$$\int_{t_0}^{t} e^{-\Psi(t)+\Psi(\tau)} d\tau \le \frac{c_4}{\mu}\big(1 - e^{-\mu(t-t_0)/4}\big) \to \frac{c_4}{\mu}, \quad \text{as} \quad t \to \infty,$$

where $c_4 = 4\exp(\nu\lambda_1 G^2 J^2/\mu)$. It follows that

$$\limsup_{t\to\infty} \mathbb{E}(\|v(t)\|_V^2) \le c_4\mu\Sigma = 4\mu\exp(\nu\lambda_1 G^2 J^2/\mu)\operatorname{Tr}[A^{1/2}QA^{1/2}],$$

which is the first inequality.

To obtain the second inequality, we use

$$\eta = \frac{4C_{\mathrm{B}}|AU|_H}{\nu\lambda_1} \quad \text{and} \quad r = \frac{|Av|_H^2}{\lambda_1\|v\|_V^2}$$

in Lemma 4.4 to obtain

$$\begin{aligned} d\|v\|_V^2 + \frac{\nu}{2}|Av|_H^2 dt + \frac{1}{2}\Big\{\mu - \frac{\tilde{J}^2}{\mu}|AU|_H^2\Big\}\|v\|_V^2 dt& \\ \le 2\mu\langle Av, dW\rangle dt + \mu^2\Sigma dt& \end{aligned} \tag{52}$$

where

$$\tilde{J} = 4C_{\mathrm{B}}\log 4C_{\mathrm{B}}c^{1/2}(1+G)^2 \le 2J.$$

Let $t_1 \ge t_0$ be large enough such that

$$\mathbb{E}(\|v(t)\|_V^2) \le 2c_4\mu\Sigma \qquad \text{for} \qquad t \ge t_1.$$

After integrating inequality (52) and taking expected values, we obtain for times $t > t_1$ that

$$\begin{aligned} \frac{\nu}{2}\int_t^{t+T} \mathbb{E}\,|Av(\tau)|_H^2 d\tau &\le 2c_4\mu\Sigma + c_4\tilde{J}^2\Sigma\int_t^{t+T}|AU(\tau)|_H^2 d\tau + \mu^2 T\Sigma \\ &\le 2c_4\mu\Sigma + 8c_4J^2\Sigma(1+T\nu\lambda_1)\nu\lambda_1 G^2 + \mu^2 T\Sigma. \end{aligned}$$

This finishes the proof. □

The bounds on h^{-2} are proportional to $G(1+\log(1+G))$ which is similar to the deterministic case. However, the bounds on the expected value of $\|u-U\|_V^2$ depend exponentially on G. Therefore, unless the variance in the stochastic error represented by $\operatorname{Tr}[A^{1/2}QA^{1/2}]$ is very small, this bound will be very large. However, this exponential dependence on G may be removed by taking $\mu = \nu\lambda_1 G^2 J^2$ and h^{-2} correspondingly larger. This yields

Corollary 4.6. *Suppose* $\mu = c_5\nu\lambda_1 G^2(1+\log(1+G))^2$ *and* h *is small enough such that*

$$\nu \geq 2c_3\mu h^2,$$

where c_3 *is defined as in Theorem 4.5 and* $c_5 = 4C_{\mathrm{B}}^2(2+\log 2C_{\mathrm{B}}c^{1/2})^2$. *Then*

$$\limsup_{t\to\infty} \mathbb{E}(\|u-U\|_V^2) \leq 4e\mu \operatorname{Tr}[A^{1/2}QA^{1/2}]$$

and

$$\limsup_{t\to\infty} \frac{\nu}{T}\int_t^{t+T} \mathbb{E}(|Au(\tau)-AU(\tau)|_H^2)d\tau$$
$$\leq \Big\{\frac{20}{T}+16\nu\lambda_1+\frac{\mu}{2e}\Big\}4e\mu \operatorname{Tr}[A^{1/2}QA^{1/2}].$$

