Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications 121 (2011) 701–724
stochastic processes and their applications
www.elsevier.com/locate/spa

Parameter estimation for the stochastically perturbed Navier–Stokes equations

Igor Cialenco[a,*], Nathan Glatt-Holtz[b]

[a] *Department of Applied Mathematics, Illinois Institute of Technology, 10 West 32nd Str, Bld E1, Room 208, Chicago, IL 60616-3793, United States*
[b] *Department of Mathematics, Indiana University, 303 Swain East, Bloomington, IN 47405, United States*

Received 19 June 2010; received in revised form 23 October 2010; accepted 15 December 2010
Available online 22 December 2010

Abstract

We consider a parameter estimation problem of determining the viscosity ν of a stochastically perturbed 2D Navier–Stokes system. We derive several different classes of estimators based on the first N Fourier modes of a single sample path observed on a finite time interval. We study the consistency and asymptotic normality of these estimators. Our analysis treats strong, pathwise solutions for both the periodic and bounded domain cases in the presence of an additive white (in time) noise.

MSC: 60H15; 35Q30; 65L09

Keywords: Parameter estimation; Inverse problems; Nonlinear stochastic partial differential equations; Navier–Stokes equations; Maximum likelihood estimators; Stochastic evolution equations; Estimation of viscosity

1. Introduction

The theory of stochastic partial differential equations (SPDEs) is a rapidly developing field of pure and applied mathematics. These equations are used to describe the evolution of dynamical systems in the presence of persistent spatial–temporal uncertainties. When considering nonlinear processes one encounters many new, fundamental and mathematically challenging problems for SPDEs, with important applications in physics and applied sciences.

While the general form of a particular SPDE is commonly derived from the fundamental properties of the underlying processes under study, frequently parameters arise in the formulation

* Corresponding author. Tel.: +1 312 567 3131.
E-mail addresses: igor@math.iit.edu (I. Cialenco), negh@indiana.edu (N. Glatt-Holtz).
URLs: http://math.iit.edu/~igor (I. Cialenco), http://mypage.iu.edu/~negh (N. Glatt-Holtz).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2010.12.007

which need to be specified or determined on the basis of some sort of empirical observation. In such situations, the so called problem of *parameter estimation* arises naturally: *under the assumption that a phenomenon of interest follows the dynamics of an* SPDE, *and given that some realizations of this process are measured, we wish to find the unknown parameters appearing in the model, such that the equations fit or predict as much as possible the observed data.*

Actually, the development of methods for estimating parameters appearing in a model serve practical purposes for two reasons. On the one hand we may be confident in the model, but have an incomplete knowledge of the physical parameters appearing therein. An "estimator" of the true parameter therefore provides a means for measuring these unknowns. On the other hand, we may already possess accurate knowledge of the physical quantities involved in the model, but lack confidence in the validity of the underlying model. In this situation finding an "estimator" will be the first step in testing and validating the model.

Since the solution of an SPDE is a random variable, this inverse problem of finding the true parameters is treated by methods from stochastic analysis and statistics. In this work we will follow a continuous time approach and assume that the solution $U = U_\nu(t, \omega)$ of the SPDE is observed for every time t over an interval $[0, T]$. We note that different kinds of methods and approaches are used to study inverse problems for deterministic PDEs, and we refer the reader to [20,21] and references therein.

A core notion in the theory of statistical inference for stochastic processes is the so called 'regularity' of the family of probability measures associated with the set of possible values Θ of the parameter of interest ν. Note that ν could be a vector in general. Let H be the function space where the solution evolves and for each $\nu \in \Theta$ denote by $\mathbb{P}_\nu^T$ the probability measures on $C([0, T]; H)$ generated by the solution $U_\nu(t)$, $0 \le t \le T$. We say that a model is 'regular' if any two probability measures from the family $\{\mathbb{P}_\nu^T, \nu \in \Theta\}$ are mutually absolutely continuous. On the other hand the model is said to be 'singular' if these measures are mutually singular.

For regular models one approach to the parameter estimation problem is to consider the maximum likelihood estimator (MLE) $\widehat{\nu}$ of ν. This type of estimator is obtained by fixing a reference value ν_0 and then maximizing the Radon–Nikodym derivative or likelihood ratio $\mathrm{d}\mathbb{P}_\nu^T/\mathrm{d}\mathbb{P}_{\nu_0}^T$ with respect to ν. Usually $\widehat{\nu} \neq \nu$ and the problem is to study the convergence of these estimators to the true parameter as more information arrives (for example as time passes or by decreasing the amplitude of the noise). In contrast, each singular model requires an individual approach, and usually the true parameter can be found exactly, without any limiting procedure (at least if the solution is observed continuously).

Statistical inference for finite dimensional systems of stochastic differential equations (SDEs) has been studied widely and provides instructive examples of both 'regular' and 'singular' problems. Typically estimating the drift coefficient for an SDE is a regular problem which may be treated with an MLE. Here the likelihood ratio can be determined by Girsanov-type theorems. By contrast, estimating the diffusion coefficient is a singular problem and in this case one can find the diffusion coefficient by measuring the quadratic variation of the process. In general there exist necessary and sufficient conditions for the regularity for (finite dimensional) SDEs. See the monographs [22,25], and references therein, for a comprehensive treatment.

It turns out that the parameter estimation problem for infinite dimensional systems (SPDEs) is, in many cases, a singular problem where one can find the parameter ν "exactly" on any finite interval of time. In particular this has been shown in the case of linear stochastic parabolic equations with the parameter of interest in the drift appearing next to the highest order differential operator. Note that this is in direct contrast to the case for most of the corresponding finite dimensional processes where one has to observe a sample path over an infinite time horizon or to decrease the amplitude of the noise term in order to get similar results. One of the first significant works in the theory of statistical inference for SPDEs that explores this singularity is [17]. The idea in this work is to approximate the original singular problem by a sequence of regular problems for which MLEs exist; this approximation is carried out by considering Galerkin-type projections of the solution onto a finite dimensional space where the estimation problem becomes regular. They prove that as the dimension of the projection increases, the corresponding MLEs converge to the true parameter. In [18,19,27,28], the problem has been extended to a general class of linear parabolic SPDEs driven by additive noise and the convergence of the estimators has been classified in terms of the order of the corresponding differential operators. For recent developments and other kinds of inference problems for *linear* SPDEs see the survey paper [26] and references therein.

While the linear theory has been extensively studied in the framework described above it seems that, to the best of our knowledge, no similar results have been established for *nonlinear* SPDEs. We therefore embark in this and concurrent work [5] on a study of parameter estimation problems for certain fundamental nonlinear SPDEs from fluid dynamics.

Note that for the linear case, key properties such as efficiency and asymptotic normality of the estimators are proven by making essential use of the exact long time behavior of the moments of the Fourier coefficients of the solutions. In the case of nonlinear equations, for example stochastic equations from mathematical fluid dynamics, the problem is much more delicate, due to the (highly nontrivial) coupling of the Fourier modes.

From the point of view of applications this work is motivated in particular by recent developments in the area of geophysical fluid dynamics (GFD) where the theory of SPDEs is now playing an important role. See, for example, [31–33,11,16,15,9]. For this developing field, novel ‘inverse’ methods are clearly needed. While the problems that we consider initially are toy models in comparison to large scale circulation models such as the primitive equations, we are optimistic that the methods and insights developed for simple nonlinear SPDEs will eventually serve the wider goal of extending our understanding to a more physically realistic setting.

In this work we consider the $2D$ Navier–Stokes equations forced with an additive white noise:

$$dU + ((U \cdot \nabla)U - \nu \Delta U + \nabla P)dt = \sigma dW, \tag{1.1a}$$

$$\nabla \cdot U = 0, \tag{1.1b}$$

$$U(0) = U_0, \tag{1.1c}$$

which describe the flow of a viscous, incompressible fluid. Here $U = (U_1, U_2)$ and P respectively represent the velocity field and the pressure. The coefficient $\nu > 0$ corresponds to the kinematic viscosity of the fluid, and it will be the parameter of interest. The goal of our analysis will be to find a suitable estimator $\hat{\nu} = \hat{\nu}(\omega)$ which is a functional of a single sample path $U(\omega)$ observed over a finite and fixed time interval $[0, T]$.

We assume that the governing equations (1.1) evolve over a domain $\mathcal{D}$. Throughout this work we will consider two possible boundary conditions. On the one hand we may suppose that the flow occurs over all of $\mathbb{R}^2$, take $\mathcal{D} = [-L/2, L/2]^2$ for some $L > 0$ and prescribe the *periodic boundary condition*:

$$U(\mathbf{x} + L\mathbf{e}_j, t) = U(\mathbf{x}, t), \quad \text{for all } \mathbf{x} \in \mathbb{R}^2, t \geq 0; \quad \int_{\mathcal{D}} U(\mathbf{x})d\mathbf{x} = \mathbf{0}.^{1} \tag{1.2}$$

[1] This second condition may be added with no loss of generality and slightly simplifies the analysis. See for instance [37].

We also consider the case when $\mathcal{D}$ is a bounded subset of $\mathbb{R}^2$ with a smooth boundary $\partial\mathcal{D}$ and assume the *Dirichlet (no-slip) boundary condition*:

$$U(\mathbf{x}, t) = 0 \quad \text{for all } \mathbf{x} \in \partial\mathcal{D}, \quad t \geq 0. \tag{1.3}$$

The stochastic forcing that we consider is an additive space–time noise colored in space. Formally, we may write

$$\sigma \mathrm{d}W = \sum_k \lambda_k^{-\gamma} \Phi_k \mathrm{d}W_k, \tag{1.4}$$

where Φ_k are the eigenfunctions of the Stokes operator, λ_k represent the associated eigenvalues, and $W_k, k \geq 1$, are one-dimensional independent Brownian motions. We assume that γ is a real parameter greater than 1 which guarantees some spatial smoothness in the forcing. We may also formally derive (see e.g. [8]) the space–time correlation structure of the noise term

$$\mathbb{E}(\sigma \mathrm{d}W(\mathbf{x}, t)\sigma \mathrm{d}W(\mathbf{y}, s)) = K(\mathbf{x}, \mathbf{y})\delta_{t-s},$$

where $K(\mathbf{x}, \mathbf{y}) = \sum_{k\geq 1} \lambda_k^{-2\gamma} \Phi_k(\mathbf{x}) \Phi_k(\mathbf{y})$.

We should mention that the stochastic Navier–Stokes equations in both two and three dimensions and under much more general stochastic forcing conditions have been extensively studied. See, for instance, [2–4,7,12,14,30] and references therein.

Since the parameter of interest ν appears next to the highest order differential operator, the linear analogue of (1.1) is singular, as we described above. With this in mind we expected that the full nonlinear model might also be singular. As developed below, ν may be found exactly from a single observation over a finite time window which suggests that this singular structure is preserved in this nonlinear case.

