Determining white noise forcing from Eulerian observations in the Navier-Stokes equation

**Viet Ha Hoang · Kody J. H. Law ·
Andrew M. Stuart**

Abstract The Bayesian approach to inverse problems is of paramount importance in quantifying uncertainty about the input to, and the state of, a system of interest given noisy observations. Herein we consider the forward problem of the forced 2D Navier-Stokes equation. The inverse problem is to make inference concerning the forcing, and possibly the initial condition, given noisy observations of the velocity field. We place a prior on the forcing which is in the form of a spatially-correlated and temporally-white Gaussian process, and formulate the inverse problem for the posterior distribution. Given appropriate spatial regularity conditions, we show that the solution is a continuous function of the forcing. Hence, for appropriately chosen spatial regularity in the prior, the posterior distribution on the forcing is absolutely continuous with respect to the prior and is hence well-defined. Furthermore, it may then be shown that the posterior distribution is a continuous function of the data. We complement these theoretical results with numerical simulations showing the feasibility of computing the posterior distribution, and illustrating its properties.

Keywords Bayesian inversion · Navier-Stokes equation · White noise forcing

V. H. Hoang (✉)
Division of Mathematical Sciences, School of Physical and Mathematical Sciences,
Nanyang Technological University, Singapore 637371, Singapore
e-mail: vhhoang@ntu.edu.sg

K. J. H. Law · A. M. Stuart
Mathematics Institute, Warwick University, Coventry CV4 7AL, UK
e-mail: a.m.stuart@warwick.ac.uk

Present address:
K. J. H. Law
Computer, Electrical and Mathematical Sciences & Engineering Division, King Abdullah University
of Science and Technology, Thuwal 23955-6900, Saudi Arabia
e-mail: kody.law@kaust.edu.sa

1 Introduction

The Bayesian approach to inverse problems has grown in popularity significantly over the last decade, driven by algorithmic innovation and steadily increasing computer power [10]. Recently there have been systematic developments of the theory of Bayesian inversion on function spaces [3,11–14,18] and this has led to new sampling algorithms which perform well under mesh-refinement [2,15,21]. In this paper we add to this growing interest in the Bayesian formulation of inversion, in the context of a specific PDE inverse problem, motivated by geophysical applications such as data assimilation in the atmosphere and ocean sciences, and demonstrate that fully Bayesian probing of the posterior distribution is feasible.

The primary goal of this paper is to demonstrate that the Bayesian formulation of inversion for the forced Navier-Stokes equation, introduced in [3], can be extended to the case of white noise forcing. The paper [3] assumed an Ornstein-Uhlenbeck structure in time for the forcing, and hence did not include the white noise case. It is technically demanding to extend to the case of white noise forcing, but it is also of practical interest. This practical importance stems from the fact that the Bayesian formulation of problems with white noise forcing corresponds to a statistical version of the continuous time weak constraint 4DVAR methodology [22]. The 4DVAR approach to data assimilation currently gives the most accurate global short term weather forecasts available [16] and this is arguably the case because, unlike ensemble filters which form the major competitor, 4DVAR has a rigorous statistical interpretation as a maximum a posteriori (or MAP) estimator—the point which maximizes the posterior probability. It is therefore of interest to seek to embed our understanding of such methods in a broader Bayesian context.

To explain the connection between our work and the 4DVAR methodology and, in particular, to explain the terminology used in the data assimilation community, it is instructive to consider the finite dimensional differential equation

$$\frac{du}{dt} = f(u) + \xi, \quad u(0) = u_0$$

on $\mathbb{R}^n$. Assume that we are given noisy observations $\{y_j\}$ of the solution $u_j = u(t_j)$ at times $t_j = jh$ so that

$$y_j = u_j + \eta_j, \quad j = 1, \ldots, J,$$

for some noises η_j. An important inverse problem is to determine the initial condition u_0 and forcing ξ which best fit the data. If we view the solution u_j as a function of the initial condition and forcing, then a natural regularized least squares problem is to determine u_0 and ξ to minimize

$$I(u_0, \xi) = \sum_{j=1}^{J} \left| \Gamma^{-\frac{1}{2}} \left(y_j - u_j(u_0, \xi) \right) \right|^2 + |\Sigma^{-\frac{1}{2}} u_0|^2 + \|Q^{-\frac{1}{2}} \xi\|^2$$

where $|\cdot|$, $\|\cdot\|$ denote the $\mathbb{R}^n-$Euclidean and $L^2\big([0, T]; \mathbb{R}^n\big)$ norms respectively, Γ, Σ denote covariance matrices and Q a covariance operator. This is a continuous time analogue of the 4DVAR or variational methodology, as described in the book of Bennett [1]. In numerical weather prediction the method is known as weak constraint 4DVAR, and as 4DVAR if ξ is set to zero (so that the ODE

$$\frac{du}{dt} = f(u), \quad u(0) = u_0$$

is satisfied as a hard constraint), the term $\|\mathsf{Q}^{-\frac{1}{2}}\xi\|^2$ dropped, and the minimization is over u_0 only. As explained in [10], these minimization problems can be viewed as probability maximizers for the posterior distribution of a Bayesian formulation of the inverse problem—the so-called MAP estimators. In this interpretation the prior measures on u_0 and ξ are centred Gaussians with covariances Σ and Q respectively. Making this connection opens up the possibility of performing rigorous statistical inversion, and thereby estimating uncertainty in predictions made.

The ODEs arising in atmosphere and ocean science applications are of very high dimension due to discretizations of PDEs. It is therefore conceptually important to carry through the program in the previous paragraph, and in particular Bayesian formulation of the inversion, for PDEs; the paper [5] explains how to define MAP estimators for measures on Hilbert spaces and the connection to variational problems. The Navier-Stokes equation in 2D provides a useful canonical example of a PDE of direct relevance to the atmosphere and ocean sciences. When the prior covariance operator Q is chosen to be that associated to an Ornstein-Uhleneck operator in time, the Bayesian formulation for the 2D Navier-Stokes equation has been carried out in [3]. Our goal in this paper is to extend to the more technically demanding case where Q is the covariance operator associated with a white noise in time, with spatial correlation Q. We will thus use the prior model $\xi dt = dW$ where W is a $Q-$Wiener process in an appropriate Hilbert space, and consider inference with respect to W and u_0. In the finite dimensional setting the differences between the case of coloured and white noise forcing, with respect to the inverse problem, are much less substantial and the interested reader may consult [7] for details.

The key tools required in applying the function space Bayesian approach in [18] are the proof of continuity of the forward map from the function space of the unknowns to the data space, together with estimates of the dependence of the forward map upon its point of application, sufficient to show certain integrability properties with respect to the prior. This program is carried out for the 2D Navier-Stokes equation with Ornstein-Uhlenbeck priors on the forcing in the paper [3]. However to use priors which are white in time adds further complications since it is necessary to study the stochastically forced 2D Navier-Stokes equation and to establish continuity of the solution with respect to small changes in the Brownian motion W which defines the stochastic forcing. We do this by employing the solution concept introduced by Flandoli in [6], and using probabilistic estimates on the solution derived by Mattingly in [17]. In Sect. 2 we describe the relevant theory of the forward problem, employing the setting of Flandoli. In Sect. 3 we build on this theory, using the estimates of Mattingly to verify the conditions in [18], resulting in a well-posed Bayesian inverse problem for

which the posterior is Lipschitz in the data with respect to Hellinger metric. Section 4 extends this to include making inference about the initial condition as well as the forcing. Finally, in Sect. 5, we present numerical results which demonstrate feasibility of sampling from the posterior on white noise forces, and demonstrate the properties of the posterior distribution.

2 Forward problem

In this section we study the forward problem of the Navier-Stokes equation driven by white noise. Section 2.1 describes the forward problem, the Navier-Stokes equation, and rewrites it as an ordinary differential equation in a Hilbert space. In Sect. 2.2 we define the functional setting used throughout the paper. Section 2.3 highlights the solution concept that we use, leading in Sect. 2.4 to proof of the key fact that the solution of the Navier-Stokes equation is continuous as a function of the rough driving of interest and the initial condition. All our theoretical results in this paper are derived in the case of Dirichlet (no flow) boundary conditions. They may be extended to the problem on the periodic torus $\mathbb{T}^d$, but we present the more complex Dirichlet case only for brevity.

2.1 Overview

Let $D \in \mathbb{R}^2$ be a bounded domain with smooth boundary. We consider in D the Navier-Stokes equation

$$\partial_t u - \nu \Delta u + u \cdot \nabla u = f - \nabla p, \quad (x, t) \in D \times (0, \infty)$$
$$\nabla \cdot u = 0, \quad (x, t) \in D \times (0, \infty)$$
$$u = 0, \quad (x, t) \in \partial D \times (0, \infty),$$
$$u = u_0, \quad (x, t) \in D \times \{0\}. \tag{1}$$

We assume that the initial condition u_0 and the forcing $f(\cdot, t)$ are divergence-free. We will in particular work with Eq. (3) below, obtained by projecting (1) into the space of divergence free functions—the Leray projector [19]. We denote by V the space of all divergence-free smooth functions from D to $\mathbb{R}^2$ with compact support, by $\mathbb{H}$ the closure of V in $(L^2(D))^2$, and by $\mathbb{H}^1$ the closure of V in $(H^1(D))^2$. Let $\mathbb{H}^2 = (H^2(D))^2 \bigcap \mathbb{H}^1$. The initial condition u_0 is assumed to be in $\mathbb{H}$. We define the linear Stokes' operator $A : \mathbb{H}^2 \to \mathbb{H}$ by $Au = -\Delta u$ noting that the assumption of compact support means that Dirichlet boundary condition is imposed on the Stokes' operator A. Since A is selfadjoint, A possesses eigenvalues $0 < \lambda_1 \leq \lambda_2 \leq \cdots$ with the corresponding eigenvectors $e_1, e_2, \ldots \in \mathbb{H}^2$.

We denote by $\langle \cdot, \cdot \rangle$ the inner product in $\mathbb{H}$, extended to the dual pairing on $\mathbb{H}^{-1} \times \mathbb{H}^1$. We then define the bilinear form $B : \mathbb{H}^1 \times \mathbb{H}^1 \to \mathbb{H}^{-1}$

$$\langle B(u, v), z \rangle = \int_D z(x) \cdot (u(x) \cdot \nabla) v(x) dx$$

which must hold for all $z \in \mathbb{H}^1$. From the incompressibility condition we have, for all $z \in \mathbb{H}^1$,

$$\langle B(u, v), z \rangle = -\langle B(u, z), v \rangle. \tag{2}$$

By projecting problem (1) into $\mathbb{H}$ we may write it as an ordinary differential equation in the form

$$du(t) = -\nu A u \, dt - B(u, u) dt + dW(t), \quad u(0) = u_0 \in \mathbb{H}, \tag{3}$$

where $dW(t)$ is the projection of the forcing $f(x, t) dt$ into $\mathbb{H}$. We will define the solution of this equation pathwise, for suitable W, not necessarily differentiable in time.