We apply Corollary 4.6 to nodal observations described by (10) to obtain

Corollary 4.7. *Suppose that the observational measurements are given by nodal observations* (10) *plus a noise term of the form* (22), *where each* b_d *is an independent one-dimensional Brownian motion with variance* $\sigma^2/2$. *Interpolate the noisy observations using* (23) *where* ℓ_d *are given by* (29). *Suppose that* $\sqrt{\nu/(2c_3\mu)} < L$ *where* $\mu = c_5\nu\lambda_1 G^2(1+\log(1+G))^2$ *and choose* $N = K^2$ *such that*

$$h = L/K \leq \sqrt{\nu/(2c_3\mu)} < L/(K-1).$$

Then the solution U *to* (21) *satisfies*

$$\limsup_{t\to\infty} \mathbb{E}(\|u(t)-U(t)\|_V^2) \leq \kappa_3\nu\lambda_1 G^4(1+\log(1+G))^4\sigma^2$$

and

$$\limsup_{t\to\infty} \frac{\nu}{T}\int_t^{t+T} \mathbb{E}(|Au(\tau)-AU(\tau)|_H^2)d\tau$$
$$\leq \Big\{\frac{20}{T}+16\nu\lambda_1+\frac{c_5\nu\lambda_1}{2e}G^2(1+\log(1+G))^2\Big\}$$
$$\times\, \kappa_3\nu\lambda_1 G^4(1+\log(1+G))^4\sigma^2$$

where $\kappa_3 = 128\pi^2 ecc_3c_5^2$ *is an absolute constant.*

Proof. By Proposition 2.6 equation (31), we obtain

$$\operatorname{Tr}[A^{1/2}QA^{1/2}] \leq c\sigma^2\frac{L^2}{h^2} \leq c\sigma^2K^2 \leq 4c\sigma^2(K-1)^2$$
$$< 32\pi^2 cc_3c_5 G^2(1+\log(1+G))^2\sigma^2.$$

Since μ and h satisfy the hypothesis of Corollary 4.6, then

$$\limsup_{t\to\infty} \mathbb{E}(\|u-U\|_V^2) \leq 128\pi^2 ecc_3c_5^2\nu\lambda_1 G^4(1+\log(1+G))^4\sigma^2$$

and the second inequality follows similarly. □

We end by noting that the oversampling argument used to reduce the error in Corollary 4.3 can also be used with nodal measurements. Along these lines we obtain

Corollary 4.8. *Suppose that the observational measurements are given by nodal observations* (10) *plus a noise term of the form* (22) *where each* b_d *is an independent one-dimensional Brownian motion with variance* $\sigma^2/2$*. Let* μ *be as in Corollary 4.6 and* $\epsilon \in (0,1)$*. Then, there exists an interpolant observable based on nodal measurements with observation density* h *such that*

$$\frac{\kappa_4 G^2(1+\log(1+G))^2}{L^2} \le \frac{\epsilon}{h^2} \le \frac{\max(\epsilon, 16\kappa_4 G^2(1+\log(1+G))^2)}{L^2}$$

where $\kappa_4 = 32\pi^2 c_3 C_{\mathrm{B}}^2(2+\log 2C_{\mathrm{B}}c^{1/2})^2$ *is an absolute constant and for which*

$$\limsup_{t\to\infty} \mathbb{E}(\|u(t)-U(t)\|_V^2) \le 32ecc_3\mu^2\nu^{-1}\sigma^2L^2\epsilon$$

and

$$\begin{aligned}\limsup_{t\to\infty} \frac{\nu}{T}\int_t^{t+T} & \mathbb{E}\,|Au(\tau)-AU(\tau)|_H^2 d\tau \\ & \le \Big\{\frac{20}{T}+16\nu\lambda_1+\frac{\mu}{2e}\Big\} 32ecc_3\mu^2\nu^{-1}\sigma^2L^2\epsilon.\end{aligned}$$

Proof. Define h', K_2, M, N, q, Q_n and Q'_m as in the proof of Corollary 4.3 where we have taken c_3 in place of c_1. Let $x_n \in Q_n$ for $n = 1, 2, \ldots, N$. Since the Q_n are disjoint then the x_n are distinct. Inside each large square Q'_m fit q^2 smaller squares Q_n and therefore q^2 points x_n. Denote

$$\{\, x_n : x_n \in Q_m \,\} = \{\, x'_{m,j} : j = 1, 2, \ldots, q^2 \,\}.$$

Since $x'_{m,j} \in Q_m$ for each $j = 1, \ldots, q^2$, we may view $\mathcal{O}_h$ as a family of q^2 observations of nodes $\mathcal{O}^j_{h'} \colon [\dot H^2]^2 \to \mathbb{R}^{2M}$ given by