The starting point of our analysis, the derivation of an estimator for ν, follows methods already developed for the linear case (see references mentioned above). We project (1.1) down to a finite dimensional space, and for each N we arrive at a system of the form

$$\mathrm{d}U^N + (\nu A U^N + P_N B(U))\mathrm{d}t = P_N \sigma \mathrm{d}W, \quad U(0) = U_0,$$

where P_N is the projection operator on the finite dimensional space generated by the first N Fourier eigenvalues of the Stokes operator. We then formally compute the MLEs associated with these systems, and take them as an ansatz for our estimators. In the course of the analysis we introduce an additional degree of freedom, a parameter α, which we may carefully tune to compensate for the nonlinear term. We arrive finally at the following three classes of estimators:

$$\begin{aligned}
\widetilde{\nu}_N &= -\frac{\int_0^T \langle A^{1+2\alpha} U^N, \mathrm{d}U^N\rangle + \int_0^T \langle A^{1+2\alpha} U^N, P_N B(U)\rangle \mathrm{d}t}{\int_0^T |A^{1+\alpha} U^N|^2 \mathrm{d}t}, \\
\check{\nu}_N &= -\frac{\int_0^T \langle A^{1+2\alpha} U^N, \mathrm{d}U^N\rangle + \int_0^T \langle A^{1+2\alpha} U^N, P_N B(U^N)\rangle \mathrm{d}t}{\int_0^T |A^{1+\alpha} U^N|^2 \mathrm{d}t}, \\
\hat{\nu}_N &= -\frac{\sum_{k=1}^N \lambda_k^{1+2\alpha} (u_k^2(T) - u_k^2(0) - T\lambda_k^{-2\gamma})}{2\sum_{k=1}^N \lambda_k^{2+2\alpha} \int_0^T u_k^2 \mathrm{d}t}.
\end{aligned} \tag{1.5}$$

Here $u_k := (U, \Phi_k)$ represents the kth (generalized) Fourier mode of the solution U.

The main result in this work establishes the following properties for the proposed estimators:

Theorem 1.1. *Suppose that $U = U(\omega)$ is a single sample path solution of (1.1), (1.2) or (1.1), (1.3) observed on a finite interval of time $[0, T]$. Assume that (1.1) is forced with a white noise process of the form (1.4) where $\gamma > 1$,*[2] *and suppose that $\alpha > \gamma - 1$. Then, given a suitably regular initial vector field U_0,*

(i) *the functionals $\widetilde{\nu}_N, \check{\nu}_N, \widehat{\nu}_N$ defined by (1.5) are weakly consistent estimators of the parameter ν, i.e.*
$$\lim_{N\to\infty} \widetilde{\nu}_N = \lim_{N\to\infty} \check{\nu}_N = \lim_{N\to\infty} \widehat{\nu}_N = \nu$$
in probability;

(ii) *if we assume further that $\alpha > \gamma - 1/2$, then $\widetilde{\nu}_N$ is asymptotically normal with rate N, i.e.*
$$N(\widetilde{\nu}_N - \nu) \xrightarrow{d} \eta$$
(converges in distribution) where η is a mean zero, normally distributed random variable.

While we are able to prove the strongest convergence results for $\widetilde{\nu}_N$, this estimator is intractable numerically and even analytically. This is because $\widetilde{\nu}_N$ depends on all of the Fourier modes of the solution in a highly nonlinear fashion. At the other extreme is $\hat{\nu}_N$ which is much more straightforward to compute but is expected to have a slower rate of convergence to the actual parameter ν. The estimator $\check{\nu}_N$ is a compromise between the two extremes since it depends only on the knowledge of the first N eigenmodes but retains some of the complex structure of the nonlinear term. Although at the present time we are not able to prove this, we expect $\check{\nu}_N$ to have a faster rate of convergence than $\hat{\nu}_N$. We conjecture, in Section 4.3, that $\check{\nu}_N$ is also asymptotically normal with the same variance and rate of convergence as $\widetilde{\nu}_N$. Given the explicit formulas for the estimators, (1.5), all these questions, including the effect of the free parameter α on the rate of convergence, can be studied by means of numerical simulations, which the authors plan to undertake in a separate forthcoming paper.

While the form of the proposed estimators and the general statements of the main results in this work are similar to those in previous works in the linear case, fundamental new difficulties arise which require one to take a novel approach for the analysis. This is of course due to the complex structure of the nonlinear term appearing in (1.1) which couples, in an intricate way, all of the modes $u_k = (U, \Phi_k)$. In contrast to the linear case, we lose for example any explicit spectral information about the elements u_k. This coupling also means that the u_k are not expected to be independent.

To overcome these difficulties the analysis relies on a careful decomposition of the solution $U = \bar{U} + R$. Here $\bar{U}$ satisfies a linear system where the modes are independent. Crucially, a complete spectral picture is obtainable for $\bar{U}$. On the other hand, R, while depending in a complicated way on the full solution U, is more regular in comparison to $\bar{U}$. This is because R is not directly forced by the noise terms σdW. For this point the analysis, particularly in the case of bounded domains, requires a delicate treatment of the nonlinear term.

Due to these technical issues, we were able to establish asymptotic normality only for $\widetilde{\nu}_N$. It is interesting that $\widehat{\nu}_N$ is a consistent estimator for ν and it is the same as the MLE of the corresponding linear equation (the stochastic Stokes equation). This effect can be explained as

[2] In the case (1.3) we assume, for technical reasons, an upper bound on γ, $\gamma < 1 + 1/4$, as well. See below.

follows: since the nonlinear term $B(U)$ is in some sense 'lower order' it fails to destroy the information about ν; ν remains observable like in the linear case.

The exposition of the paper is organized as follows. In Section 2 we lay the theoretical foundations for this work, reviewing the relevant mathematical theory for the stochastic Navier–Stokes equations. We establish some crucial spectral information concerning the linear system associated with (1.1). We also recall in this section some particular variants on the law of large numbers and the central limit theorem. Section 3 sketches the derivation of the estimators $\widetilde{\nu}, \check{\nu}, \widehat{\nu}_N$. We conclude the section with a strict formulation of the main results. The proof of the main theorem is carried out in Section 4 in a series of modular substeps. We first study the regularity of the 'residual' R that appears after we 'subtract off' the noise term appearing in (1.1) via the linear Stokes equation. As an immediate application we are able to determine some precise rates for the denominators appearing in the estimators (1.5). Using these rates we successively analyze the consistency of the estimators. The final subsection treats the question of asymptotic normality with the help of a central limit theorem for martingales.

2. The mathematical setting of the problem

We begin by recalling the mathematical background for the stochastic Navier–Stokes equations and then review some general results from probability theory that will be used in the sequel.

2.1. The stochastic Navier–Stokes equation

We first describe how (1.1) is recast as an infinite dimensional stochastic evolution equation of the form

$$\begin{aligned} &\mathrm{d}U + (\nu AU + B(U))\mathrm{d}t = \sigma \mathrm{d}W, \\ &U(0) = U_0. \end{aligned} \tag{2.1}$$

The basic function spaces are designed to capture both the boundary conditions and the divergence free nature of the flow.

We first consider the spaces associated with a *Dirichlet boundary condition* (1.3). Let $H := \{U \in L^2(\mathcal{D})^2 : \nabla \cdot U = 0, U \cdot n = 0\}$, where n is the outer pointing unit normal to $\partial\mathcal{D}$. H is endowed as a Hilbert space with the L^2 inner product $(U^\flat, U^\sharp) = \int_{\mathcal{D}} U^\flat U^\sharp \mathrm{d}x$ and associated norm $|U| = (U, U)^{1/2}$. The Leray–Hopf projector, P_H, is defined as the orthogonal projection of $L^2(\mathcal{M})^d$ onto H. We next take $V := \{U \in H_0^1(\mathcal{D})^2 : \nabla \cdot U = 0\}$ and endow this space with the inner product $((U^\flat, U^\sharp)) = \int_{\mathcal{M}} \nabla U^\flat \cdot \nabla U^\sharp \mathrm{d}\mathcal{M}$. Due to the Dirichlet boundary condition, (1.3), the Poincaré inequality $|U| \leq c\|U\|$ holds for $U \in V$, justifying this definition.

The definitions for H and V are slightly different for the case of *periodic boundary conditions* (1.2). We take $\mathcal{D} = [-L/2, L/2]^2$ and define the spaces $L^2_{per}(\mathcal{D})^2$, $H^1_{per}(\mathcal{D})^2$ to be the families of vector fields $U = U(\mathbf{x})$ which are L periodic in each direction and which belong respectively to $L^2(\mathcal{O})^2$ and $H^1(\mathcal{O})^2$ for every open bounded set $\mathcal{O} \subset \mathbb{R}^2$. We now define

$$H = \left\{ U \in L^2_{per}(\mathcal{D})^2 : \nabla \cdot U = 0, \int_{\mathcal{D}} U(\mathbf{x})\mathrm{d}\mathbf{x} = \mathbf{0} \right\},$$

and

$$V = \left\{ U \in H^1_{per}(\mathcal{D})^2 : \nabla \cdot U = 0, \int_{\mathcal{D}} U(\mathbf{x})\mathrm{d}\mathbf{x} = \mathbf{0} \right\}.$$

H and V are endowed with the norms $|\cdot|$ and $\|\cdot\|$ as above. Note that we impose the mean zero condition for H and V so that the Poincaré inequality holds. As mentioned in the introduction, there is no loss of generality in imposing this extra assumption. See, e.g., [37].

The linear portion of (1.1) is captured in the Stokes operator $A = -P_H\Delta$, which is an unbounded operator from H to H with the domain $D(A) = H^2(\mathcal{D}) \cap V$. Since A is self-adjoint, with a compact inverse $A^{-1} : H \to D(A)$, we may apply the standard theory of compact, symmetric operators to guarantee the existence of an orthonormal basis $\{\Phi_k\}_{k\geq 1}$ for H of eigenfunctions of A with the associated eigenvalues $\{\lambda_k\}_{k\geq 0}$ forming an unbounded, increasing, sequence. Moreover,

$$\lambda_k \approx \lambda_1 k, \tag{2.2}$$

where the notation $a_n \approx b_n$ means that $\lim_{n\to\infty} a_n/b_n = 1$. Also, we will write $a_n \sim b_n$ when there exists a finite, nonzero constant c such that $\lim_{n\to\infty} a_n/b_n = c$. For more details about the asymptotical behavior of $\{\lambda_k\}_{k\geq 1}$ see for instance [1,29] for the no-slip case (1.3), and [6] for the spatially periodic case (1.2). Define $H_N = \mathrm{Span}\{\Phi_1, \ldots, \Phi_N\}$, and take P_N to be the projection from H onto this space. We let $Q_N := I - P_N$.