2.2 Function spaces

For any $s \geq 0$ we define $\mathbb{H}^s \subset \mathbb{H}$ to be the Hilbert space of functions $u = \sum_{k=1}^{\infty} u_k e_k \in \mathbb{H}$ such that

$$\sum_{k=1}^{\infty} \lambda_k^s u_k^2 < \infty;$$

we note that the $\mathbb{H}^j$ for $j \in \{0, 1, 2\}$ coincide with the preceding definitions of these spaces. The space $\mathbb{H}^s$ is endowed with the inner product

$$\langle u, v \rangle_{\mathbb{H}^s} = \sum_{k=1}^{\infty} \lambda_k^s u_k v_k,$$

for $u = \sum_{k=1}^{\infty} u_k e_k$, $v = \sum_{k=1}^{\infty} v_k e_k$ in $\mathbb{H}$. We denote by $\mathbb{V}$ the particular choice $s = \frac{1}{2} + \epsilon$, namely $\mathbb{H}^{\frac{1}{2} + \epsilon}$, for given $\epsilon > 0$. In what follows we will be particularly interested in continuity of the mapping from the forcing W into linear functionals of the solution of (3). To this end it is helpful to define the Banach space $\mathbb{X} := C([0, T]; \mathbb{V})$ with the norm

$$\|W\|_{\mathbb{X}} = \sup_{t \in (0, T)} \|W(t)\|_{\mathbb{V}}.$$

2.3 Solution concept

In what follows we define a solution concept for Eq. (3) for each forcing function W which is continuous, but not necessarily differentiable, in time. We always assume that $W(0) = 0$. Following Flandoli [6], for each $W \in \mathbb{X}$ we define the weak solution $u(\cdot; W) \in C([0, T]; \mathbb{H}) \bigcap L^2([0, T]; \mathbb{H}^{1/2})$ of (3) as a function that satisfies

$$\langle u(t), \phi \rangle + \nu \int_0^t \langle u(s), A\phi \rangle ds - \int_0^t \langle B\big(u(s), \phi\big), u(s) \rangle dx = \langle u_0, \phi \rangle + \langle W(t), \phi \rangle,$$

$$(4)$$

for all $\phi \in \mathbb{H}^2$ and all $t \in (0, T)$; note the integration by parts on the Stokes' operator and the use of (2) to derive this identity from (3). Note further that if u and W are sufficiently smooth, (4) is equivalent to (3).

To employ this solution concept we first introduce the concept of a solution of the linear equation

$$dz(t) = -\nu A z dt + dW(t), \quad z(0) = 0 \in \mathbb{H} \tag{5}$$

where W is a deterministic continuous function obtaining values in $\mathbb{X}$ but not necessarily differentiable. We define a weak solution of this equation as a function $z \in C([0, T]; \mathbb{H})$ such that

$$\langle z(t), \phi \rangle + \nu \int_0^t \langle z(s), A\phi \rangle ds = \langle W(t), \phi \rangle \tag{6}$$

for all $\phi \in \mathbb{H}^2$.

Then for this function $z(t)$ we consider the solution v of the equation

$$dv(t) = -\nu A v dt - B(z + v, z + v) dt, \quad v(0) = u_0 \in \mathbb{H}. \tag{7}$$

As we will show below, $z(t)$ possesses sufficiently regularity so (7) possesses a weak solution v. We then deduce that $u = z + v$ is a weak solution of (3) in the sense of (4). When we wish to emphasize the dependence of u on W (and similarly for z and v) we write $u(t; W)$.

We will now show that the function z defined by

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

$$= W(t) - \int_0^t \nu A e^{-\nu A(t-s)} W(s) ds \tag{8}$$

satisfies the weak formula (6). Let $w_k = \langle W, e_k \rangle$, that is

$$W(t) := \sum_{k=1}^{\infty} w_k(t) e_k \in \mathbb{X}. \tag{9}$$

We then deduce from (8) that

$$z(t; W) = W(t) - \sum_{k=1}^{\infty} \left(\int_0^t w_k(s) v \lambda_k e^{(t-s)(-v\lambda_k)} ds \right) e_k. \qquad (10)$$

We have the following regularity property for z:

Lemma 1 *For each $W \in \mathbb{X}$, the function $z = z(\cdot; W) \in C([0, T]; \mathbb{H}^{1/2})$.*

Proof We first show that for each t, $z(t; W)$ as defined in (10) belongs to $\mathbb{H}^{1/2}$. Fixing an integer $M > 0$, using inequality $a^{1-\epsilon/2} e^{-a} < c$ for all $a > 0$ for an appropriate constant c, we have

$$\sum_{k=1}^{M} \lambda_k^{1/2} \left(\int_0^t v \lambda_k e^{(t-s)(-v\lambda_k)} w_k(s) ds \right)^2$$

$$\leq \sum_{k=1}^{M} \lambda_k^{1/2} \left(\int_0^t \frac{c}{(t-s)^{1-\epsilon/2}} \lambda_k^{\epsilon/2} |w_k(s)| dx \right)^2.$$

Therefore,

$$\left\| \sum_{k=1}^{M} \int_0^t v \lambda_k e^{(t-s)(-v\lambda_k)} w_k(s) e_k ds \right\|_{\mathbb{H}^{1/2}} \leq \left\| \sum_{k=1}^{M} \int_0^t \frac{c}{(t-s)^{1-\epsilon/2}} \lambda_k^{\epsilon/2} |w_k(s)| e_k ds \right\|_{\mathbb{H}^{1/2}}$$

$$\leq \int_0^t \frac{c}{(t-s)^{1-\epsilon/2}} \left\| \sum_{k=1}^{M} \lambda_k^{\epsilon/2} |w_k(s)| e_k \right\|_{\mathbb{H}^{1/2}} ds$$

$$\leq \max_{s \in (0,T)} \| W(s) \|_{\mathbb{H}^{1/2+\epsilon}} \int_0^t \frac{c}{(t-s)^{1-\epsilon/2}} ds,$$

which is uniformly bounded for all M. Therefore,

$$\sum_{k=1}^{\infty} \left(\int_0^t w_k(s) v \lambda_k e^{(t-s)(-v\lambda_k)} ds \right) e_k \in \mathbb{H}^{1/2}.$$

It follows from (10) that, since $W \in \mathbb{X}$, for each t, $z(t; W) \in \mathbb{H}^{1/2}$ as required. Furthermore, for all $t \in (0, T)$

$$\| z(t; W) \|_{\mathbb{H}^{1/2}} \leq c \| W \|_{\mathbb{X}}. \qquad (11)$$

Now we turn to the continuity in time. Arguing similarly, we have that

$$\left\| \sum_{k=M}^{\infty} \left(\int_0^t w_k(s) v\lambda_k e^{(t-s)(-v\lambda_k)} ds \right) e_k \right\|_{\mathbb{H}^{1/2}}$$

$$\leq \int_0^t \frac{c}{(t-s)^{1-\epsilon/2}} \left\| \sum_{k=M}^{\infty} w_k(s) e_k \right\|_{\mathbb{H}^{1/2+\epsilon}} ds$$

$$\leq \left(\int_0^t \frac{c}{(t-s)^{(1-\epsilon/2)p}} ds \right)^{1/p} \left(\int_0^t \left\| \sum_{k=M}^{\infty} w_k(s) e_k \right\|_{\mathbb{H}^{1/2+\epsilon}}^q ds \right)^{1/q},$$

for all $p, q > 0$ such that $1/p + 1/q = 1$. From the Lebesgue dominated convergence theorem,

$$\lim_{M\to\infty} \int_0^t \left\| \sum_{k=M}^{\infty} w_k(s) e_k \right\|_{\mathbb{H}^{1/2+\epsilon}}^q ds = 0;$$

and when p sufficiently close to 1,

$$\int_0^t \frac{c}{(t-s)^{(1-\epsilon/2)p}} ds$$

is finite. We then deduce that

$$\lim_{M\to\infty} \left\| \sum_{k=M}^{\infty} \left(\int_0^t w_k(s) v\lambda_k e^{(t-s)(-v\lambda_k)} ds \right) e_k \right\|_{\mathbb{H}}^{1/2} = 0,$$

uniformly for all t.

Fixing $t \in (0, T)$ we show that

$$\lim_{t'\to t} \|z(t; W) - z(t'; W)\|_{\mathbb{H}^{1/2}} = 0.$$

We have

$$\|z(t; W) - z(t'; W)\|_{\mathbb{H}^{1/2}} \leq \|W(t) - W(t')\|_{\mathbb{H}^{1/2}}$$

$$+ \left\| \sum_{k=1}^{M-1} \left(\int_0^t w_k(s) v\lambda_k e^{(t-s)(-v\lambda_k)} ds - \int_0^{t'} w_k(s) v\lambda_k e^{(t'-s)(-v\lambda_k)} ds \right) e_k \right\|_{\mathbb{H}^{1/2}}$$

$$+ \left\| \sum_{k=M}^{\infty} \left(\int_0^t w_k(s) \nu \lambda_k e^{(t-s)(-\nu\lambda_k)} ds \right) e_k \right\|_{\mathbb{H}^{1/2}}$$

$$+ \left\| \sum_{k=M}^{\infty} \left(\int_0^{t'} w_k(s) \nu \lambda_k e^{(t'-s)(-\nu\lambda_k)} ds \right) e_k \right\|_{\mathbb{H}^{1/2}}.$$

For $\delta > 0$, when M is sufficiently large, the argument above shows that

$$\left\| \sum_{k=M}^{\infty} \left(\int_0^t w_k(s) \nu \lambda_k e^{(t-s)(-\nu\lambda_k)} ds \right) e_k \right\|_{\mathbb{H}^{1/2}}$$

$$+ \left\| \sum_{k=M}^{\infty} \left(\int_0^{t'} w_k(s) \nu \lambda_k e^{(t'-s)(-\nu\lambda_k)} ds \right) e_k \right\|_{\mathbb{H}^{1/2}} < \delta/3.$$

Furthermore, when $|t' - t|$ is sufficiently small,

$$\left\| \sum_{k=1}^{M-1} \left(\int_0^t w_k(s) \nu \lambda_k e^{(t-s)(-\nu\lambda_k)} ds - \int_0^{t'} w_k(s) \nu \lambda_k e^{(t'-s)(-\nu\lambda_k)} ds \right) e_k \right\|_{\mathbb{H}^{1/2}} < \delta/3.$$

Finally, since $W \in \mathbb{X}$, for $|t' - t|$ is sufficiently small we have

$$\|W(t) - W(t')\|_{\mathbb{H}^{1/2}} < \delta/3$$

Thus when $|t' - t|$ is sufficiently small, $\|z(t; W) - z(t'; W)\|_{\mathbb{H}^{1/2}} < \delta$. The conclusion follows. □

Having established regularity, we now show that z is indeed a weak solution of (5).