$$\mathcal{O}^j_{h'}(\Phi) = (\varphi_{1,j}, \ldots, \varphi_{2M,j}) \quad \text{where} \quad \begin{bmatrix} \varphi_{2m-1,j} \\ \varphi_{2m,j} \end{bmatrix} = \Phi(x'_{m,j})$$

and $m = 1, 2, \ldots, M$. This leads to a family of q^2 independent noisy observations $\tilde{\mathcal{O}}^j_{h'}(U(t))$. It follows that the average of the noisy observations

$$\tilde{\mathcal{O}}_{h'}(U(t)) = \frac{1}{q^2}\sum_{j=1}^{q^2} \tilde{\mathcal{O}}^j_{h'}(U(t)) = \frac{1}{q^2}\sum_{j=1}^{q^2} \mathcal{O}^j_{h'}(U(t)) + \mathcal{F}(t)$$

where

$$\mathcal{F}(t)dt = (d\beta_1(t), d\beta_2(t), \ldots, d\beta_{2M})$$

and the β_k are one-dimensional independent Brownian motions such that $\mathbb{E}(\beta_k(t)) = 0$ and $\mathbb{E}(\beta_k(t)^2) = t\sigma^2/(2q^2)$ for $k = 1, 2, \ldots, 2M$. Therefore, just as in the case with finite volume elements, we have reduced the variance in the noise term by averaging. In particular, the noise term is now equivalent to an $[\dot{H}^1]^2$-valued Q'-Brownian motion with

$$\mathrm{Tr}[A^{1/2}Q'A^{1/2}] \leq c\sigma^2\Big(\frac{L}{h'}\Big)^2\frac{1}{q^2} \leq 8cc_3\mu\nu^{-1}\sigma^2L^2\epsilon.$$

We now define the interpolant observable

$$\mathcal{R}_{h'} = \frac{1}{q^2}\sum_{j=1}^{q^2}\mathcal{L}_{h'} \circ \mathcal{O}^j_{h'}(U(t)).$$

Since $\mathcal{R}_{h'}$ satisfies (7) with the same constants as before, then applying Corollary 4.6 now completes the proof. $\square$

5 Conclusions

We have shown the continuous data assimilation algorithm proposed in [3] continues to be well posed when the observational measurements contain errors represented by stochastic noise. Provided the resolution of the observational data is fine enough, we have shown that the expected value of the difference between the approximate solution, recovered by this data assimilation algorithm, and the exact solution is bounded by a factor depending on the Grashof number times the variance of the noise, asymptotically in time. This occurs for general interpolant operator observables satisfying either one of the approximate identity properties (6) or (7), and, in particular, for interpolant observables based on volume elements and nodal measurements.

In the case of Theorem 4.5 the resolution of the observational data needed for the algorithm to work for noisy measurements is roughly the same as without noise; however, to remove the exponential dependency on the Grashof number in the error bounds, Corollary 4.6, requires increasing the resolution by its square. Once the resolution needed to remove the exponential term is achieved, no further benefits are obtained by increasing the resolution. To benefit from additional resolution in the observational measurements, we note that oversampling an already very high resolution observation, and then by locally averaging the oversampled observation, can produce a observation that still has sufficient resolution but with reduced variance in the noise. In our case, we assumed the random errors were independent; however, this may not be the case in practical problems. For example, Budd, Freitag and Nichols [4] obtain great benefits by using adaptive filters based on assumptions about the independence of the measurements errors in real-world weather forecasting applications. The effect oversampling has on reducing the errors in our theoretical bounds is consistent with the observed effects of filtering in applications.

Computer simulations done by Gesho [26] have shown that in the absence of measurements errors the algorithm studied in this paper performs much better than analytical estimates would suggest. In the case of nodal measurements, the actual resolution requirements for the observation density is orders of magnitude less than the upper bounds given by the analysis. This phenomenon, that the numerics perform much better than the analysis, was also noted for a different data assimilation algorithm in [35] and [36]. It is plausible that in the presence of stochastic noise the data algorithm studied here will also perform numerically much better than our analytic bounds. Work is underway to study the numerical performance of this data assimilation algorithm when the observation density is much less than our analytic bounds and to understand how the variance in the stochastic noise numerically affects the convergence of the approximating solution to the reference solution over time.