The analysis below relies extensively on the fractional powers of A. Given $\alpha > 0$, take $D(A^\alpha) = \left\{U \in H : \sum_k \lambda_k^{2\alpha} |u_k|^2 < \infty\right\}$, where $u_k = (U, \Phi_k)$. On this set we may define A^α according to $A^\alpha U = \sum_k \lambda_k^\alpha u_k \Phi_k$, for $U = \sum_k u_k \Phi_k$. Classically we have the generalized Poincaré and inverse Poincaré estimates

$$|A^{\alpha_2} P_N U| \leq \lambda_N^{\alpha_2-\alpha_1} |A^{\alpha_1} P_N U|, \qquad |A^{\alpha_1} Q_N U| \leq \frac{1}{\lambda_N^{\alpha_2-\alpha_1}} |A^{\alpha_2} Q_N U|, \tag{2.3}$$

for any $\alpha_1 < \alpha_2$.

We next describe the stochastic terms in (1.1). Fix a stochastic basis $\mathcal{S} := (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}, \{W_k\}_{k\geq 1})$, that is a filtered probability space with $\{W_k\}_{k\geq 1}$ a sequence of independent standard Brownian motions relative to filtration $\{\mathcal{F}_t\}_{t\geq 0}$. In order to avoid unnecessary complications below we may assume that $\mathcal{F}_t$ is complete and right continuous (see [8] for more details). Writing formally $W = \sum_{k\geq 0} \Phi_k W_k$, W may be viewed as a cylindrical Brownian motion on H.

We briefly recall the classical formalism for the infinite dimensional Wiener process as in [8,34]. Consider the collection of Hilbert–Schmidt operators mapping H into $D(A^\beta)$, $\beta \geq 0$. We denote this family by $L_2(H, D(A^\beta))$. Throughout this work we assume that σ, understood as an operator, has the form

$$\sigma \Phi_k = \lambda_k^{-\gamma} \Phi_k. \tag{2.4}$$

We will write $\sigma \mathrm{d}W(t) = \sum_{k\geq 1} \lambda_k^{-\gamma} \Phi_k \mathrm{d}W_k(t),\ t \geq 0$. One may check that, for every $\epsilon > 0$, $\sigma \in L_2(H, D(A^{\gamma-1/2-\epsilon}))$. In particular, given the standing assumption that $\gamma > 1$, we have $\sigma \in L_2(H, D(A^{1/2}))$.

2.2. The stochastic Stokes equation and limit theorems

We next consider the linear system associated with (2.1), which we write in the abstract form

$$\mathrm{d}\bar{U} + \nu A\bar{U}\mathrm{d}t = \sum_k \lambda_k^{-\gamma} \Phi_k \mathrm{d}W_k, \quad \bar{U}(0) = \bar{U}_0. \tag{2.5}$$

For the purposes here this system can be analyzed as a 2D stochastic heat equation driven by an additive cylindrical Brownian motion (for general results we refer readers to [8,35].)

Let us denote by $\bar{u}_k, k \geq 1$, the Fourier coefficients of the solution $\bar{U}$ with respect to the system $\{\Phi_k\}_k$ in H, i.e. $\bar{u}_k = (\bar{U}, \Phi_k), k \geq 1$. By (2.5), we note that each Fourier mode $\bar{u}_k$ represents a one-dimensional stable Ornstein–Uhlenbeck process with the dynamics

$$d\bar{u}_k + \nu\lambda_k \bar{u}_k dt = \lambda_k^{-\gamma} dW_k, \quad \bar{u}_k(0) = \bar{u}_{0k}, \quad k \geq 1. \tag{2.6}$$

It follows from (2.6) that

$$\bar{u}_k(t) = \bar{u}_k(0) e^{-\nu\lambda_k t} + \lambda_k^{-\gamma} \int_0^t e^{-\nu\lambda_k(t-s)} dW_k(t), \quad k \geq 1,\ t \geq 0. \tag{2.7}$$

In what follows we will use the following auxiliary results concerning asymptotics of the first moments of the Fourier modes $\bar{u}_k, k \geq 1$ (see also Theorem 2.1 in [26].)

Lemma 2.1. *Suppose that $\bar{U}$ is a solution of (2.5) and let $\bar{U}^N := P_N \bar{U},\ N \geq 1$.*

(i) *Assume that $\gamma' < \gamma$ and that $\mathbb{E}|A^{\gamma'-1/2}\bar{U}_0|^2 < \infty$. Then*

$$\bar{U} \in L^2(\Omega; L^2_{\text{loc}}([0,\infty); D(A^{\gamma'}))) \cap L^2(\Omega; C([0,\infty); D(A^{\gamma'-1/2}))). \tag{2.8}$$

(ii) *Suppose that $\bar{U}_0 = 0$; then*

$$\mathbb{E} \int_0^T \bar{u}_k^2 dt \approx \frac{T\lambda_k^{-(1+2\gamma)}}{2\nu} \approx \frac{T\lambda_1^{-(1+2\gamma)}}{2\nu} k^{-(1+2\gamma)}, \tag{2.9}$$

and

$$\text{Var}\left[\int_0^T \bar{u}_k^2 dt\right] \sim \lambda_k^{-(3+4\gamma)} \sim k^{-(3+4\gamma)}. \tag{2.10}$$

(iii) *Moreover, for $\beta > \gamma$,*

$$\mathbb{E} \int_0^T |A^\beta \bar{U}^N|^2 dt \approx \frac{T\lambda_1^{2\beta-2\gamma-1}}{2\nu(2\beta-2\gamma)} N^{2\beta-2\gamma}. \tag{2.11}$$

Proof. The first item is classical and may, for example, be justified with a Galerkin scheme or other suitable techniques from the general theory of existence and uniqueness of the solutions for stochastic parabolic equations. See e.g. [8,35]. Using (2.7), (ii) follows by direct computations of the corresponding moments, and for the final item we deduce

$$\begin{aligned}\mathbb{E} \int_0^T |A^\beta \bar{U}^N|^2 dt &= \mathbb{E} \int_0^T \left| \sum_{k=1}^N \lambda_k^\beta \bar{u}_k \Phi_k \right|^2 dt = \sum_{k=1}^N \lambda_k^{2\beta} \mathbb{E} \int_0^T \bar{u}_k^2 dt \\ &\approx \frac{T}{2\nu} \sum_{k=1}^N \lambda_k^{2\beta-1-2\gamma} \approx \frac{T\lambda_1^{2\beta-2\gamma-1}}{2\gamma} \frac{N^{2\beta-2\gamma}}{2\beta-2\gamma},\end{aligned}$$

where we have made use of (ii), (2.2) in conjunction with

$$\sum_{k=1}^N k^a \approx \frac{N^{1+a}}{a+1}, \quad a > -1. \tag{2.12}$$

The proof is complete. $\square$

We finally recall some particular versions of the law of large numbers (LLN) and the central limit theorem (CLT) which are used to prove consistency and asymptotic normality of the class of estimators given by (1.5).

Lemma 2.2 (*The Law of Large Numbers*)**.** *Let ξ_n, $n \geq 1$, be a sequence of random variables and b_n, $n \geq 1$, be an increasing sequence of positive numbers such that $\lim_{n\to\infty} b_n = +\infty$, and*

$$\sum_{n=1}^{\infty} \frac{\operatorname{Var} \xi_n}{b_n^2} < \infty. \tag{2.13}$$

(i) *If we assume that the random variables ξ_n, $n \geq 1$, are independent then*

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} (\xi_k - \mathbb{E}\xi_k)}{b_n} = 0 \quad a.s.$$

(ii) *If we suppose only that ξ_n, $n \geq 1$, are merely uncorrelated random variables, then*

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} (\xi_k - \mathbb{E}\xi_k)}{b_n} = 0, \tag{2.14}$$

in probability.

Proof. See, for example, Shiryaev [36, Theorem IV.3.2] for the proof of (i). The second item, (ii), similar to the proof of the weak LLN, follows from the Markov inequality. For a fixed $\epsilon > 0$, and for all pairs $m < n$, we have

$$\begin{aligned}
\mathbb{P}\left(\frac{\sum_{k=1}^{n} (\xi_k - \mathbb{E}\xi_k)}{b_n} > \epsilon\right) &\leq \frac{1}{\epsilon^2 b_n^2} \mathbb{E}\left(\sum_{k=1}^{n} (\xi_k - \mathbb{E}\xi_k)\right)^2 \leq \frac{1}{\epsilon^2 b_n^2} \sum_{k=1}^{n} \operatorname{Var} \xi_k \\
&\leq \frac{1}{\epsilon^2 b_n^2} \sum_{k=1}^{m} \operatorname{Var} \xi_k + \frac{1}{\epsilon^2} \sum_{k=m}^{n} \frac{\operatorname{Var} \xi_k}{b_k^2} \\
&\leq \frac{1}{\epsilon^2 b_n^2} \sum_{k=1}^{m} \operatorname{Var} \xi_k + \frac{1}{\epsilon^2} \sum_{k=m}^{\infty} \frac{\operatorname{Var} \xi_k}{b_k^2}.
\end{aligned}$$

Since $b_n \to \infty$, (2.14) follows. □

The following central limit theorem is a special case of a more general result for martingales; see, for instance, [24, Theorem 5.5.4(II)].

Lemma 2.3 (*CLT for Stochastic Integrals*)**.** *Let $\mathcal{S} = (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\geq 0}, \{W_k\}_{k\geq 1})$ be a stochastic basis. Suppose that $\sigma_k \in L^2(\Omega; L^2([0,T]))$ is a sequence of real valued predictable processes such that*

$$\lim_{N\to\infty} \frac{\sum_{k=1}^{N} \int_0^T \sigma_k^2 \mathrm{d}t}{\sum_{k=1}^{N} \mathbb{E} \int_0^T \sigma_k^2 \mathrm{d}t} = 1 \quad \textit{in Probability.}$$

Then

$$\frac{\sum_{k=1}^{N}\int_0^T \sigma_k \mathrm{d}W_k}{\left(\sum_{k=1}^{N}\mathbb{E}\int_0^T \sigma_k^2 \mathrm{d}t\right)^{1/2}}$$

converges in distribution to a standard normal random variable as $N \to \infty$.