Lemma 2 *For each $\phi \in \mathbb{H}^2$, $z(t) = z(t; W)$ satisfies* (6).

Proof It is sufficient to show this for $\phi = e_k$. We have

$$\int_0^t \langle z(s), Ae_k \rangle ds = \int_0^t \langle W(s), Ae_k \rangle ds - \int_0^t \int_0^s w_k(\tau) \nu \lambda_k^2 e^{(s-\tau)(-\nu\lambda_k)} d\tau ds$$

$$= \lambda_k \int_0^t w_k(s) ds - \nu \lambda_k^2 \int_0^t w_k(\tau) \left(\int_\tau^t e^{(s-\tau)(-\nu\lambda_k)} ds \right) d\tau$$

$$= \lambda_k \int_0^t w_k(s)ds - \lambda_k \int_0^t w_k(\tau)d\tau + \lambda_k \int_0^t w_k(\tau)e^{(t-\tau)(-\nu\lambda_k)}d\tau$$

$$= \lambda_k \int_0^t w_k(\tau)e^{(t-\tau)(-\nu\lambda_k)}d\tau.$$

On the other hand,

$$\langle z(t), e_k \rangle = \langle W(t), e_k \rangle - \nu\lambda_k \int_0^t w_k(s)e^{(t-s)(-\nu\lambda_k)}ds.$$

The result then follows.   □

We now turn to the following result, which concerns v and is established on page 416 of [6], given the properties of $z(\cdot; W)$ established in the preceding two lemmas.

Lemma 3 *For each $W \in \mathbb{X}$, problem* (7) *has a unique solution v in the function space* $C(0, T; \mathbb{H}) \bigcap L^2(0, T; \mathbb{H}^1)$.

We then have the following existence and uniqueness result for the Navier-Stokes Eq. (3), more precisely for the weak form (4), driven by rough additive forcing [6]:

Proposition 1 *For each $W \in \mathbb{X}$, problem* (4) *has a unique solution $u \in C(0, T; \mathbb{H}) \bigcap L^2(0, T; \mathbb{H}^{1/2})$ such that $u - z \in L^2(0, T; \mathbb{H}^1)$.*

Proof A solution u for (4) can be taken as

$$u(t; W) = z(t; W) + v(t; W). \tag{12}$$

From the regularity properties of z and v in Lemmas 1 and 3, we deduce that $u \in C(0, T; \mathbb{H}) \bigcap L^2(0, T; \mathbb{H}^{1/2})$. Assume that $\bar{u}(t; W)$ is another solution of (4). Then $\bar{v}(t; W) = \bar{u}(t; W) - z(t; W)$ is a solution in $C(0, T; \mathbb{H}) \bigcap L^2(0, T; \mathbb{H}^1)$ of (7). However, (7) has a unique solution in $C(0, T; \mathbb{H}) \bigcap L^2(0, T; \mathbb{H}^1)$. Thus $\bar{v} = v$.   □

2.4 Continuity of the forward map

The purpose of this subsection is to establish continuity of the forward map from W into the weak solution u of (3), as defined in (4), at time $t > 0$. In fact we prove continuity of the forward map from (u_0, W) into u and for this it is useful to define the space $\mathcal{H} = \mathbb{H} \times \mathbb{X}$ and denote the solution u by $u(t; u_0, W)$.

Theorem 1 *For each $t > 0$, the solution $u(t; \cdot, \cdot)$ of* (3) *is a continuous map from $\mathcal{H}$ into $\mathbb{H}$.*

Proof First we fix the initial condition and just write $u(t; W)$ for simplicity. We consider Eq. (3) with driving $W \in \mathbb{X}$ given by (9) and by $W' \in \mathbb{X}$ defined by

$$W'(s) = \sum_{k=1}^{\infty} w_k{}'(s) e_k \in \mathbb{X}.$$

We will prove that, for W, W' from a bounded set in $\mathbb{X}$, there is $c = c(T) > 0$, such that

$$\sup_{t \in (0,T)} \|z(t; W) - z(t; W')\|_{\mathbb{H}^{1/2}} \leq c \|W - W'\|_{\mathbb{X}} \tag{13}$$

and, for each $t \in (0, T)$,

$$\|v(t; W) - v(t; W')\|_{\mathbb{H}}^2 \leq c \sup_{s \in (0,T)} \|z(s; W) - z(s; W')\|_{L^4(D)}^2. \tag{14}$$

This suffices to prove the desired result since Sobolev embedding yields, from (14),

$$\|v(t; W) - v(t; W')\|_{\mathbb{H}}^2 \leq c \sup_{s \in (0,T)} \|z(s; W) - z(s; W')\|_{\mathbb{H}^{\frac{1}{2}}}^2. \tag{15}$$

Since $u = z + v$ we deduce from (13) and (15) that u as a map from $\mathbb{X}$ to $\mathbb{H}$ is continuous.

To prove (13) we note that

$$\|z(t; W) - z(t; W')\|_{\mathbb{H}^{\frac{1}{2}}} \leq \|W(t) - W'(t)\|_{\mathbb{H}^{\frac{1}{2}}}$$
$$+ \left\| \int_0^t \nu A e^{-\nu A(t-s)} \left(W(s) - W'(s) \right) ds \right\|_{\mathbb{H}^{\frac{1}{2}}}$$

so that

$$\sup_{t \in (0,T)} \|z(t; W) - z(t; W')\|_{\mathbb{H}^{\frac{1}{2}}} \leq \|W - W'\|_{\mathbb{X}}$$
$$+ \sup_{t \in (0,T)} \left\| \int_0^t \nu A e^{-\nu A(t-s)} \left(W(s) - W'(s) \right) ds \right\|_{\mathbb{H}^{\frac{1}{2}}}.$$

Thus it suffices to consider the last term on the right hand side. We have

$$
\left\| \int_0^t A e^{-\nu A(t-s)} \left(W(s) - W'(s) \right) ds \right\|_{\mathbb{H}^{\frac{1}{2}}}^2
$$

$$
= \left\| \sum_{k=1}^{\infty} \int_0^t \lambda_k e^{(t-s)(-\nu\lambda_k)} (w_k'(s) - w_k(s)) e_k ds \right\|_{\mathbb{H}^{1/2}}^2
$$

$$
= \sum_{k=1}^{\infty} \lambda_k^{1/2} \left(\int_0^t \lambda_k e^{(t-s)(-\nu\lambda_k)} (w_k'(s) - w_k(s)) ds \right)^2
$$

$$
\leq \sum_{k=1}^{\infty} \lambda_k^{1/2} \left(\int_0^t \lambda_k e^{(t-s)(-\nu\lambda_k)} |w_k'(s) - w_k(s)| ds \right)^2
$$

$$
\leq \sum_{k=1}^{\infty} \lambda_k^{1/2} \left(\int_0^t \frac{c}{(t-s)^{1-\epsilon/2}} \lambda_k^{\epsilon/2} |w_k'(s) - w_k(s)| ds \right)^2
$$

where we have used the fact that $a^{1-\epsilon/2} e^{-a} < c$ for all $a > 0$ for an appropriate constant c. From this, we deduce that

$$
\left\| \int_0^t A e^{-\nu A(t-s)} \left(W(s) - W'(s) \right) ds \right\|_{\mathbb{H}^{\frac{1}{2}}}
$$

$$
\leq \left\| \sum_{k=1}^{\infty} \int_0^t \frac{c}{(t-s)^{1-\epsilon/2}} \lambda_k^{\epsilon/2} |w_k'(s) - w_k(s)| e_k ds \right\|_{\mathbb{H}^{1/2}}
$$

$$
\leq \int_0^t \frac{c}{(t-s)^{1-\epsilon/2}} \left\| \sum_{k=1}^{\infty} \lambda_k^{\epsilon/2} |w_k'(s) - w_k(s)| e_k \right\|_{\mathbb{H}^{1/2}} ds
$$

$$
\leq \int_0^t \frac{c}{(t-s)^{1-\epsilon/2}} ds \sup_{s \in (0,T)} \left\| \sum_{k=1}^{\infty} \lambda_k^{\epsilon/2} |w_k'(s) - w_k(s)| e_k \right\|_{\mathbb{H}^{1/2}}
$$

$$
\leq c \sup_{s \in (0,T)} \| W'(s) - W(s) \|_{\mathbb{V}}.
$$

Therefore (13) holds.

We now prove (14). We will use the following estimate for the solution v of (7) which is proved in Flandoli [6], page 412, by means of a Gronwall argument:

$$\sup_{s \in (0,T)} \|v(s)\|_{\mathbb{H}}^2 + \int_0^T \|v(s)\|_{\mathbb{H}^1}^2 \le C\left(T, \sup_{s \in (0,T)} \|z(s)\|_{L^4(D)}\right). \tag{16}$$

We show that the map $C([0,T]; L^4(D)) \ni z(\cdot; W) \mapsto v(\cdot; W) \in \mathbb{H}$ is continuous.