Acknowledgements

The authors would like to dedicate this work to Professor Ciprian Foias on the occasion of his 80th birthday as a token of appreciation for his continuous support, friendship and inspiration. H.B. and E.S.T. are thankful to the IMA for its kind hospitality where this project has started during their joint visit to the IMA. E.S.T. is also thankful to the warm hospitality of the Instituto Nacional de Matemática Pura e Aplicada (IMPA), where part of this work was completed. H.B. was supported in part by the Simons Foundation grant #283308. The work of E.O. was supported in part by EPSRC grant EP/G007470/1 and by sabbatical leave from the University of Nevada Reno. The work of E.S.T. was supported in part by the NSF grants DMS-1009950, DMS-1109640 and DMS-1109645. Also by the CNPq-CsF grant #401615/2012-0, through the program Ciência sem Fronteiras.

References

[1] J.P. Aubin, Un théorème de compacité, *C.R. Acad. Sci. Paris Sér. I Math.*, Vol. 256, 1963, pp. 5042–5044.

[2] A. Azouani, E.S. Titi, Feedback control of nonlinear dissipative systems by finite determining parameters—a reaction-diffusion paradigm. arXiv:1301.6992

[3] A. Azouani, E. Olson, E.S. Titi, Continuous data assimilation using general interpolant observables, *Journal of Nonlinear Science*, Vol. 24, No. 2, 2013, pp. 1–27.

[4] C.J. Budd, M.A. Freitag, N.K. Nichols, Regularization techniques for ill-posed inverse problems in data assimilation, *Comput. & Fluids*, Vol. 46, 2011, pp. 168-173.

[5] H. Bessaih, B. Ferrario, Inviscid limit of stochastic damped 2D Navier-Stokes equations, *Nonlinearity*, Vol. 27, 2013, pp. 1–15.

[6] D. Blömker, K.J.H. Law, A.M. Stuart, K. C. Zygalakis, Accuracy and stability of the continuous-times 3DVAR filter for the Navier–Stokes equations, *Nonlinearity*, Vol. 26, 2013, pp. 2193–2219.

[7] S.C. Brenner, R. Scott, *The Mathematical Theorey of Finite Element Methods*, Springer, 2007.

[8] H. Brézis, T. Gallouet, Nonlinear Schrödinger evolution equations, *Nonlinear Anal.*, Vol. 4, No. 4, 1980, pp. 677–681.

[9] G.L. Browning, W.D. Henshaw, H.O. Kreiss, A numerical investigation of the interaction between the large and small scales of the two-dimensional incompressible Navier–Stokes equations, Research report LA-UR-98-1712, Los Alamos National Laboratory, 1998.

[10] I. Chueshov, A. Millet, Stochastic 2D hydrodynamical type system: Well posedness and large deviation, *Appl. Math. Optim.*, Vol. 61, 2010, pp. 379–420.

[11] P. Ciarlet, *The Finite Element Method for Elliptic Problems,* North-Holland, 1978.

[12] B. Cockburn, D.A. Jones, E.S. Titi, Estimating the number of asymptotic degrees of freedom for nonlinear dissipative systems, *Mathematics of Computation*, Vol. 66, No. 219, 1997, pp. 1073–1087.

[13] P. Constantin, C. Foias, *Navier–Stokes Equations*, University of Chicago Press, 1988.

[14] R. Daley, *Atmospheric Data Analysis*, Cambridge Atmospheric and Space Science Series, Cambridge University Press, 1991.

[15] G. Da Prato, J. Zabczyk: *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, Cambridge, 1992.

[16] R. Dascaliuc, C. Foias, M.S. Jolly, Estimates on enstrophy, palinstrophy, and invariant measures for 2-D turbulence, *Journal of Differential Equations*, Vol. 248, 2010, pp. 792–819.

[17] F. Flandoli, Dissipativity and invariant measures for stochastic Navier-Stokes equations, *Nonlinear Differential Equations Appl.*, Vol. 1, No. 4, 1994, pp. 403–423.

[18] C. Foias, M. Jolly, R. Kravchenko and E.S. Titi, A determining form for the 2D Navier-Stokes equations - the Fourier modes case, *Journal of Mathematical Physics*, Vol. 53, 115623, 2012.

[19] C. Foias, M. Jolly, R. Kravchenko and E.S. Titi, A unified approach to determining forms for the 2D Navier-Stokes equations – the general interpolants case, *Russian Mathematical Surveys*, (2013), (in press). arXiv:1309.0247

[20] C. Foias, M. Jolly, R. Lan, R. Rupam, Y. Yang, Analyticity with higher norm estimates for the 2D Navier Stokes equations, arXiv:1312.0929.