2.3. The nonlinear term

The nonlinear term appearing in (2.1) is given by $B(U, U^\sharp) := P_H((U\cdot\nabla)U^\sharp) = P_H(\sum_{j=1}^2 U_j\partial_j U^\sharp)$, which is defined for $U \in V$ and $U^\sharp \in D(A)$. Note that, for brevity of notation, we will often write $B(U)$ for $B(U, U)$ as for example in (2.1). We have the following properties of B:

Lemma 2.4. (i) *B is bilinear and continuous from $V \times V$ into V' and from $V \times D(A)$ into H. For $U, U^\sharp \in V$, B satisfies the cancelation property*

$$\langle B(U, U^\sharp), U^\sharp\rangle = 0. \tag{2.15}$$

If $U, U^\sharp, U^\flat$ are elements in V, then

$$\left|\langle B(U, U^\sharp), U^\flat\rangle\right| \le c|U|^{1/2}\|U\|^{1/2}\|U^\sharp\||U^\flat|^{1/2}\|U^\flat\|^{1/2}. \tag{2.16}$$

On the other hand if $U \in V$, $U^\sharp \in D(A)$, and $U^\flat \in H$, then we have

$$\left|(B(U, U^\sharp), U^\flat)\right| \le c\begin{cases}|U|^{1/2}\|U\|^{1/2}\|U^\sharp\|^{1/2}|AU^\sharp|^{1/2}|U^\flat|.\\ |U|^{1/2}|AU|^{1/2}\|U^\sharp\||U^\flat|.\end{cases} \tag{2.17}$$

(ii) *In the case of either periodic, (1.2), or Dirichlet, (1.3), boundary conditions, $B(U) \in D(A^\beta)$ for every $0 < \beta < 1/4$ and every $U \in D(A)$. Moreover, for such values of β,*

$$|A^\beta B(U)|^2 \le c\|U\|^2|AU|^2. \tag{2.18}$$

(iii) *In the case of periodic boundary conditions (1.2), $B(U, U^\sharp) \in D(A^\beta)$ whenever $\beta > 1/2$, $U \in D(A^\beta)$, $U^\sharp \in D(A^{\beta+1/2})$, and for such U, $U^\sharp$,*

$$|A^\beta B(U, U^\sharp)|^2 \le c|A^\beta U|^2|A^{\beta+1/2}U|^2. \tag{2.19}$$

Proof. The properties outlined in (i) and (iii) are classical; see, for instance, [38], or [6, Lemma 10.4] for (2.19).

The properties in (ii) are established via interpolation and the equivalence of certain fractional order spaces; see [13]. Since [13] emphasized the case of spatial dimension 3, for the sake completeness, we briefly recall the arguments.

For any element $U \in D(A)$, standard estimates imply that

$$|B(U)|^2 \le c\|U\|^3|AU|,$$
$$\|B(U)\|^2_{H^1(\mathcal{M})^2} \le c\|U\||AU|^3.$$

Let $\tilde{V} = H \cap H^1(\mathcal{D})^2$ and for $s \in (0, 1)$ we define the interpolation spaces $\tilde{V}_s = [\tilde{V}, H]_{1-s}$. See [23] for the general theory. In [13], it is established that $D(A^\beta) = \tilde{V}_{2\beta}$, $\beta < 1/4$ in the Dirichlet case (1.3).[3] Note that $\tilde{V}$ does not incorporate boundary conditions and so $B(U) \in \tilde{V}$,

[3] In the periodic case, (1.2), $\tilde{V} = V$, so $D(A^\beta) = \tilde{V}_{2\beta}$, as a direct consequence of the fact that $D(A^{1/2}) = V$.

for $U \in D(A)$. In consequence, for any such $U \in D(A)$ and allowed values of $0 < \beta < 1/4$ we have, by interpolation,

$$\begin{aligned} |A^{\beta} B(U)|^2 &= |B(U)|^2_{\tilde{V}^{2\beta}} \leq (|B(U)|^{1-2\beta} \|B(U)\|^{2\beta}_{H^1})^2 \\ &\leq c(\|U\|^3 |AU|)^{1-2\beta} (\|U\| |AU|^3)^{2\beta} \\ &\leq c(\|U\|^3 |AU|)^{1/2} (\|U\| |AU|^3)^{1/2} \leq c\|U\|^2 |AU|^2. \end{aligned}$$

Combining these observations gives (ii), completing the proof. □

Remark 2.5. When we consider the case (1.3) it is not true in general that $B(U) \in D(A^{\beta})$, even for $U \in D(A^{\beta+1/2})$, $\beta \geq 1/4$. This is due to the fact that while the Leray projector P_H is continuous on $H^m(\mathcal{D})$, $m \geq 1$, we do not expect P_H to map $H^m_0(\mathcal{D})$ into $H^m_0(\mathcal{D})$. See [38] and also [13]. For this reason we may not expect an inequality like (2.19) for such Dirichlet boundary conditions. As a result, (2.18) relies on a delicate analysis of small fractional order space where the boundary is not present; see [13,23].

2.4. Existence, uniqueness and higher regularity

With these mathematical formalities in place we now define precisely (2.1), in the usual time integrated sense, and recall some now well established existence, uniqueness and regularity results for these equations. Note that for this work the solutions that we consider correspond to so called ‘strong solutions’ in the deterministic setting (see [38]). In the context of stochastic analysis, since we may suppose that the stochastic basis $\mathcal{S}$ is fixed in advance, we may say that the solutions considered are ‘strong’ (or less confusingly ‘pathwise’) in the probabilistic sense as well.

Theorem 2.6. (i) *Suppose that we impose either (1.2) or (1.3) and assume that $U_0 \in V$, $\sigma \in L_2(H, V)$. Then there exists a unique, H-valued, $\mathcal{F}_t$-adapted process U with*

$$U \in L^2_{\text{loc}}([0, \infty); D(A)) \cap C([0, \infty); V) \quad a.s. \tag{2.20}$$

and so for each $t \geq 0$,

$$U(t) + \int_0^t (\nu AU + B(U))\mathrm{d}t' = U_0 + \sum_k \sigma \Phi_k W^k(t),$$

with the equality understood in H.

(ii) *In the case of periodic boundary conditions (1.2), if $\beta > 1/2$ and so $\sigma \in L_2(H, D(A^{\beta}))$, $U_0 \in D(A^{\beta})$, then*

$$U \in L^2_{\text{loc}}([0, \infty), D(A^{\beta+1/2})) \cap C([0, \infty), D(A^{\beta})). \tag{2.21}$$

Remark 2.7. (i) As noted above, when σ is defined via (2.4), $\sigma \in L_2(H, V)$ whenever $\gamma > 1$. Indeed we have $\sigma \in L_2(H, D(A^{\beta}))$ for every $\beta < \gamma - 1/2$.

(ii) We suspect that higher regularity similar to that of Theorem 2.6, (ii), may be established in the case of Dirichlet boundary conditions, (1.3). However since (2.19) does not apply (see Remark 2.5) a proof different to that outlined here is needed.

Proof. The well-posedness of (2.1) has been studied by many authors, as discussed in the introduction. Since we are considering the case of an additive noise, the proof is close to the deterministic case after we perform a suitable change of variables. For completeness, we briefly recall some of the formal arguments and note that the computations may be rigorously justified with a suitable Galerkin scheme. Consider first the linear system (2.5) with initial condition $\bar{U}(0) = U_0$. As in Lemma 2.1 above, we have that $\bar{U}$ in $L^2_{\text{loc}}([0,\infty); D(A)) \cap C([0,\infty); V)$ (or in $L^2_{\text{loc}}([0,\infty), D(A^{\beta+1/2})) \cap C([0,\infty), D(A^\beta))$, under the conditions of item (ii)). We now consider the shifted variable $\tilde{U} = U - \bar{U}$, which satisfies

$$\frac{\mathrm{d}\tilde{U}}{\mathrm{d}t} + \nu A\tilde{U} + B(\tilde{U} + \bar{U}) = 0 \quad \tilde{U}(0) = 0. \tag{2.22}$$

The estimates that lead to (2.21) are standard. We first multiply (2.22) by U, integrate over the domain and use (2.15), (2.16) and (2.8) to infer that $\tilde{U} \in L^2_{\text{loc}}([0,\infty); V) \cap L^\infty_{\text{loc}}([0,\infty); H)$. With this regularity in hand we next multiply (2.22) by $A\tilde{U}$ and apply (2.17) and (2.8) in order to conclude (2.20).

For $\beta > 1/2$ we multiply (2.22) by $A^{2\beta}\tilde{U}$ and infer

$$\frac{\mathrm{d}|A^\beta \tilde{U}|^2}{\mathrm{d}t} + 2\nu |A^{\beta+1/2}\tilde{U}|^2 - 2\langle A^\beta B(\tilde{U} + \bar{U}), A^\beta U\rangle = 0. \tag{2.23}$$

Since $\beta > 1/2$, we may apply (2.19) and estimate

$$\begin{aligned}\frac{\mathrm{d}|A^\beta \tilde{U}|^2}{\mathrm{d}t} + 2\nu |A^{\beta+1/2}\tilde{U}|^2 &\le c|A^\beta(\tilde{U} + \bar{U})||A^{\beta+1/2}(\tilde{U} + \bar{U})||A^\beta \tilde{U}| \\ &\le c(|A^\beta \tilde{U}|^2 + |A^\beta \bar{U}|^2)|A^\beta \tilde{U}|^2 + \nu|A^{\beta+1/2}\tilde{U}|^2 + \nu|A^{\beta+1/2}\bar{U}|^2.\end{aligned}$$

Rearranging,

$$\frac{\mathrm{d}|A^\beta \tilde{U}|^2}{\mathrm{d}t} + 2\nu |A^{\beta+1/2}\tilde{U}|^2 \le c(|A^\beta \tilde{U}|^2 + |A^\beta \bar{U}|^2)|A^\beta \tilde{U}|^2 + \nu|A^{\beta+1/2}\bar{U}|^2.$$

Observe that, due to the Gronwall lemma, if $\tilde{U} \in L^2_{\text{loc}}([0,\infty); D(A^\beta))$ then we infer that $\tilde{U} \in L^2_{\text{loc}}([0,\infty); D(A^{\beta+1/2})) \cap L^\infty_{\text{loc}}([0,\infty); D(A^\beta))$. The desired result therefore follows from an inductive argument on β starting with the base case assumption $\beta \in [1/2, 1)$ which is satisfied as a consequence of (2.20). □

3. Estimators for ν: heuristic derivation and the main results

In this section we sketch the heuristic derivations of the estimators based on a particular version of the Girsanov theorem. We then restate, now in precise terms, the main results of this paper.