For W and W' in $\mathbb{X}$, define $v = v(t; W), v' = v(t; W'), z = z(t; W), z' = z(t; W'), e = v - v'$ and $\delta = z - z'$. Then we have

$$\frac{de}{dt} + \nu A e + B(v + z, v + z) - B(v' + z', v' + z') = 0. \tag{17}$$

From this, we have

$$\frac{1}{2}\frac{d\|e\|_{\mathbb{H}}^2}{dt} + \nu\|e\|_{\mathbb{H}^1}^2 = -\langle B(v + z, v + z), e\rangle \\ + \langle B(v' + z', v' + z'), e\rangle.$$

From (2) we obtain

$$\frac{1}{2}\frac{d\|e\|_{\mathbb{H}}^2}{dt} + \nu\|e\|_{\mathbb{H}^1}^2 = +\langle B(v + z, e), v + z\rangle \\ - \langle B(v' + z', e), v' + z'\rangle = \langle B(v + z, e), v + z - v' - z'\rangle \\ - \langle B(v' + z' - v - z, e), v' + z'\rangle = \langle B(v + z, e), e + \delta\rangle \\ + \langle B(e + \delta, e), v' + z'\rangle \le (\|e\|_{L^4(D)} + \|\delta\|_{L^4(D)})(\|v\|_{L^4(D)} + \|z\|_{L^4(D)} \\ + \|v'\|_{L^4(D)} + \|z'\|_{L^4(D)})\|e\|_{\mathbb{H}^1}.$$

We now use the following interpolation inequality

$$\|e\|_{L^4(D)} \le c_0 \|e\|_{\mathbb{H}^1}^{1/2} \|e\|_{\mathbb{H}}^{1/2}, \tag{18}$$

which holds for all two dimensional domains D with constant c_0 depending only on D; see Flandoli [6]. Using this we obtain

$$\frac{1}{2}\frac{d\|e\|_{\mathbb{H}}^2}{dt} + \nu\|e\|_{\mathbb{H}^1}^2 \le c_1 \left(\|e\|_{\mathbb{H}^1}^{3/2}\|e\|_{\mathbb{H}}^{1/2} + \|\delta\|_{L^4(D)}\|e\|_{\mathbb{H}^1}\right) \\ \cdot \left(\|v\|_{L^4(D)} + \|v'\|_{L^4(D)} + \|z\|_{L^4(D)} + \|z'\|_{L^4(D)}\right)$$

for a positive constant c_1. From the Young inequality, we have

$$\|e\|_{\mathbb{H}^1}^{3/2}\|e\|_{\mathbb{H}}^{1/2}\left(\|v\|_{L^4(D)} + \|v'\|_{L^4(D)} + \|z\|_{L^4(D)} + \|z'\|_{L^4(D)}\right) \\ \le \frac{3}{4}c_2^{4/3}\|e\|_{\mathbb{H}^1}^2 + \frac{1}{4c_2^4}\|e\|_{\mathbb{H}}^2\left(\|v\|_{L^4(D)} + \|v'\|_{L^4(D)} + \|z\|_{L^4(D)} + \|z'\|_{L^4(D)}\right)^4$$

and

$$\|\delta\|_{L^4(D)}\|e\|_{\mathbb{H}^1}\left(\|v\|_{L^4(D)} + \|v'\|_{L^4(D)} + \|z\|_{L^4(D)} + \|z'\|_{L^4(D)}\right)$$
$$\leq \frac{c_3^2}{2}\|e\|_{\mathbb{H}^1}^2 + \frac{1}{2c_3^2}\|\delta\|_{L^4(D)}^2\left(\|v\|_{L^4(D)} + \|v'\|_{L^4(D)} + \|z\|_{L^4(D)} + \|z'\|_{L^4(D)}\right)^2$$

for all positive constants c_2 and c_3. Choosing c_2 and c_3 so that $c_1(3c_2^{4/3}/4 + c_3^2/2) = \nu$, we deduce that there is a positive constant c such that

$$\frac{1}{2}\frac{d\|e\|_{\mathbb{H}}^2}{dt} + \nu\|e\|_{\mathbb{H}^1}^2 \leq \nu\|e\|_{\mathbb{H}^1}^2 + c\|e\|_{\mathbb{H}}^2 \cdot I_4 + c\|\delta\|_{L^4(D)}^2 \cdot I_2 \qquad (19)$$

where we have defined

$$I_2 = \|v\|_{L^4(D)}^2 + \|v'\|_{L^4(D)}^2 + \|z\|_{L^4(D)}^2 + \|z'\|_{L^4(D)}^2$$
$$I_4 = \|v\|_{L^4(D)}^4 + \|v'\|_{L^4(D)}^4 + \|z\|_{L^4(D)}^4 + \|z'\|_{L^4(D)}^4.$$

From Gronwall's inequality, we have

$$\|e(t)\|_{\mathbb{H}}^2 \leq c \int_0^t \left(e^{\int_s^t I_4(s')ds'}\right)\|\delta(s)\|_{L^4(D)}^2 I_2(s)ds. \qquad (20)$$

Applying the interpolation inequality (18) to $v(s'; W)$, we have that

$$\int_0^T \|v(s'; W)\|_{L^4(D)}^4 ds' \leq c \sup_{s' \in (0,T)} \|v(s'; W)\|_{\mathbb{H}}^2 \int_0^T \|v(s'; W)\|_{\mathbb{H}^1}^2 ds',$$

which is bounded uniformly when W belongs to a bounded subset of $\mathbb{X}$ due to (16). Using this estimate, and a similar estimate on v', together with (11) and Sobolev embedding of $\mathbb{H}^{\frac{1}{2}}$ into $L^4(D)$, we deduce that

$$\|e(t)\|_{\mathbb{H}}^2 \leq c \sup_{0 \leq s \leq T} \|\delta(s)\|_{L^4(D)}^2.$$

We now extend to include continuity with respect to the initial condition. We show that $u(\cdot, t; u_0, W)$ is a continuous map from $\mathcal{H}$ to $\mathbb{H}$. For $W \in \mathbb{X}$ and $u_0 \in \mathbb{H}$, we consider the following equation:

$$\frac{dv}{dt} + Av + B(v + z, v + z) = 0, \quad v(0) = u_0. \qquad (21)$$

We denote the solution by $v(t) = v(t; u_0, W)$ to emphasize the dependence on initial condition and forcing which is important here. For $(u_0, W) \in \mathcal{H}$ and $(u_0', W') \in \mathcal{H}$,

from (19) and Gronwall's inequality, we deduce that

$$\|v(t; u_0, W) - v(t; u_0', W')\|_{\mathbb{H}}^2 \leq \|u_0 - u_0'\|_{\mathbb{H}}^2 e^{\int_0^t I_4(s'))ds'}$$
$$+ c \int_0^t \left(e^{\int_s^t I_4(s')ds'} \cdot \|z(s; W) - z(s; W')\|_{L^4(D)}^2 \cdot I_2(s) \right) ds.$$

We then deduce that

$$\|v(t; u_0, W) - v(t; u_0', W')\|_{\mathbb{H}}^2 \leq c\|u_0 - u_0'\|_{\mathbb{H}}^2 + c \sup_{0 \leq s \leq T} \|z(s; W) - z(s; W')\|_{L^4(D)}^2$$
$$\leq c\|u_0 - u_0'\|_{\mathbb{H}}^2 \quad + c \sup_{t \in (0,T)} \|W(t) - W'(t)\|_{\mathbb{H}^{1/2+\epsilon}}.$$

This gives the desired continuity of the forward map. □

3 Bayesian inverse problems with model error

In this section we formulate the inverse problem of determining the forcing to Eq. (3) from knowledge of the velocity field; more specifically we formulate the Bayesian inverse problem of determining the driving Brownian motion W from noisy pointwise observations of the velocity field. Here we consider the initial condition to be fixed and hence denote the solution of (3) by $u(t; W)$; extension to the inverse problem for the pair (u_0, W) is given in the following section.

We set-up the likelihood in Sect. 3.1. Then, in Sect. 3.2, we describe the prior on the forcing which is a Gaussian white-in-time process with spatial correlations, and hence a spatially correlated Brownian motion prior on W. This leads, in Sect. 3.3, to a well-defined posterior distribution, absolutely continuous with respect to the prior, and Lipschitz in the Hellinger metric with respect to the data. To prove these results we employ the framework for Bayesian inverse problems developed in Cotter et al. [3] and Stuart [18]. In particular, Corollary 2.1 of [3] and Theorem 6.31 of [18] show that, in order to demonstrate the absolute continuity of the posterior measure with respect to the prior, it suffices to show that the mapping $\mathcal{G}$ in (23) is continuous with respect to the topology of $\mathbb{X}$ and to choose a prior with full mass on $\mathbb{X}$. Furthermore we then employ the proofs of Theorem 2.5 of [3] and Theorem 4.2 of [18] to show the well-posedness of the posterior measure; indeed we show that the posterior is Lipschitz with respect to data, in the Hellinger metric.

3.1 Likelihood

Fix a set of times $t_j \in (0, T)$, $j = 1, \ldots, J$. Let ℓ be a collection of K bounded linear functionals on $\mathbb{H}$. We assume that we observe, for each j, $\ell(u(\cdot, t_j; W))$ plus a draw

from a centered K dimensional Gaussian noise ϑ_j so that

$$\delta_j = \ell\big(u(\cdot, t_j; W)\big) + \vartheta_j \tag{22}$$

is known to us. Concatenating the data we obtain

$$\delta = \mathcal{G}(W) + \vartheta \tag{23}$$

where $\delta, \vartheta \in \mathbb{R}^{JK}$ and $\mathcal{G} : \mathbb{X} \to \mathbb{R}^{JK}$. The observational noise ϑ is a draw from the JK dimensional Gaussian random variable with the covariance matrix Σ.

In the following we will define a prior measure ρ on W and then determine the conditional probability measure $\rho^\delta = \mathbb{P}(W|\delta)$. We will then show that ρ^δ is absolutely continuous with respect to ρ and that the Radon–Nikodym derivative between the measures is given by

$$\frac{d\rho^\delta}{d\rho} \propto \exp\left(-\Phi(W; \delta)\right), \tag{24}$$

where

$$\Phi(W; \delta) = \frac{1}{2} \left| \Sigma^{-\frac{1}{2}} \left(\delta - \mathcal{G}(W)\right) \right|^2. \tag{25}$$

The right hand side of (24) is the likelihood of the data δ.

3.2 Prior

We construct our prior on the time-integral of the forcing, namely W. Let Q be a linear operator from the Hilbert space $\mathbb{H}^{\frac{1}{2}+\epsilon}$ into itself with eigenvectors e_k and eigenvalues σ_k^2 for $k = 1, 2, \ldots$. We make the following assumption

Assumption 1 There is an $\epsilon > 0$ such that the eigenvalues $\{\sigma_k\}$ satisfy

$$\sum_{k=1}^{\infty} \sigma_k^2 \lambda_k^{1/2+\epsilon} < \infty$$

where λ_k are the eigenvalues of the operator A defined in Sect. 2.1.

As

$$\sum_{k=1}^{\infty} \langle Q e_k, e_k \rangle_{\mathbb{H}^{\frac{1}{2}+\epsilon}} = \sum_{k=1}^{\infty} \sigma_k^2 \lambda_k^{\frac{1}{2}+\epsilon} < \infty,$$

Q is a trace class operator in $\mathbb{H}^{\frac{1}{2}+\epsilon}$.