[21] C. Foias, O. Manley, R. Rosa, R. Temam, *Navier–Stokes Equations and Turbulence*, Encyclopedia of Mathematics and Its Applications 83, Cambridge University Press, 2001.

[22] C. Foias, G. Prodi, Sur le comportement global des solutions non-stationnaires des équations de Navier-Stokes en dimension 2, *Rend. Sem. Mat. Univ. Padova*, Vol. 39, 1967, pp. 1–34.

[23] C. Foias; R. Temam, Sur la détermination d'un écoulement fluide par des observations discrètes [On the determination of fluid flows by discrete observations, *C. R. Acad. Sci. Paris Sér. I Math.* Vol. 295, No. 3, 1982, pp. 239–241 (and the continuation in Vol. 295, No 9, 1982, pp. 523–525).

[24] C. Foias, R. Temam, Determination of the solutions of the Navier-Stokes equations by a set of nodal values, *Math. Comp.* Vol. 43, No. 167, 1984, pp. 117–133.

[25] C. Foias, E.S. Titi, Determining nodes, finite difference schemes and inertial manifolds, *Nonlinearity*, Vol. 4, No. 1, 1991, pp. 135–153.

[26] M. Gesho, *A Numerical Study of Continuous Data Assimilation Using Nodal Points in Space for the Two-dimensional NavierStokes Equations*, Masters Thesis, University of Nevada, Department of Mathematics and Statistics, 2013.

[27] K. Hayden, E. Olson, E.S. Titi, Discrete data assimilation in the Lorenz and 2D Navier-Stokes equations, *Physica D: Nonlinear Phenomena*, Vol. 240, No. 18, 2011, pp. 1416–1425.

[28] W.D. Henshaw, H.O. Kreiss, J. Yström, Numerical experiments on the interaction between the large and small-scale motions of the Navier–Stokes equations, *Multiscale Modeling and Simulation*, Vol. 1, No. 1, 2003, pp. 119–149.

[29] M.J. Holst and E.S. Titi, Determining projections and functionals for weak solutions of the Navier–Stokes equations, *Contemporary Mathematics* Vol. 204, 1997, pp. 125–138.

[30] D.A. Jones, E.S. Titi, Determining finite volume elements for the 2D Navier–Stokes equations, *Physica D*, Vol. 60, 1992, pp. 165–174.

[31] D.A. Jones, E.S. Titi, Upper bounds on the number of determining modes, nodes and volume elements for the Navier–Stokes equations, *Indiana Univ. Math. J.*, Vol. 42, No. 3, 1993, pp. 875–887.

[32] P. Korn, Data assimilation and determining observations for the Lagrangian-averaged Navier-Stokes$-\alpha$ equations, *Physica D*, Vol. 238, 2009, pp. 1957–1974.

[33] J.L. Lions, E. Magenes, *Nonhomogeneous Boundary Value Problems*, Springer, Berlin, 1972.

[34] J.C. Mattingly, Ergodicity of the 2D Navier–Stokes Equations with Random Forcing and Large Viscosity, *Comm. Math. Phys.*, Vol. 206, No. 2, 1999, pp. 273–288.

[35] E. Olson, E.S. Titi, Determining modes for continuous data assimilation in 2D turbulence, *Journal of Statistical Physics*, Vol. 113, No. 5–6, 2003, pp. 799–840.

[36] E. Olson, E.S. Titi, Determining modes and Grashoff number in 2D turbulence, *Theoretical and Computational Fluid Dynamics*, Vol. 22, No. 5, 2008, pp. 327–339.

[37] Pazy, A.: *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Applied Mathematical Sciences, **44**, Springer-Verlag, New York, 1983.

[38] J. Robinson, *Infinite-Dimensional Dynamical Systems*, Cambridge Texts in Applied Mathematics, 2001.

[39] R. Temam, *Navier–Stokes Equations and Nonlinear Functional Analysis*, CBMS Regional Conference Series, No. 41, SIAM, Philadelphia, 1983.

[40] R. Temam, *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, second edition, Applied Mathematics Sciences, Vol. 68, Springer Verlag, 1997.

[41] R. Temam, *Navier–Stokes Equations: Theory and Numerical Analysis*, revised edition, AMS Chelsea Publising, 2001.

[42] R. Wait, A.R. Mitchell, *Finite Element Analysis and Applications*, John Wiley & Sons, 1985.