As before, we denote by U^N the projection of the solution U of the original equation (2.1) onto $H_N = P_N H \cong \mathbb{R}^N$. Note that U^N satisfies the following finite dimensional system:

$$\mathrm{d}U^N = -(\nu A U^N + \psi^N)\mathrm{d}t + P_N \sigma \mathrm{d}W, \quad U^N(0) = U_0^N, \tag{3.1}$$

where $\psi^N(t) := P_N(B(U))$. To obtain an initial guess of the form of the estimator for the parameter ν, we treat ψ^N as an external known quantity, independent of ν, and view (3.1) as a stochastic equation evolving in $\mathbb{R}^N$. Let us denote by $\mathbb{P}^{N,T}_\nu$ the probability measure in $C([0,T]; \mathbb{R}^N)$ generated by U^N. Formally, we compute the Radon–Nikodym derivative or likelihood ratio

$\mathrm{d}\mathbb{P}_{\nu}^{N,T}/\mathrm{d}\mathbb{P}_{\nu_0}^{N,T}$ (see e.g. [25, Section 7.6.4])

$$\frac{\mathrm{d}\mathbb{P}_{\nu}^{N,T}(U^N)}{\mathrm{d}\mathbb{P}_{\nu_0}^{N,T}} = \exp\Bigg(-\int_0^T (\nu-\nu_0)(AU^N)'G^2\mathrm{d}U^N(t) - \frac{1}{2}\int_0^T (\nu^2-\nu_0^2)(AU^N)'G^2AU^N\mathrm{d}t - \int_0^T (\nu-\nu_0)(AU^N)'G^2\psi^N\mathrm{d}t\Bigg),$$

where $G := (P_N\sigma)^{-1} = \mathrm{diag}[\sigma_1^{-1},\dots,\sigma_N^{-1}] = \mathrm{diag}[\lambda_1^{\gamma},\dots,\lambda_N^{\gamma}]$ and v' denotes the transpose of the vector $v\in\mathbb{R}^N$. By maximizing the likelihood ratio with respect to the parameter of interest ν, we may compute the (formal) maximum likelihood estimator (MLE) ν_N of the parameter ν. A direct computation yields

$$\nu_N = -\frac{\int_0^T (AU^N)'G^2\mathrm{d}U^N + \int_0^T (AU^N)'G^2P_N(B(U))\mathrm{d}t}{\int_0^T (AU^N)'G^2AU^N\mathrm{d}t}. \tag{3.2}$$

As expected, ν_N is a valid estimator and in fact one can show that it is a consistent estimator of the true parameter ν. This consistency makes essential use of the fact that the denominator $\int_0^T (AU^N)'G^2AU^N\mathrm{d}t$ diverges to infinity as $N\uparrow\infty$. See Lemma 4.3 below. With this in mind, we introduce a slight modification to the MLE (3.2), and propose the following class of estimators:

$$\widetilde{\nu}_N = -\frac{\int_0^T \langle A^{1+2\alpha}U^N, \mathrm{d}U^N\rangle + \int_0^T \langle A^{1+2\alpha}U^N, P_NB(U)\rangle\mathrm{d}t}{\int_0^T |A^{1+\alpha}U^N|^2\mathrm{d}t}, \tag{3.3}$$

where α is a free parameter with a range specified later on. Note that this formulation appears in the functional language developed above and is derived using that the action of G^2 on H_N is equivalent to $A^{2\gamma}$. Also we observe that ν_N is a particular case of $\widetilde{\nu}_N$ with $\alpha=\gamma$.

While the estimator $\widetilde{\nu}_N$ has desirable theoretical properties, it also assumes that $P_N(B(U))$ is computable, which could be quite a difficult task. Since our goal is to provide estimators that can be eventually implemented in practice (evaluated numerically), we propose two further classes of estimators. One class is naturally derived from (3.3) by approximating $P_N(B(U))$ with $P_N(B(U^N))$:

$$\check{\nu}_N = -\frac{\int_0^T \langle A^{1+2\alpha}U^N, \mathrm{d}U^N\rangle + \int_0^T \langle A^{1+2\alpha}U^N, P_NB(U^N)\rangle\mathrm{d}t}{\int_0^T |A^{1+\alpha}U^N|^2\mathrm{d}t}. \tag{3.4}$$

Note that $\check{\nu}_N$ now depends only on the first N Fourier modes. However, even in this case the expression for $P_NB(U^N)$ is very complicated due to the nontrivial coupling of the modes. See e.g. [10]. It turns out, as shown rigorously below (see Proposition 4.6), that the second term appearing in (3.3),

$$\kappa_N := -\frac{\int_0^T \langle A^{1+2\alpha}U^N, P_NB(U)\rangle\mathrm{d}t}{\int_0^T |A^{1+\alpha}U^N|^2\mathrm{d}t}, \tag{3.5}$$

is of lower order and tends to zero, as $N\to\infty$. Hence we get the following consistent estimators of the parameter ν:

$$\hat{\nu}_N = -\frac{\int_0^T \langle A^{1+2\alpha}U^N, \mathrm{d}U^N\rangle}{\int_0^T |A^{1+\alpha}U^N|^2\mathrm{d}t} = -\frac{\sum_{k=1}^N \lambda_k^{1+2\alpha}\int_0^T u_k \mathrm{d}u_k}{\sum_{k=1}^N \lambda_k^{2+2\alpha}\int_0^T u_k^2\mathrm{d}t}$$

$$= -\frac{\sum_{k=1}^N \lambda_k^{1+2\alpha}(u_k^2(T) - u_k^2(0) - T\lambda_k^{-2\gamma})}{2\sum_{k=1}^N \lambda_k^{2+2\alpha}\int_0^T u_k^2\mathrm{d}t}. \tag{3.6}$$

Clearly this last estimator is easiest to compute numerically. On the other hand it may lack the speed of convergence of the first two.

We conclude this section with the main result of this paper:

Theorem 3.1. *Suppose that U solves (2.1) with either (1.2) or (1.3) in the sense of and under the conditions imposed by Theorem 2.6. Assume that $\gamma > 1$ and in the case (1.3), additionally that $\gamma < 1 + 1/4$. Also, assume that $U_0 \in D(A^\beta)$, for some $\beta > \gamma - 1/2$.*

(i) *If $\alpha > \gamma - 1$, then $\widetilde{\nu}_N$, $\check{\nu}_N$ and $\hat{\nu}_N$ as given by (3.3), (3.4) and (3.6) are weakly consistent estimators of the parameter ν, i.e.*

$$\lim_{N\to\infty} \widetilde{\nu}_N = \lim_{N\to\infty} \check{\nu}_N = \lim_{N\to\infty} \hat{\nu}_N = \nu$$

in probability.

(ii) *If $\alpha > \gamma - 1/2$, then $\tilde{\nu}_N$ is asymptotically normal with rate N, i.e.*

$$N(\tilde{\nu}_N - \nu) \xrightarrow{d} \eta, \tag{3.7}$$

where η is Gaussian random variable with mean zero and variance $\frac{2\nu(\alpha-\gamma+1)^2}{\lambda_1 T(\alpha-\gamma+1/2)}$.

4. Proof of the main theorem

We establish the proof of Theorem 1.1 in a series of propositions. As mentioned in the introduction, we do not have precise spectral information about Fourier coefficients $u_k = (U, \Phi_k)$, $k \geq 1$, in contrast to the linear case (see Section 2.2). To overcome this, we proceed by decomposing the solution into a linear and a nonlinear part, $U = \bar{U} + R$. We assume that $\bar{U}$ is the solution of the linear stochastic Stokes equation (2.5) with $\bar{U}(0) = 0$. The residual R must therefore satisfy

$$\partial_t R + \nu AR = -B(U), \quad R(0) = R_0. \tag{4.1}$$

First, we study the regularity properties of R and show that R is slightly smoother than $\bar{U}$. Subsequently, we make crucial use of this extra regularity and establish the consistency of the proposed estimators by showing that the second term in (3.3) converges to zero. The final section treats the asymptotic normality using the CLT introduced in Section 2.2.

Remark 4.1. For simplicity and clarity of presentation we shall assume a more regular initial condition $U_0 \in D(A^\gamma)$ in comparison to the statement of Theorem 3.1. The more general case when we assume merely that $U_0 \in D(A^\beta)$ for some $\beta > \gamma - 1/2$ may be treated by writing $U = \bar{U} + R + S$, where $\bar{U}$ satisfies (2.5) with $\bar{U}_0 = 0$, R satisfies (4.1), this time with $R_0 = 0$, and finally S is the solution of $\partial_t S + \nu AS = 0$, with $S(0) = U_0$.

4.1. Regularity properties for the residual

Proposition 4.2. *Suppose that* $\beta \geq 0$, $\gamma > 1$ *and that* R *solves* (4.1) *with* U *the solution of* (2.1) *corresponding to an initial condition* $U_0 \in D(A^{1/2+\beta})$.

(i) *If* U *and* R *satisfy Dirichlet boundary conditions* (1.3), *and* $\beta < 1/4$, *then for every* $T > 0$ *we have*

$$\sup_{t\in[0,T]} |A^{1/2+\beta} R|^2 + \int_0^T |A^{1+\beta} R|^2 < \infty. \tag{4.2}$$

Moreover, for an increasing sequence of stopping times τ_n *with* $\tau_n \uparrow \infty$,

$$\mathbb{E}\left(\sup_{t\in[0,\tau_n]} |A^{1/2+\beta} R|^2 + \int_0^{\tau_n} |A^{1+\beta} R|^2\right) < \infty. \tag{4.3}$$

(ii) *In the case where both* U *and* R *satisfy periodic boundary conditions* (1.2) *and we assume* $\beta < \gamma - 1/2$, *the same conclusions hold.*

Proof. As above in Theorem 2.6 the computations given here may be rigorously justified via Galerkin approximations. Multiplying (4.1) by $A^{1+2\beta} R$, integrating and using the symmetry of the powers of A, we infer

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|A^{\beta+1/2} R|^2 + \nu|A^{\beta+1} R|^2 = -\langle A^\beta B(U), A^{\beta+1} R\rangle. \tag{4.4}$$

For the case of a bounded domain, (1.3), we infer from (2.18) and Theorem 2.6, (2.20) that

$$\int_0^T |A^\beta B(U)|^2 \mathrm{d}t \leq c \int_0^T \|U\|^2 |AU|^2 \mathrm{d}t < \infty \quad \text{a.s.}$$

By integrating (4.4) in time and making standard estimates with Young's inequality, (4.2) now follows in this case.

In the case of the periodic domain we estimate $|A^\beta B(U)|$ differently. Define $\beta' = \max\{\beta, \gamma/2\}$ such that $1/2 < \beta' < \gamma - 1/2$. By applying the higher regularity estimates (2.19) we find that

$$\begin{aligned}
|\langle A^\beta B(U), A^{\beta+1} R\rangle| &\leq |A^\beta B(U)||A^{\beta+1} R| \\
&\leq c|A^{\beta'} B(U)||A^{\beta+1} R| \\
&\leq c|A^{\beta'} U||A^{\beta'+1/2} U||A^{\beta+1} R| \\
&\leq c|A^{\beta'} U|^2 |A^{\beta'+1/2} U|^2 + \frac{\nu}{2}|A^{\beta+1} R|^2.
\end{aligned}$$

Due to Theorem 2.6, (ii), we have, for any $T > 0$, that

$$\int_0^T |A^{\beta'} U|^2 |A^{\beta'+1/2} U|^2 \mathrm{d}t < \infty \quad \text{a.s.},$$

and (4.2) follows once again.