We assume that our prior is the Q-Wiener process W with values in $\mathbb{H}^{\frac{1}{2}+\epsilon}$ where $W(s_1) - W(s_2)$ is Gaussian in $\mathbb{H}^{\frac{1}{2}+\epsilon}$ with covariance $(s_1 - s_2)Q$ and mean 0. This process can be written as

$$W(t) = \sum_{k=1}^{\infty} \sigma_k e_k \omega_k(t), \tag{26}$$

where $\omega_k(t)$ are pair-wise independent Brownian motions (see Da Prato and Zabczyk [4], Proposition 4.1) and where the convergence of the infinite series is in the mean square norm with respect to the probability measure of the probability space that generates the randomness of W. We define by ρ the measure generated by this Q-Wiener process on $\mathbb{X}$.

Remark 1 We have constructed the solution to (3) for each deterministic continuous function $W \in \mathbb{X}$. As we equip $\mathbb{X}$ with the prior probability measure ρ, we wish to employ the results from [6] concerning the solution of (3) when W is considered as a Brownian motion obtaining values in $\mathbb{X}$. However, the solution of (3) is constructed in a slightly different way in [6] from that used in the preceding developments. We therefore show that under Assumption 1, ρ almost surely, solution u of (4) defined in (12) for each individual function W equals the unique progressively measurable solution in $C([0, T]; \mathbb{H}) \bigcap L^2([0, T]; \mathbb{H}^1)$ constructed in Flandoli [6] when the noise W is sufficiently spatially regular. This allows us to employ the existence of the second moment of $\|u(\cdot, t; W)\|^2_{\mathbb{H}}$, i.e. the finiteness of the energy $\mathbb{E}^\rho[\|u(\cdot, t; W)\|^2_{\mathbb{H}}]$, established in Mattingly [17], which we need later.

For the infinite dimensional Brownian motion W defined in (26) where

$$\sum_{k=1}^{\infty} \lambda_k^{2\beta_0 - 1/2} \sigma_k^2 < \infty,$$

for some $\beta_0 > 0$, where we employ the same notation as in [6] for ease of exposition. Flandoli [6] employs the Ornstein-Uhlenbeck process

$$z_\alpha(t) = \int_{-\infty}^{t} e^{-(\nu A + \alpha)(t-s)} dW(s) \tag{27}$$

which, considered as the stochastic process, is a solution of the Ornstein-Uhlenbeck equation

$$dz_\alpha(t) + Az_\alpha(t)dt + \alpha z_\alpha(t)dt = dW(t),$$

where α is a constant, in order to define a solution of (4). Note that if $\beta_0 > \frac{1}{2}$ then Assumption 1 is satisfied. With respect to the probability space $(\Omega, \mathcal{F}_t, \mathbb{P})$, the expectation $\mathbb{E}\|z_\alpha(t)\|^2_{\mathbb{H}^{1/2+2\beta}}$ is finite for $\beta < \beta_0$. Thus almost surely with respect to $(\Omega, \mathcal{F}_t, \mathbb{P})$, $z_\alpha(t)$ is sufficiently regular so that problem (7) with the initial condition

$v(0; W) = u_0 - z_\alpha(0)$ is well posed. The stochastic solution to the problem (3) is defined as

$$u(\cdot, t; W) = z_\alpha(t; W) + v(t; W) \tag{28}$$

which is shown to be independent of α in [6]. When $\beta_0 > \frac{1}{2}$, $\mathbb{E}\|z_\alpha(t)\|_{\mathbb{H}^1}^2$ is finite so $u(\cdot, t; W) \in C([0, T]; \mathbb{H}) \bigcap L^2([0, T]; \mathbb{H}^1)$. Flandoli [6] leaves open the question of the uniqueness of a generalized solution to (4) in $C([0, T]; \mathbb{H}) \bigcap L^2([0, T]; \mathbb{H}^{1/2})$. However, there is a unique solution in $C([0, T]; \mathbb{H}) \bigcap L^2([0, T]; \mathbb{H}^1)$.

Almost surely with respect to the probability measure ρ, solution u of (4) constructed in (12) equals the solution constructed by Flandoli [6] in (28). To see this, note that the stochastic integral

$$\int_0^t e^{-\nu A(t-s)} dW(s) \tag{29}$$

can be written in the integration by parts form (10). Therefore, with respect to ρ,

$$\mathbb{E}^\rho \left[\|z(t)\|_{L^2(0,T;\mathbb{H}^1)}^2 \right] = \int_0^T \mathbb{E}^\rho \left[\sum_{k=1}^\infty \lambda_k \sigma_k^2 \left(\int_0^t e^{-\nu\lambda_k(t-s)} d\omega_k(s) \right)^2 \right] dt$$

$$= \int_0^T \left(\sum_{k=1}^\infty \lambda_k \sigma_k^2 \int_0^t e^{-2\nu\lambda_k(t-s)} ds \right) dt$$

$$= \frac{1}{2\nu} \int_0^T \sum_{k=1}^\infty \sigma_k^2 (1 - e^{-2\nu\lambda_k t}) dt$$

which is finite. Therefore ρ almost surely, $z(t) \in L^2(0, T; \mathbb{H}^1)$. Thus $u(t; W) \in C(0, T; \mathbb{H}) \bigcap L^2(0, T; \mathbb{H}^1)$. We can then argue that ρ almost surely, the solution u constructed in (12) equals Flandoli's solution in (28) which we denote by u_α (even though it does not depend on α) as follows. As $u_\alpha \in C([0, T]; H) \bigcap L^2([0, T]; \mathbb{H}^1)$, $v_\alpha(t; W) = u_\alpha(t; W) - z(t; W) \in C([0, T]; H) \bigcap L^2([0, T]; \mathbb{H}^1)$ and satisfies (7). As for each W, (7) has a unique solution in $C([0, T]; H) \bigcap L^2([0, T]; \mathbb{H}^1)$, so $v_\alpha(t; W) = v(t; W)$. Thus almost surely, the Flandoli [6] solution equals the solution u in (12). This is also the argument to show that (3) has a unique solution in $C([0, T]; \mathbb{H}) \bigcap L^2([0, T]; \mathbb{H}^1)$.

3.3 Posterior

Theorem 2 *The conditional probability measure $\mathbb{P}(W|\delta) = \rho^\delta$ is absolutely continuous with respect to the prior measure ρ with the Radon–Nikodym derivative being*

given by (24). *Furthermore, when* $|\delta| < r$ *and* $|\delta'| < r$ *there is a constant* $c = c(r)$ *so that*

$$d_{\text{Hell}}(\rho^\delta, \rho^{\delta'}) \le c|\delta - \delta'|.$$

Proof Note that $\rho(\mathbb{X}) = 1$. It follows from Corollary 2.1 of Cotter et al. [3] and Theorem 6.31 of Stuart [18] that, in order to demonstrate that $\rho^\delta \ll \rho$, it suffices to show that the mapping $\mathcal{G} : \mathbb{X} \to \mathbb{R}^{JK}$ is continuous; then the Randon-Nikodym derivative (24) defines the density of ρ^δ with respect to ρ. As ℓ is a collection of bounded continuous linear functionals on $\mathbb{H}$, the continuity of $\mathcal{G}$ with respect to the topology of $\mathbb{X}$ follows from Theorem 1.

We now turn to the Lipschitz continuity of the posterior in the Hellinger metric. The method of proof is very similar to that developed in the proofs of Theorem 2.5 in [3] and Theorem 4.2 in [18]. We define

$$Z(\delta) := \int_{\mathbb{X}} \exp(-\Phi(W; \delta)) d\rho(W).$$

Mattingly [17] shows that for each t, the second moment $\mathbb{E}^\rho(\|u(\cdot, t; W)\|_{\mathbb{H}}^2)$ is finite. Fixing a large constant M, the ρ measure of the set of paths W such that $\max_{j=1,\ldots,J} \|u(\cdot, t_j; W)\|_{\mathbb{H}} \le M$ is larger than $1 - cJ/M^2 > 1/2$. For those paths W in this set we have,

$$\Phi(W; \delta) \le c(|\delta| + M).$$

From this, we deduce that $Z(\delta) > 0$. Next, we have that

$$|Z(\delta) - Z(\delta')| \le \int_{\mathbb{X}} |\Phi(W; \delta) - \Phi(W; \delta')| d\rho(W)$$

$$\le c \int_{\mathbb{X}} \left(|\delta| + |\delta'| + 2\sum_{j=1}^{J} |\ell(u(t_j; W))|_{\mathbb{R}^K}\right) |\delta - \delta'| d\rho(W)$$

$$\le c|\delta - \delta'|.$$

We then have

$$2d_{\text{Hell}}(\rho^\delta, \rho^{\delta'})^2 \le \int_{\mathbb{X}} \left(Z(\delta)^{-1/2} \exp(-\tfrac{1}{2}\Phi(W; \delta)) - Z(\delta')^{-1/2} \exp(-\tfrac{1}{2}\Phi(W'; \delta'))\right)^2$$

$$d\rho(W)$$

$$\le I_1 + I_2,$$

where

$$I_1 = \frac{2}{Z(\delta)} \int_{\mathbb{X}} \left(\exp(-\frac{1}{2}\Phi(W;\delta)) - \exp(-\frac{1}{2}\Phi(W;\delta')) \right)^2 d\rho(W),$$

and

$$I_2 = 2|Z(\delta)^{-1/2} - Z(\delta')^{-1/2}|^2 \int_{\mathbb{X}} \exp(-\Phi(W;\delta'))d\rho(W).$$

Using the facts that

$$\left| \exp(-\frac{1}{2}\Phi(W;\delta)) - \exp(-\frac{1}{2}\Phi(W;\delta')) \right| \le \frac{1}{2}|\Phi(W;\delta) - \Phi(W;\delta')|$$

and that $Z(\delta) > 0$, we deduce that

$$
\begin{aligned}
I_1 &\le c \int_{\mathbb{X}} |\Phi(W;\delta) - \Phi(W;\delta')|^2 d\rho(W) \\
&\le c \int_{\mathbb{X}} (|\delta| + |\delta'| + 2\sum_{j=1}^{J} |\ell(u(t_j;W))|_{\mathbb{R}^K})^2 |\delta - \delta'|^2 d\rho(W) \\
&\le c|\delta - \delta'|^2.
\end{aligned}
$$

Furthermore,

$$|Z(\delta)^{-1/2} - Z(\delta')^{-1/2}|^2 \le c \max(Z(\delta)^{-3}, Z(\delta')^{-3})|Z(\delta) - Z(\delta')|^2.$$

From these inequalities it follows that $d_{\text{Hell}}(\rho^\delta, \rho^{\delta'}) \le c|\delta - \delta'|$.   □

4 Inferring the initial condition

In the previous section we discussed the problem of inferring the forcing from the velocity field. In practical applications it is also of interest to infer the initial condition, which corresponds to a Bayesian interpretation of 4DVAR, or the initial condition and the forcing, which corresponds to a Bayesian interpretation of weak constraint 4DVAR. Thus we consider the Bayesian inverse problem for inferring the initial condition u_0 and the white noise forcing determined by the Brownian driver W. Including the initial condition does not add any further technical difficulties as the dependence on the pathspace valued forcing is more subtle than the dependence on initial condition, and this dependence on the forcing is dealt with in the previous section. As a consequence we do not provide full details.