For the stopping times τ_n, we define

$$\tau_n := \inf_{t\geq 0}\left\{\sup_{t'\leq t} \|U\|^2 + \int_0^t |AU|^2 \mathrm{d}t' > n\right\}$$

when (1.3) is assumed and

$$\tau_n := \inf_{t\geq 0}\left\{\sup_{t'\leq t}|A^{\beta'}U|^2 + \int_0^t |A^{\beta'+1/2}U|^2 \mathrm{d}t' > n\right\}$$

for (1.2). In either case it is clear that $\{\tau_n\}_{n\geq 1}$ is increasing. Moreover, in the case (1.3), since $\mathbb{P}(\tau_n < T) = \mathbb{P}(\sup_{t'\leq T}\|U\|^2 + \int_0^T |AU|^2\mathrm{d}t' \geq n)$, it follows from (2.20) and the fact that τ_n is increasing that $\lim_{n\to}\tau_n = \infty$ a.s. Arguing in the same manner for the case (1.2), the proof is complete. □

Remark. Comparing Proposition 4.2, (4.2) with Lemma 2.1 (2.8) and (2.11) we see that R has been shown to be just shy of a derivative more regular than $\bar{U}$. More precisely we have that, for any $\epsilon > 0$,

$$\bar{U} \in L^2(\Omega; L^2_{\mathrm{loc}}([0,\infty); D(A^{\gamma-\epsilon}))), \qquad \bar{U} \notin L^2(\Omega; L^2_{\mathrm{loc}}([0,\infty); D(A^{\gamma+\epsilon}))),$$

while on the other hand,

$$R \in L^2(\Omega; L^2_{\mathrm{loc}}([0,\infty); D(A^{\gamma+1/2-\epsilon}))).$$

As an immediate application of these properties of the residual R we have the following result:

Lemma 4.3. *Suppose that U and $\bar{U}$ are the solutions of (2.1) and (2.5) respectively. For both (1.2) and (1.3) we suppose that $\gamma > 1$, $U(0) = U_0 \in D(A^\gamma)$*[4] *and $\bar{U}_0 = 0$. Additionally, in the case (1.3), we assume that $\gamma < 1 + 1/4$. Then, for any $\alpha > \gamma - 1$,*

$$\lim_{N\to\infty}\frac{\int_0^T |A^{1+\alpha}U^N|^2\mathrm{d}t}{\mathbb{E}\int_0^T |A^{1+\alpha}\bar{U}^N|^2\mathrm{d}t} = 1 \tag{4.5}$$

with probability 1.

Proof. Note that

$$|A^{1+\alpha}U^N|^2 \leq |A^{1+\alpha}\bar{U}^N|^2 + |A^{1+\alpha}R^N|^2 + 2|A^{1+\alpha}\bar{U}^N||A^{1+\alpha}R^N|,$$
$$|A^{1+\alpha}U^N|^2 \geq |A^{1+\alpha}\bar{U}^N|^2 + |A^{1+\alpha}R^N|^2 - 2|A^{1+\alpha}\bar{U}^N||A^{1+\alpha}R^N|,$$

and therefore (4.5) follows once we have shown that

$$I_1^N := \frac{\int_0^T |A^{1+\alpha}\bar{U}^N|^2}{\mathbb{E}\int_0^T |A^{1+\alpha}\bar{U}^N|^2} \to 1 \quad \text{a.s.}$$

and that

$$I_2^N := \frac{\int_0^T |A^{1+\alpha}R^N|^2\mathrm{d}t}{\mathbb{E}\int_0^T |A^{1+\alpha}\bar{U}^N|^2} \to 0 \quad \text{a.s.} \tag{4.6}$$

For the first item, I_1^N, we apply the law of large numbers (LLN), Lemma 2.2, with $\xi_n := \lambda_n^{2\alpha+2}\int_0^T \bar{u}_n^2(t)\mathrm{d}t$ and $b_n := \sum_{k=1}^n \mathbb{E}[\xi_k]$. Notice that, due to (2.2) and (2.9),

$$b_n \sim \sum_{k=1}^n \lambda_k^{2\alpha+2}\lambda_k^{-1-2\gamma} \sim \sum_{k=1}^n k^{2\alpha-2\gamma+1}. \tag{4.7}$$

[4] At the cost of further evaluations, this condition may be weakened to the conditions imposed in Theorem 3.1. This applies both here and below for Propositions 4.4 and 4.6. See Remark 4.1.

Given the assumptions $\alpha > \gamma - 1$, we have that $\lim_{n\to\infty} b_n = \infty$. Moreover, combining (4.7) with (2.2), (2.10) and (2.12),

$$\sum_{n\geq 1} \frac{\operatorname{Var}\xi_n}{b_n^2} \sim \sum_{n\geq 1} \frac{\lambda_n^{4\alpha+4-3-4\gamma}}{\left(\sum_{k=1}^{n} \lambda_k^{2\alpha-2\gamma+1}\right)^2} \sim \sum_{n\geq 1} \frac{\lambda_n^{4\alpha-4\gamma+1}}{(\lambda_n^{2\alpha-2\gamma+2})^2} \sim \sum_{n\geq 1} \frac{1}{n^3} < \infty.$$

Thus, by the LLN we conclude that $\lim_{N\to\infty} I_1^N = 1$ with probability 1.

Since $1+\alpha > \gamma$, by (2.11) we infer

$$\mathbb{E} \int_0^T |A^{1+\alpha} \bar{U}^N| \mathrm{d}t \sim N^{2\alpha-2\gamma+2}.$$

Pick any $\alpha' \in (\gamma - 1, \min\{\alpha, 1/4\})$, in the case (1.3), or any $\alpha' \in (\gamma - 1, \min\{\alpha, \gamma - 1/2\})$ under the assumption (1.2). By applying (4.2) for R established in Proposition 4.2, we have in both cases that

$$\int_0^T |A^{1+\alpha'} R|^2 \mathrm{d}t < \infty \quad \text{a.s.}$$

Combining these observations and making use of (2.3), we have

$$I_2^N \leq c \frac{\int_0^T |A^{1+\alpha} R^N|^2 \mathrm{d}t}{N^{2\alpha-2\gamma+2}} \leq c \frac{\lambda_N^{2(\alpha-\alpha')} \int_0^T |A^{1+\alpha'} R^N|^2 \mathrm{d}t}{N^{2\alpha-2\gamma+2}} \leq c \frac{\int_0^T |A^{1+\alpha'} R|^2 \mathrm{d}t}{N^{2\alpha'-2\gamma+2}}.$$

Due to the restrictions on the choice of α', we have that $2\alpha' - 2\gamma + 2 > 0$, and hence $I_2^N \to 0$, as $N \to \infty$, with probability 1. The proof is complete. □

4.2. *Consistency of the estimators*

Using the dynamics of U^N, i.e. substituting (3.1) into (3.3), we get the following representation for the estimator $\widetilde{\nu}_N$:

$$\begin{aligned} \widetilde{\nu}_N &= \nu - \frac{\int_0^T \langle A^{1+2\alpha} U^N, P_N \sigma \mathrm{d}W \rangle}{\int_0^T |A^{1+\alpha} U^N|^2 \mathrm{d}t} \\ &= \nu - \frac{\int_0^T \left\langle A^{1+2\alpha-\gamma} U^N, \sum_{k=1}^{N} \Phi_k \mathrm{d}W_k \right\rangle}{\int_0^T |A^{1+\alpha} U^N|^2 \mathrm{d}t}. \end{aligned} \tag{4.8}$$

Similarly, we deduce

$$\begin{aligned} \check{\nu}_N &= \nu - \frac{\int_0^T \left\langle A^{1+2\alpha-\gamma} U^N, \sum_{k=1}^{N} \Phi_k \mathrm{d}W_k \right\rangle}{\int_0^T |A^{1+\alpha} U^N|^2 \mathrm{d}t} \\ &\quad + \frac{\int_0^T \langle A^{1+2\alpha} U^N, P_N B(U) - P_N B(U^N) \rangle \mathrm{d}t}{\int_0^T |A^{1+\alpha} U^N|^2 \mathrm{d}t}. \end{aligned} \tag{4.9}$$

Note that $\hat{\nu}_N = \widetilde{\nu}_N - \kappa_N$, with κ_N defined by (3.5). Thus,

$$\widehat{\nu}_N = \nu - \kappa_N - \frac{\int_0^T \left\langle A^{1+2\alpha-\gamma} U^N, \sum_{k=1}^N \Phi_k \mathrm{d}W_k \right\rangle}{\int_0^T |A^{1+\alpha} U^N|^2 \mathrm{d}t}. \tag{4.10}$$

With the above representations for the estimators, the consistency will follow if we show that each stochastic term on the right hand side of (4.8)–(4.10) converges to zero.

Proposition 4.4. *Assume the conditions and notation from Lemma 4.3. Then:*

(i) *For every $\delta_1 < \min\{2 + 2\alpha - 2\gamma, 1\}$,*

$$\lim_{N\to\infty} N^{\delta_1} \frac{\int_0^T \left\langle A^{1+2\alpha-\gamma} \bar{U}^N, \sum_{k=1}^N \Phi_k \mathrm{d}W_k \right\rangle}{\int_0^T |A^{1+\alpha} U^N|^2 \mathrm{d}t} = 0 \quad a.s. \tag{4.11}$$

(ii) *Whenever $\delta_2 < \min\{2 + 2\alpha - 2\gamma, 3/2\}$ in the case (1.2), or whenever $\delta_2 < \min\{2 + 2\alpha - 2\gamma, 5/4 + 1 - \gamma\}$ in the case (1.3), we have*

$$\lim_{N\to\infty} N^{\delta_2} \frac{\int_0^T \left\langle A^{1+2\alpha-\gamma} R^N, \sum_{k=1}^N \Phi_k \mathrm{d}W_k \right\rangle}{\int_0^T |A^{1+\alpha} U^N|^2 \mathrm{d}t} = 0 \tag{4.12}$$

in probability.

Proof. Due to Lemma 4.3, (4.5) and (2.11) the desired result follows once we show that each of the sequences

$$J_N^1 := \frac{\int_0^T \left\langle A^{1+2\alpha-\gamma} \bar{U}^N, \sum_{k=1}^N \Phi_k \mathrm{d}W_k \right\rangle}{\lambda_N^{2+2\alpha-2\gamma-\delta_1}} = \frac{\sum_{k=1}^N \lambda_k^{1+2\alpha-\gamma} \int_0^T \bar{u}_k \mathrm{d}W_k}{\lambda_N^{2+2\alpha-2\gamma-\delta_1}}$$

and

$$J_N^2 := \frac{\int_0^T \left\langle A^{1+2\alpha-\gamma} R^N, \sum_{k=1}^N \Phi_k \mathrm{d}W_k \right\rangle}{\lambda_N^{2+2\alpha-2\gamma-\delta_2}}$$

converges to zero as $N \to \infty$.

For the first term, J_N^1, define $\bar{\xi}_k := \lambda_k^{1+2\alpha-\gamma} \int_0^T \bar{u}_k \mathrm{d}W_k$ and $b_n := \lambda_n^{2+2\alpha-2\gamma-\delta_1}$. Under the given conditions, $\lim_{n\to\infty} b_n = \infty$. With the Itō isometry and (2.9), we have

$$\mathrm{Var}\,[\bar{\xi}_k] = \mathbb{E}[\bar{\xi}_k^2] \sim \lambda_k^{2+4\alpha-2\gamma} \lambda_k^{-(1+2\gamma)} = \lambda_k^{1+4\alpha-4\gamma}.$$

Thus,

$$\sum_{n\geq 1} \frac{\mathrm{Var}\,\xi_n}{b_n^2} \sim \sum_{n\geq 1} \frac{\lambda_n^{1+4\alpha-4\gamma}}{\lambda_n^{4+4\alpha-4\gamma-2\delta_1}} = \sum_{n\geq 1} \frac{1}{\lambda_n^{3-2\delta_1}} \sim \sum_{n\geq 1} \frac{1}{n^{3-2\delta_1}} < \infty.$$

Note that under the given conditions, $\delta_1 < 1$. This justifies the assertion that the final sum is finite. We conclude, by the LLN, Lemma 2.2, that $\lim_{N\to\infty} J_N^1 = 0$.

We turn to J_N^2. Let $r_k := (R, \Phi_k)$, $k \geq 1$, and for any stopping time τ we define

$$\zeta_k^\tau := \lambda_k^{1+2\alpha-\gamma} \int_0^\tau r_k \mathrm{d}W_k.$$

Note that the random variables ζ_k^τ, $k \geq 1$, are uncorrelated. Like in the above arguments, we let $b_n := \lambda_n^{2+2\alpha-2\gamma-\delta_2}$ and observe that this sequence is increasing and unbounded. Up to any stopping time τ such that $\operatorname{Var} \zeta_k^\tau < \infty$, we have

$$\begin{aligned}
\sum_{k\geq 1} \frac{\operatorname{Var}[\zeta_k^\tau]}{b_k^2} &= \sum_{k\geq 1} \frac{\lambda_k^{2+4\alpha-2\gamma}}{\lambda_k^{4+4\alpha-4\gamma-2\delta_2}} \mathbb{E} \int_0^\tau r_k^2 \mathrm{d}t \\
&= \sum_{k\geq 1} \lambda_k^{2\gamma-2+2\delta_2} \mathbb{E} \int_0^\tau r_k^2 \mathrm{d}t \\
&= \mathbb{E} \int_0^\tau |A^{\gamma-1+\delta_2} R|^2 \mathrm{d}t. \qquad (4.13)
\end{aligned}$$

Note that under the initial assumptions, in the case of a bounded domain, (1.3), $\gamma - 1 + \delta_2 < 5/4$, and in the periodic case, (1.2), we have $\gamma - 1 + \delta_2 < \gamma + 1/2$. In either case, by taking τ_n as in Proposition 4.2, we infer from (4.13) with (4.3) that, for every n,

$$\sum_{k\geq 1} \frac{\operatorname{Var} \zeta_k^{T\wedge\tau_n}}{b_k^2} < \infty.$$

By applying Lemma 2.2, we conclude that, for each n fixed,

$$\lim_{N\to\infty} \frac{\int_0^{T\wedge\tau_n} \left\langle A^{1+2\alpha-\gamma} R^N, \sum_{k=1}^N \Phi_k \mathrm{d}W_k \right\rangle}{\lambda_N^{2+2\alpha-2\gamma-\delta_2}} = 0 \quad \text{in Probability.}$$

Since τ_n is increasing, $\tilde{\Omega} = \cup_n \{\tau_n > T\}$ is a set of full measure, and a simple estimate yields that

$$\lim_{N\to\infty} \frac{\int_0^{T} \left\langle A^{1+2\alpha-\gamma} R^N, \sum_{k=1}^N \Phi_k \mathrm{d}W_k \right\rangle}{\lambda_N^{2+2\alpha-2\gamma-\delta_2}} = 0 \quad \text{in Probability.}$$

The proof is complete. $\square$

Corollary 4.5. *Putting the admissible values $\delta_1 = \delta_2 = 0$ in (4.11), (4.12), and taking into account that $U^N = \bar{U}^N + R^N$, we conclude that*

$$\lim_{N\to\infty} \frac{\int_0^{T} \left\langle A^{1+2\alpha-\gamma} U^N, \sum_{k=1}^N \Phi_k \mathrm{d}W_k \right\rangle}{\int_0^T |A^{1+\alpha} U^N|^2 \mathrm{d}t} = 0 \quad \textit{in Probability.}$$

Thus, by representation (4.8) we have that $\widetilde{\nu}_N$ is weakly consistent estimator of the true parameter ν.

We turn next to the 'nonlinear terms' appearing in (4.9).

Proposition 4.6. *Assume the conditions and notation imposed for Lemma 4.3 above. We suppose that $\delta \in [0, \min\{5/4 - \gamma, \alpha - \gamma + 1\})$ in the case (1.3) or that $\delta \in [0, \min\{1/2, \alpha - \gamma + 1\})$ when we assume (1.2). Then*

$$\lim_{N\to\infty} N^{\delta} \frac{\int_0^T \left\langle A^{1+2\alpha} U^N, P_N B(U)\right\rangle \mathrm{d}t}{\int_0^T |A^{1+\alpha} U^N|^2 \mathrm{d}t} = 0 \quad a.s. \tag{4.14}$$

Proof. By the Cauchy–Schwartz inequality,

$$\left| \frac{\int_0^T \langle A^{1+2\alpha} U^N, P_N B(U)\rangle \mathrm{d}t}{\int_0^T |A^{1+\alpha} U^N|^2 \mathrm{d}t} \right| \le \left(\frac{\int_0^T |A^{\alpha} P_N B(U)|^2 \mathrm{d}t}{\int_0^T |A^{1+\alpha} U^N|^2 \mathrm{d}t} \right)^{1/2}. \tag{4.15}$$

Due to Lemma 4.3, (4.5) and (2.11) it is therefore sufficient to show that

$$\lim_{N\to\infty} \lambda_N^{2(\delta-(\alpha-\gamma+1))} \int_0^T |A^{\alpha} P_N B(U)|^2 \mathrm{d}t = 0 \quad \text{a.s.} \tag{4.16}$$

We begin with the boundary conditions (1.2) and consider two possibilities corresponding to different values of α. First suppose that $\alpha < \gamma - 1/2$, so $\delta - (\alpha - \gamma + 1) < 0$. Pick any $\beta \in (\max\{\alpha, 1/2\}, \gamma - 1/2)$. Making use of (2.19) and then applying Theorem 2.6, (ii), we observe that

$$\begin{aligned}\int_0^T |A^{\alpha} P_N B(U)|^2 \mathrm{d}t &\le \int_0^T |A^{\beta} B(U)|^2 \mathrm{d}t \\ &\le c \int_0^T |A^{\beta} U|^2 |A^{\beta+1/2} U|^2 \mathrm{d}t < \infty \quad \text{a.s.}\end{aligned}$$

and (4.16) follows.

Now suppose that $\alpha \ge \gamma - 1/2$. In this case we pick an element $\alpha' \in (\max\{\delta + \gamma - 1, 1/2\}, \gamma - 1/2)$. Note that, by assumption, $\delta < 1/2$, so this interval is nontrivial. Clearly $\alpha' < \alpha$ and we apply (2.3) and again (2.19) in order to estimate

$$\begin{aligned}\int_0^T |A^{\alpha} P_N B(U)|^2 \mathrm{d}t &\le \lambda_N^{2(\alpha-\alpha')} \int_0^T |A^{\alpha'} P_N B(U)|^2 \mathrm{d}t \\ &\le c\lambda_N^{2(\alpha-\alpha')} \int_0^T |A^{\alpha'} U|^2 |A^{\alpha'+1/2} U|^2 \mathrm{d}t < \infty \quad \text{a.s.}\end{aligned} \tag{4.17}$$

As above, we find that the quantity on the right hand side is finite due to Theorem 2.6, (ii). Noting that $\delta - (\alpha - \gamma + 1) + \alpha - \alpha' < 0$, we infer that (4.16) holds true.

The case of Dirichlet boundary conditions (1.3) is addressed in a similar manner. When $\alpha < 1/4$ we directly apply (2.18) to infer (4.16). When $\alpha \ge 1/4$ we pick any $\alpha' \in (\delta+\gamma-1, 1/4)$. Noting that the conditions on δ ensure that this interval is nontrivial and that $\alpha' < \alpha$, we apply (2.3) and (2.18) in a similar manner to (4.17) and infer (4.16) for this case too. The proof is now complete. □

Corollary 4.7. *In similar manner one can establish the same results as above for $P_N B(U^N)$. In particular, for $\delta = 0$ we have*

$$\lim_{N\to\infty} \frac{\int_0^T \langle A^{1+2\alpha} U^N, P_N B(U)\rangle \mathrm{d}t}{\int_0^T |A^{1+\alpha} U^N|^2 \mathrm{d}t} = 0 \quad a.s. \tag{4.18}$$

$$\lim_{N\to\infty} \frac{\int_0^T \langle A^{1+2\alpha}U^N, P_N B(U^N)\rangle \mathrm{d}t}{\int_0^T |A^{1+\alpha}U^N|^2 \mathrm{d}t} = 0 \quad a.s. \tag{4.19}$$

Taking into account the above equalities and the representations (4.9) and (4.10) we have that $\widehat{\nu}_N$ and $\check{\nu}_N$ are consistent estimators of ν.