Let ϱ be a Gaussian measure on the space $\mathbb{H}$ and let $\mu = \varrho \otimes \rho$ be the prior probability measure on the space $\mathcal{H} = \mathbb{H} \times \mathbb{X}$. We denote the solution u of (3) by $u(x, t; u_0, W)$.

We outline what is required to extend the analysis of the previous two sections to the case of inferring both initial condition and driving Brownian motion. We simplify the presentation by assuming observation at only one time $t_0 > 0$ although this is easily relaxed. Given that at $t_0 \in (0, T)$, the noisy observation δ of $\ell(u(\cdot, t_0; u_0, W)$ is given by

$$\delta = \ell(u(\cdot, t_0; u_0, W)) + \vartheta \tag{30}$$

where $\vartheta \sim N(0, \Sigma)$. Letting

$$\Phi(u_0, W; \delta) = \frac{1}{2} |\delta - \ell(u(\cdot, t_0; u_0, W))|_\Sigma^2, \tag{31}$$

we aim to show that the conditional probability $\mu^\delta = \mathbb{P}(u_0, W | \delta)$ is given by

$$\frac{d\mu^\delta}{d\mu} \propto \exp(-\Phi(u_0, W; \delta)). \tag{32}$$

We have the following result.

Theorem 3 *The conditional probability measure $\mu^\delta = \mathbb{P}(u_0, W | \delta)$ is absolutely continuous with respect to the prior probability measure μ with the Radon–Nikodym derivative give by (32). Further, when $|\delta| < r$ and $|\delta'| < r$ there is a constant $c = c(r)$ such that*

$$d_{Hell}(\mu^\delta, \mu^{\delta'}) \le c|\delta - \delta'|.$$

Proof To establish the absolute continuity of posterior with respect to prior, together with the formula for the Radon–Nikodym derivative, the key issue is establishing continuity of the forward map with respect to initial condition and driving Brownian motion. This is established in Theorem 1. Since $\mu(\mathcal{H}) = 1$ the first part of the theorem follows.

For the Lipschitz dependency of the Hellinger distance of μ^δ on δ, we use the result of Mattingly [17] which shows that, for each initial condition u_0,

$$\mathbb{E}^\rho(\|u(t; u_0, W)\|_\mathbb{H}^2) \le \frac{\mathcal{E}_0}{2\nu\lambda_1} + e^{-2\nu\lambda_1 t}\left(\|u_0\|_\mathbb{H}^2 - \frac{\mathcal{E}_0}{2\nu\lambda_1}\right),$$

where $\mathcal{E}_0 = \sum_{k=1}^\infty \sigma_k^2$. Therefore $\mathbb{E}^\mu(\|u(t; u_0, W)\|_\mathbb{H}^2)$ is bounded. This enables us to establish positivity of the normalization constants and the remainder of the proof follows that given in Theorem 2.   □

5 Numerical results

The purpose of this section is twofold: firstly to demonstrate that the Bayesian formulation of the inverse problem described in this paper forms the basis for practical numerical inversion; and secondly to study some properties of the posterior distribution on the white noise forcing, given observations of linear functionals of the velocity field.

The numerical results move outside the strict remit of our theory in two directions. Firstly we work with periodic boundary conditions; this makes the computations fast, but simultaneously demonstrates the fact that the theory is readily extended from Dirichlet to other boundary conditions. Secondly we consider both (i) pointwise observations of the entire velocity field and (ii) observations found from the projection onto the lowest eigenfunctions of A noting that the second form of observations are bounded linear functionals on $\mathbb{H}$, as required by our theory, whilst the first form of observations are not.

To extend our theory to periodic boundary conditions requires generalization of the Flandoli [6] theory from the Dirichlet to the periodic setting, which is not a technically challenging generalization. However consideration of pointwise observation functionals requires the proof of continuity of $u(t; \cdot, \cdot)$ as a mapping from $\mathcal{H}$ into $\mathbb{H}^s$ spaces for s sufficiently large. Extension of the theory to include pointwise observation functionals would thus involve significant technical challenges, in particular to derive smoothing estimates for the semigroup underlying the Flandoli solution concept. Our numerical results will show that the posterior distribution for (ii) differs very little from that for (i), which is an interesting fact in its own right.

In Sect. 5.1 we describe the numerical method used for the forward problem. In Sect. 5.2 we describe the inverse problem and the Metropolis-Hastings MCMC method used to probe the posterior. Section 5.3 describes the numerical results.

5.1 Forward problem: numerical discretization

All our numerical results are computed using a viscosity of $\nu = 0.1$ and on the periodic domain. We work on the time interval $t \in [0, 0.1]$. We use $M = 32^2$ divergence free Fourier basis functions for a spectral Galerkin spatial approximation, and employ a time-step $\delta t = 0.01$ in a Taylor time-approximation [9]. The number of basis functions and time-step lead to a fully-resolved numerical simulation at this value of ν.

5.2 Inverse problem: metropolis hastings MCMC

Recall the Stokes' operator A. We consider the inverse problem of finding the driving Brownian motion. As a prior we take a centered Brownian motion in time with spatial covariance $\pi^4 A^{-2}$; thus the space-time covariance of the process is $C_0 := \pi^4 A^{-2} \otimes (-\triangle_t)^{-1}$, where $\triangle_t$ is the Laplacian in time with fixed homogeneous Dirichlet condition at $t = 0$ and homogeneous Neumann condition at $t = T$. It is straightforward to draw samples from this Gaussian measure, using the fact that A is diagonalized in the spectral basis. Note that if $W \sim \rho$, then $W \in C(0, T; \mathbb{H}^s)$
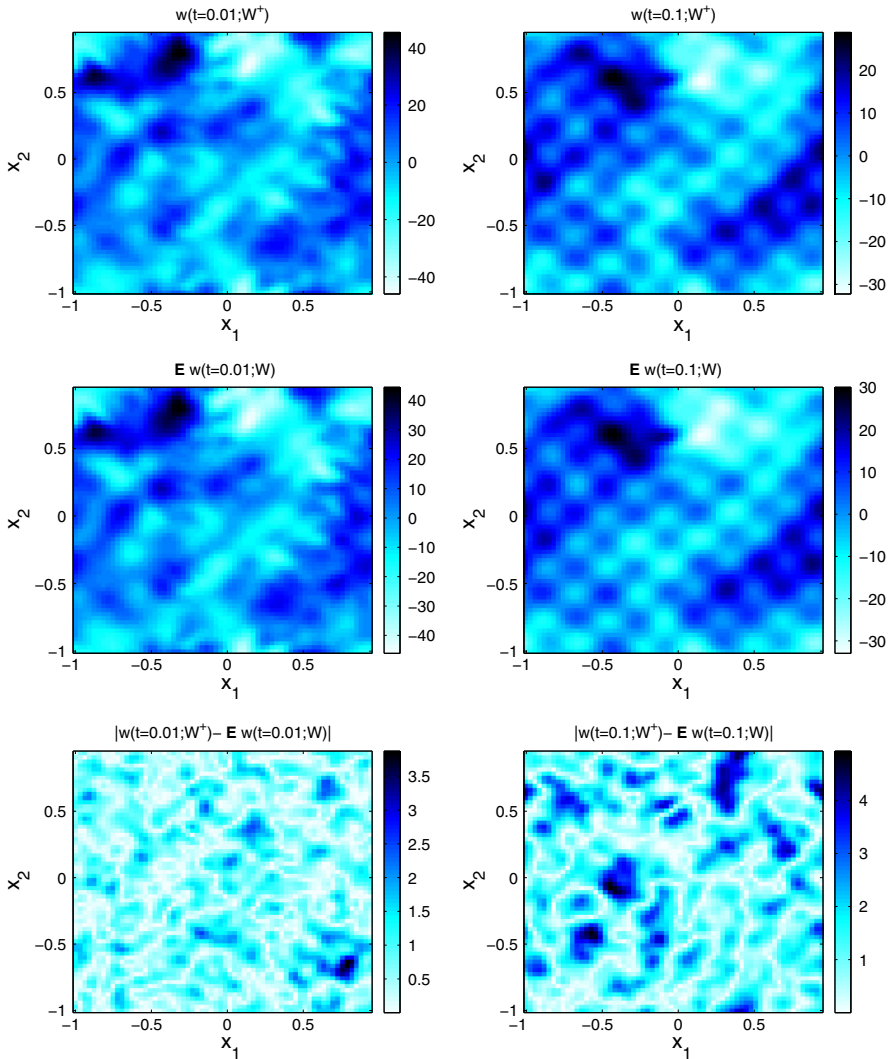
Fig. 1 Full-field point wise observations. The truth (*top*), expected value (*middle*), and absolute distance between them (*bottom*) of the vorticity $w(t; W)$, for $t = 0.01$ (*left*, relative L^2 error $e = 0.0044$) and $t = 0.1$ (*right*, $e = 0.0244$)

almost surely for all $s < 1$; in particular $W \in \mathbb{X}$. Thus $\rho(\mathbb{X}) = 1$ as required. The likelihood is defined (i) by making observations of the velocity field at every point on the 32^2 grid implied by the spectral method, at every time $t = n\delta t$, $n = 1, \cdots 10$, or (ii) by making observations of the projection onto eigenfunctions $\{\phi_k\}_{|k|<4}$ of A. The observational noise standard deviation is taken to be $\gamma = 1.6$ and all observational noises are uncorrelated.

To sample from the posterior distribution we employ a Metropolis-Hastings MCMC method. Furthermore, to ensure mesh-independent convergence properties, we use a
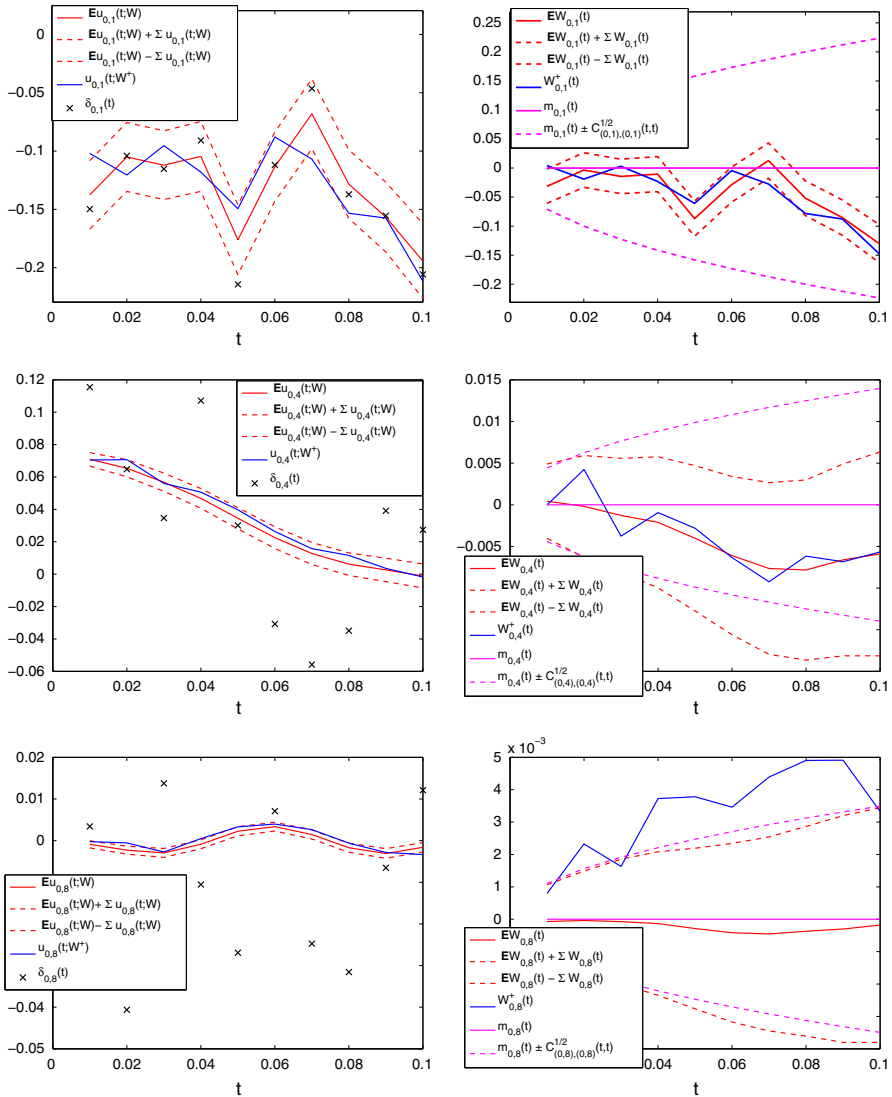
Fig. 2 Full-field point wise observations. The trajectories $u_k(t; W)$ (*left*) and W_k (*right*), with $k = (0, 1)$ (*top*), $k = (0, 4)$ (*middle*), and $k = (0, 8)$ (*bottom*). Shown are expected values and standard deviation intervals as well as true values. The *right hand* images also show the expected value and standard deviation of the prior, indicating the decreasing information content of the data for the increasing wave numbers

method which is well-defined in function space [2]. Metropolis-Hastings methods proceed by constructing a Markov kernel $\mathcal{P}$ which satisfies *detailed balance* with respect to the measure ρ^δ which we wish to sample:

$$\rho^\delta(du)\mathcal{P}(u, dv) = \rho^\delta(dv)\mathcal{P}(v, du), \quad \forall\, u, v \in \mathbb{X}. \tag{33}$$
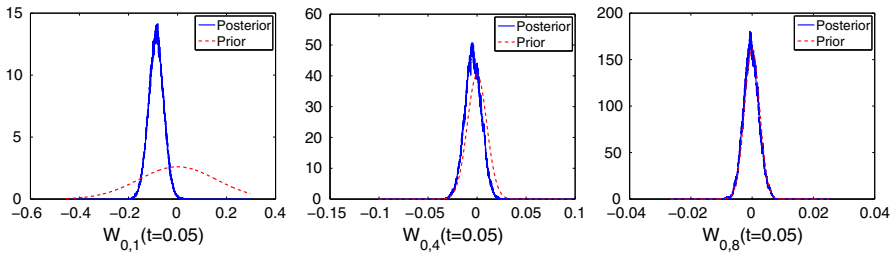
Fig. 3 Full-field point wise observations. The histograms of the posterior distribution in comparison to the prior distribution of $W_k(t = 0.05)$, for $k = (0, 1)$ (*left*), $k = (0, 4)$ (*middle*), and $k = (0, 8)$ (*right*). These plots again illustrate the decreasing information content of the data for the increasing wave numbers

Integrating with respect to u, one can see that detailed balance implies $\rho^\delta \mathcal{P} = \rho^\delta$. Metropolis-Hastings methods [8,20] prescribe an accept-reject move based on proposals from another Markov kernel $\mathcal{Q}$, in order to define a kernel $\mathcal{P}$ which satisfies detailed balance. If we define the measures

$$
\begin{aligned}
\nu(du, dv) &= \mathcal{Q}(u, dv)\rho^\delta(du) \propto \mathcal{Q}(u, dv)\exp\left(-\Phi(u; \delta)\right)\rho(du) \\
\nu^\perp(du, dv) &= \mathcal{Q}(v, du)\rho^\delta(dv) \propto \mathcal{Q}(v, du)\exp\left(-\Phi(v; \delta)\right)\rho(dv).
\end{aligned}
\tag{34}
$$

then, provided $\nu^\perp \ll \nu$, the Metropolis-Hastings method is defined as follows. Given current state u_n, a proposal is drawn $u^* \sim \mathcal{Q}(u_n, \cdot)$, and then accepted with probability

$$
\alpha(u_n, u^*) = \min\left\{1, \frac{d\nu^\perp}{d\nu}(u_n, u^*)\right\}.
\tag{35}
$$

The resulting chain is denoted by $\mathcal{P}$. If the proposal $\mathcal{Q}$ preserves the prior, so that $\rho\mathcal{Q} = \rho$, then a short calculation reveals that

$$
\alpha(u_n, u^*) = \min\left\{1, \exp\left(\Phi(u_n; \delta) - \Phi(u^*; \delta)\right)\right\};
\tag{36}
$$

thus the acceptance probability is determined by the change in the likelihood in moving from current to proposed state. We use the following pCN proposal [2] which is reversible with respect to the Gaussian prior $N(0, C_0)$:

$$
\mathcal{Q}(u_n, \cdot) = N\left(\sqrt{1 - \beta^2}u_n, \beta^2 C_0\right).
\tag{37}
$$

This hence results in the acceptance probability (36). Variants on this algorithm, which propose differently in different Fourier components, are described in [15], and can make substantial speedups in the Markov chain convergence. However for the examples considered here the basic form of the method suffices.
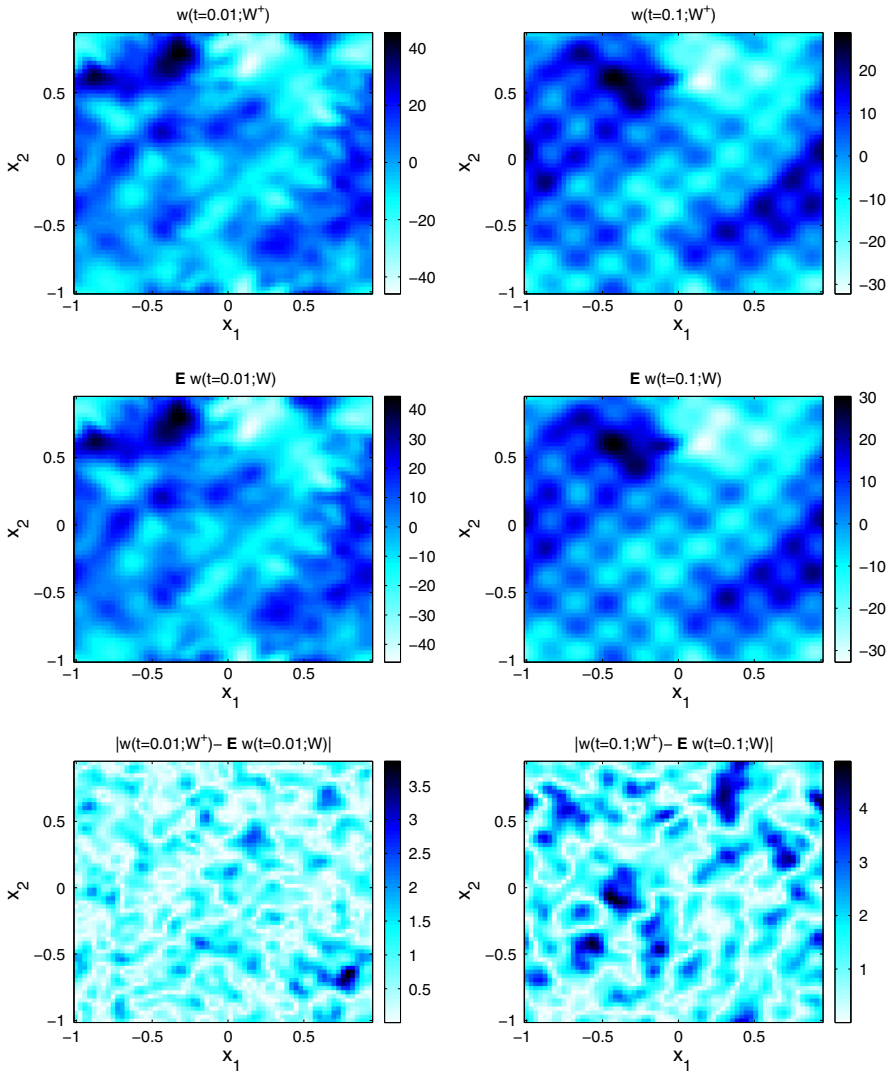
Fig. 4 Observation of Fourier modes $\{\phi_k\}_{|k|<4}$. The truth (*top*), expected value (*middle*), and absolute distance between them (*bottom*) of the vorticity $w(t; W)$, for $t = 0.01$ (*left*, relative L^2 error $e = 0.0044$) and $t = 0.1$ (*right*, $e = 0.0249$). Notice the similarity to the results of Fig. 1

5.3 Results and discussion

The true driving Brownian motion $W^\dagger$, underlying the data in the likelihood, is constructed as a draw from the prior ρ. We then compute the corresponding true trajectory $u^\dagger(t) = u(t; W^\dagger)$. We use the pCN scheme (36), (37) to sample W from the posterior distribution ρ^δ. It is important to appreciate that the object of interest here is the posterior distribution on W itself which provides estimates of the forcing, given the noisy
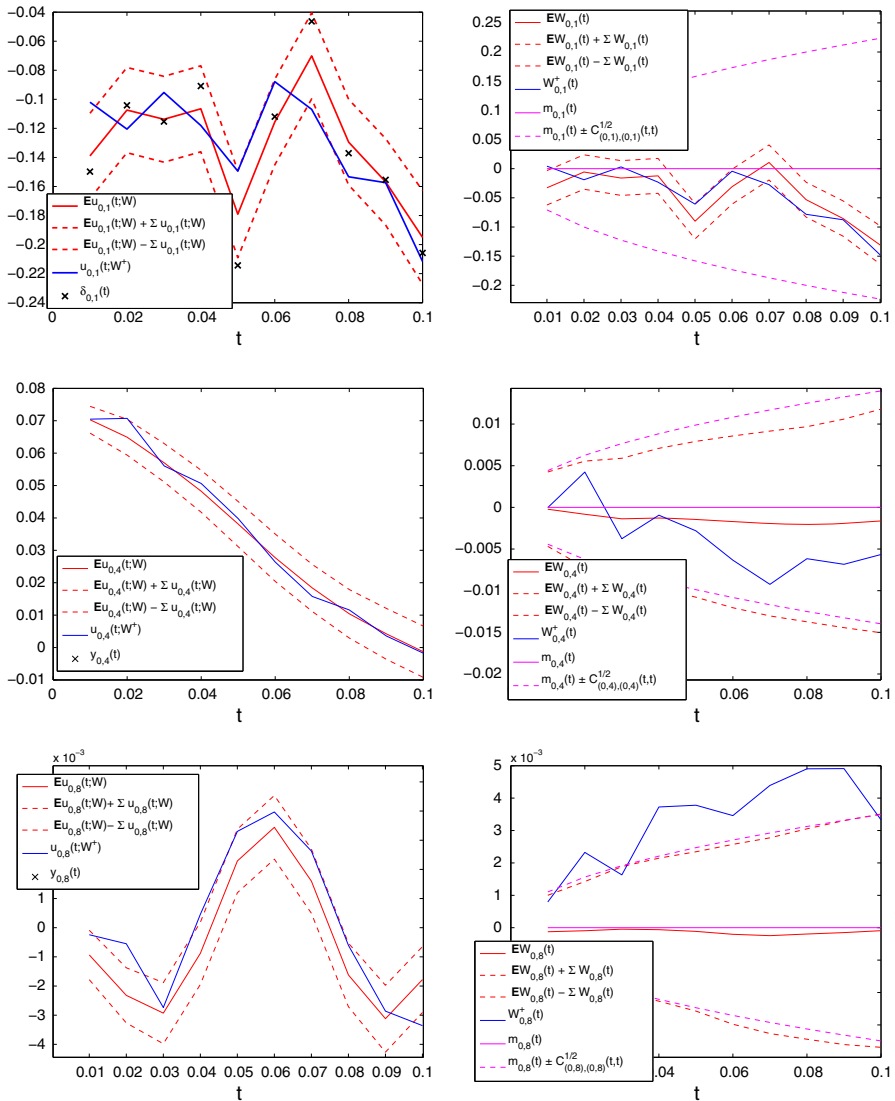
Fig. 5 Observation of Fourier modes $\{\phi_k\}_{|k|<4}$. The trajectories $u_k(t; W)$ (*left*) and W_k (*right*), with $k = (0, 1)$ (*top*), $k = (0, 4)$ (*middle*), and $k = (0, 8)$ (*bottom*). Shown are expected values and standard deviation intervals as well as true values. The *right hand* images also show the expected value and standard deviation of the prior, indicating the decreasing information content of the data for the increasing wave numbers

observations of the velocity field. This posterior distribution is not necessarily close to a Dirac measure on the truth; in fact we will show that some parameters required to define W are recovered accurately whilst others are not.

We first consider the observation set-up (i) where pointwise observations of the entire velocity field are made. The true initial and final conditions are plotted in Fig. 1,
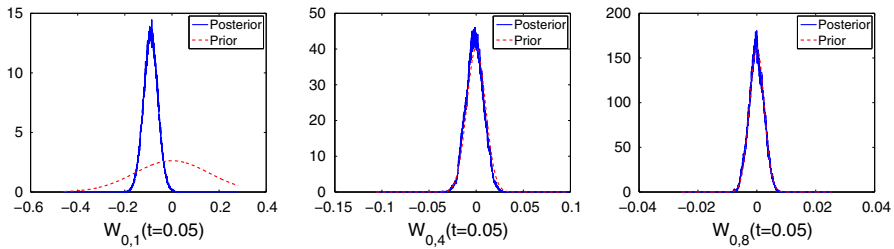
Fig. 6 Observation of Fourier modes $\{\phi_k\}_{|k|<4}$. The histograms of the posterior distribution in comparison to the prior distribution of $W_k(t = 0.05)$, for $k = (0, 1)$ (*left*), $k = (0, 4)$ (*middle*), and $k = (0, 8)$ (*right*). These plots again illustrate the decreasing information content of the data for the increasing wave numbers. Notice the *middle panel* in which one notices the posterior on $W_{0,4}(t = 0.05)$ is much closer to the prior than in Fig. 3

top two panels, for the vorticity field w; the middle two panels of Fig. 1 show the posterior mean of the same quantities and indicate that the data is fairly informative, since they closely resemble the truth; the bottom two panels of Fig. 1 show the absolute difference between the fields in the top and middle panels. The true trajectory, together with the posterior mean and one standard deviation interval around the mean, are plotted in Fig. 2, for the wavenumbers $(0, 1)$, $(0, 4)$, and $(0, 8)$, and for both the driving Brownian motion W (right) and the velocity field u (left). This figure indicates that the data is very informative about the $(0, 1)$ mode, but less so concerning the $(0, 4)$ mode, and there is very little information in the $(0, 8)$ mode. In particular for the $(0, 8)$ mode the mean and standard deviation exhibit behaviour similar to that under the prior whereas for the $(0, 1)$ mode they show considerable improvement over the prior in both position of the mean and width of standard deviations. The posterior on the $(0, 4)$ mode has gleaned some information from the data as the mean has shifted considerably from the prior; the variance remains similar to that under the prior, however, so uncertainty in this mode has not been reduced. Figure 3 shows the histograms of the prior and posterior for the same 3 modes as in Fig. 2 at the center time $t = 0.05$. One can see here even more clearly that the data is very informative about the $(0, 1)$ mode in the left panel, less so but somewhat about the $(0, 4)$ mode in the center panel, and it is not informative at all about the $(0, 8)$ mode in the right panel.

Figures 4, 5, and 6 are the same as Figs. 1, 2, and 3 except for the case of (ii) observation of low Fourier modes. Notice that the difference in the spatial fields are difficult to distinguish by eye, and indeed the relative errors even agree to threshold 10^{-3}. However, we can see that now the unobserved $(0, 4)$ mode in the center panels of Figs. 5 and 6 is not informed by the data and remains distributed approximately like the prior.

References

1. Bennett, A.F.: Inverse Modeling of the Ocean and Atmosphere. Cambridge University Press, Cambridge (2002)
2. Cotter, S., Roberts, G., Stuart, A., White, D.: MCMC methods for functions: modifying old algorithms to make them faster. Stat. Sci. **28**(3), 424–446 (2013)
3. Cotter, S.L., Dashti, M., Robinson, J.C., Stuart, A.M.: Bayesian inverse problems for functions and applications to fluid mechanics. Inverse Probl. **25**, 115008 (2009)
4. Da Prato, G., Zabczyk, J.: Stochastic Equations in Infinite Dimensions. Cambridge University Press, Cambridge (2008)
5. Dashti, M., Law, K.J.H., Stuart, A.M., Voss, J.: Map estimators and posterior consistency in Bayesian nonparametric inverse problems. Inverse Probl. **29**, 095017 (2013)
6. Flandoli, F.: Dissipative and invariant measures for stochastic Navier-Stokes equations. N0DEA 1, 403–423 (1994).
7. Hairer, M., Stuart, A.M., Voss, J.: Signal processing problems on function space: Bayesian formulation, stochastic PDEs and effective MCMC methods. In: Crisan, D., Rozovsky, B. (eds.) The Oxford Handbook of Nonlinear Filtering, pp. 833–873. Oxford University Press, Oxford (2011)
8. Hastings, W.K.: Monte Carlo sampling methods using Markov chains and their applications. Biometrika **57**, 97–109 (1970)
9. Jentzen, A., Kloeden, P.: Taylor expansions of solutions of stochastic partial differential equations with additive noise. Ann. Probab. **38**(2), 532–569 (2010)
10. Kaipio, J., Somersalo, E.: Statistical and computational inverse problems. In: Applied Mathematical Sciences. vol. 160, Springer, New York (2004).
11. Lasanen, S.: Discretizations of generalized random variables with applications to inverse problems. University of Oulu, Ann. Acad. Sci. Fenn. Math. Diss. (2002)
12. Lasanen, S.: Measurements and infinite-dimensional statistical inverse theory. PAMM **7**, 1080101–1080102 (2007)
13. Lasanen, S.: Non-Gaussian statistical inverse problems. Part I: posterior distributions. Inverse Probl. Imaging **6**(2), 215–266 (2012)
14. Lasanen, S.: Non-Gaussian statistical inverse problems. Part II: posterior convergence for approximated unknowns. Inverse Probl. Imaging **6**(2), 267–287 (2012)
15. Law, K.J.H.: Proposals which speed-up function space MCMC. J. Comput. Appl. Math **262**, 127–138 (2014)
16. Lorenc, A.C.: The potential of the ensemble Kalman filter for NWP a comparison with 4D-Var. Quart. J. R. Meteorol. Soc. **129**(595), 3183–3203 (2003)
17. Mattingly, J.C.: Ergodicity of 2D Navier-Stokes equations with random forcing and large viscosity. Commun. Math. Phys. **206**, 273–288 (1999)
18. Stuart, A.M.: Inverse problems: a Bayesian perspective. Acta Numer. **19**(1), 451–559 (2010)
19. Temam, R.: Navier-Stokes Equations. American Mathematical Society, New York (1984)
20. Tierney, L.: A note on Metropolis-Hastings kernels for general state spaces. Ann. Appl. Probab. **8**(1), 1–9 (1998)
21. Vollmer, S.J.: Dimension-independent MCMC sampling for elliptic inverse problems with non-Gaussian priors. arXiv:1302.2213, (2013).
22. Zupanski, D.: A general weak constraint applicable to operational 4DVAR data assimilation systems. Monthly Weather Rev. **125**(9), 2274–2292 (1997)