4.3. Asymptotic normality

We finally address the asymptotic normality of $\tilde{\nu}_N$ and prove the second part of Theorem 3.1. Using the representation (4.8) for $\widetilde{\nu}_N$, Lemma 4.3, and (4.5) we see that is suffices to establish that

$$\lim_{N\to\infty} N \frac{\int_0^T \left\langle A^{1+2\alpha-\gamma}\bar{U}^N, \sum_k \Phi_k \mathrm{d}W_k \right\rangle}{\mathbb{E}\int_0^T |A^{1+\alpha}\bar{U}^N|^2 \mathrm{d}t} \overset{d}{=} \eta, \tag{4.20}$$

where η is a normal random variable with mean zero and variance $\frac{2\nu(\alpha-\gamma+1)^2}{\lambda_1 T(\alpha-\gamma+1/2)}$, and that

$$\lim_{N\to\infty} N \frac{\int_0^T \left\langle A^{1+2\alpha-\gamma}R^N, \sum_k \Phi_k \mathrm{d}W_k \right\rangle}{\mathbb{E}\int_0^T |A^{1+\alpha}\bar{U}^N|^2 \mathrm{d}t} = 0 \quad \text{in Probability.} \tag{4.21}$$

We establish (4.20) with the aid of Lemma 2.3. Let $\sigma_k := \lambda_k^{1+2\alpha-\gamma}\bar{u}_k$, and $\xi_k := \int_0^T \sigma_k^2 \mathrm{d}t$, $k \geq 1$. Notice that, due to (2.9),

$$\mathbb{E}[\xi_k] \sim \lambda_k^{2+4\alpha-2\gamma}\lambda_k^{-(1+2\gamma)} = \lambda_k^{1+4\alpha-4\gamma},$$
$$\mathrm{Var}\,[\xi_k] \sim \lambda_k^{4+8\alpha-4\gamma}\lambda_k^{-(3+4\gamma)} = \lambda_k^{1+8\alpha-8\gamma}.$$

Define $b_n := \sum_{k=1}^n \mathbb{E}\xi_k$. Under the given assumptions, $1+4\alpha-4\gamma < -1$, so by (2.12) we have that $b_n \sim \lambda_n^{2+4\alpha-4\gamma}$. We infer that b_n is increasing and unbounded. Moreover,

$$\sum_{k=1}^{\infty} \frac{\mathrm{Var}\,[\xi_k]}{b_k^2} \leq c \sum_{k=1}^{\infty} k^{-3},$$

and therefore by LLN, Lemma 2.2, we conclude

$$\lim_{N\to\infty} \frac{\sum_{k=1}^N \xi_k}{\sum_{k=1}^N \mathbb{E}\xi_k} = 1 \quad \text{a.s.}$$

Consequently, by Lemma 2.3 with σ_k defined above, we have

$$\lim_{N\to\infty} \frac{\int_0^T \left\langle A^{1+2\alpha-\gamma}\bar{U}^N, \sum_k \Phi_k \mathrm{d}W_k \right\rangle}{\left(\mathbb{E}\int_0^T |A^{1+2\alpha-\gamma}\bar{U}^N|^2 \mathrm{d}t\right)^{1/2}} \overset{d}{=} \mathcal{N}(0,1). \tag{4.22}$$

Noting that both $1+\alpha > \gamma$, $1+2\alpha-\gamma > \gamma$ we may apply (2.11) and infer

$$\frac{\left(\mathbb{E}\int_0^T |A^{1+2\alpha-\gamma}\bar{U}^N|^2 \mathrm{d}t\right)^{1/2}}{\mathbb{E}\int_0^T |A^{1+\alpha}\bar{U}^N|^2 \mathrm{d}t} \approx \sqrt{\frac{2\nu}{\lambda_1 T}} \cdot \frac{\alpha-\gamma+1}{\sqrt{\alpha-\gamma+1/2}} \cdot \frac{1}{N}. \tag{4.23}$$

By combining (4.22) and (4.23) we obtain (4.20).

The other condition, (4.21), follows directly from Proposition 4.4, (4.12) with the admissible value of $\delta_2 = 1$. This completes the proof of Theorem 3.1, (ii).

Remark 4.8. Notice that the accuracy of the estimators $\widetilde{\nu}_N$, as measured by the variance from asymptotic normality, depends on α, γ and T. As one would expect, as T gets larger and more information is revealed, the quality of the estimator improves (i.e. the variance decreases). This suggests that one may show that all of the classes of estimators considered above are consistent in the large time asymptotics regime, i.e. when we fix N and send $T \to \infty$. With the spectral method that we have developed in this work establishing the long time asymptotics is more complicated in comparison to the linear case, and this will be addressed in future work.

Also, we note that $(\alpha-\gamma+1)/\sqrt{\alpha-\gamma+1/2}$, as a function of α on the domain $\alpha > \gamma - 1/2$, reaches its minimum at $\alpha = \gamma$. Thus, for fixed T and γ, the smallest asymptotic variance for the estimator $\widetilde{\nu}_N$ corresponds to $\alpha = \gamma$. Observe that when $\alpha = \gamma$, the estimator $\widetilde{\nu}_N$ reduces to the formal MLE (3.3), which in some sense is the optimal estimator in this class of estimators.

Remark 4.9. We want to emphasize that we still believe that asymptotic normality properties similar to (3.7) also hold true for the estimators $\check{\nu}_N$. However, for a rigorous proof one needs to show, for example, that Proposition 4.6 holds true for some $\delta \geq 1$, and with $P_N B(U)$ replaced by $P_N B(U) - P_N B(U^N)$. Intuitively it is clear that the difference $P_N B(U) - P_N B(U^N)$ will make the convergence to zero in (4.14) faster, allowing for a larger δ compared to those from the terms $P_N B(U)$ and $P_N B(U^N)$ considered individually. We further believe that the quality of these estimators $\check{\nu}_N$ may be optimized in terms of the free parameter α. Although, up to the present time, we remain unable to establish such quantitative results about the asymptotic normality of the estimators $\check{\nu}_N$, we plan to study these questions at least by means of numerical simulations in forthcoming work.

Acknowledgements

We would like to thank the anonymous referees for their helpful comments and suggestions which greatly improved the final manuscript. IC acknowledges support from the National Science Foundation (NSF) grant DMS-0908099. The work of NGH was partially supported by the NSF under the grants DMS-1004638, and NSF-DMS-0906440 and by the Research Fund of Indiana University.

References

[1] K.I. Babenko, On the asymptotic behavior of eigenvalues of linearized Navier–Stokes equations, Dokl. Akad. Nauk SSSR 263 (3) (1982) 521–525.
[2] A. Bensoussan, R. Temam, Équations stochastiques du type Navier–Stokes, J. Funct. Anal. 13 (1973) 195–222.
[3] H. Breckner, Galerkin approximation and the strong solution of the Navier–Stokes equation, J. Appl. Math. Stochastic Anal. 13 (3) (2000) 239–259.
[4] Z. Brzeźniak, S. Peszat, Strong local and global solutions for stochastic Navier–Stokes equations, in: Infinite Dimensional Stochastic Analysis, Amsterdam, 1999, in: Verh. Afd. Natuurkd. 1. Reeks. K. Ned. Akad. Wet., vol. 52, R. Neth. Acad. Arts Sci., Amsterdam, 2000, pp. 85–98.

[5] I. Cialenco, N. Glatt-Holtz, Statistical inference for stochastic Navier–Stokes equations driven by multiplicative noise. Working Paper.
[6] P. Constantin, C. Foias, Navier–Stokes equations, in: Chicago Lectures in Mathematics, University of Chicago Press, Chicago, IL, 1988.
[7] A.B. Cruzeiro, Solutions et mesures invariantes pour des équations d'évolution stochastiques du type Navier–Stokes, Exposition. Math. 7 (1) (1989) 73–82.
[8] G. Da Prato, J. Zabczyk, Stochastic equations in infinite dimensions, in: Encyclopedia of Mathematics and its Applications, vol. 44, Cambridge University Press, Cambridge, 1992.
[9] A. Debussche, N. Glatt-Holtz, R. Temam, Local martingale and pathwise solutions for an abstract fluids model (submitted for publication).
[10] C.R. Doering, J.D. Gibbon, Applied analysis of the Navier–Stokes equations, in: Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 1995.
[11] B. Ewald, M. Petcu, R. Temam, Stochastic solutions of the two-dimensional primitive equations of the ocean and atmosphere with an additive noise, Anal. Appl. (Singap.) 5 (2) (2007) 183–198.
[12] F. Flandoli, D. Gatarek, Martingale and stationary solutions for stochastic Navier–Stokes equations, Probab. Theory Related Fields 102 (3) (1995) 367–391.
[13] N. Glatt-Holtz, R. Temam, Cauchy Convergence Schemes for Some Nonlinear Partial Differential Equations, Applicable Analysis—Special Issue in honor of Prof. V. Solonnikov (in press).
[14] N. Glatt-Holtz, M. Ziane, Strong pathwise solutions of the stochastic Navier–Stokes system, Adv. Differ. Equ. 14 (5–6) (2009) 567–600.
[15] N. Glatt-Holtz, R. Temam, Pathwise solutions of the 2-d stochastic primitive equations, Applied Mathematics and Optimization (in press).
[16] N. Glatt-Holtz, M. Ziane, The stochastic primitive equations in two space dimensions with multiplicative noise, Discrete Contin. Dyn. Syst. Ser. B 10 (4) (2008) 801–822.
[17] M. Huebner, R. Khasminskii, B.L. Rozovskii, Two examples of parameter estimation for stochastic partial differential equations, in: Stochastic processes, Springer, New York, 1993, pp. 149–160.
[18] M. Huebner, S.V. Lototsky, B.L. Rozovskii, Asymptotic properties of an approximate maximum likelihood estimator for stochastic PDEs, in: Statistics and Control of Stochastic Processes (Moscow, 1995/1996), World Sci. Publishing, 1997, pp. 139–155.
[19] M. Huebner, B.L. Rozovskii, On asymptotic properties of maximum likelihood estimators for parabolic stochastic PDE's, Probab. Theory Related Fields 103 (2) (1995) 143–163.
[20] V. Isakov, Inverse problems for partial differential equations, 2nd ed., in: Applied Mathematical Sciences, vol. 127, Springer, New York, 2006.
[21] A. Kirsch, An introduction to the mathematical theory of inverse problems, in: Applied Mathematical Sciences, vol. 120, Springer-Verlag, New York, 1996.
[22] Yu.A. Kutoyants, Statistical inference for ergodic diffusion processes, in: Springer Series in Statistics, Springer-Verlag London Ltd., London, 2004.
[23] J.-L. Lions, E. Magenes, Non-Homogeneous Boundary Value Problems and Applications. Vol. I, Springer-Verlag, New York, 1972. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181.
[24] R.S. Liptser, A.N. Shiryayev, Theory of martingales, in: Mathematics and its Applications (Soviet Series), vol. 49, Kluwer Academic Publishers Group, 1989.
[25] R.S. Liptser, A.N. Shiryayev, Statistics of Random Processes I. General Theory, 2nd ed., Springer-Verlag, New York, 2000.
[26] S.V. Lototsky, Statistical inference for stochastic parabolic equations: a spectral approach, Publ. Mat. 53 (1) (2009) 3–45.
[27] S.V. Lototsky, B.L. Rozovskii, Spectral asymptotics of some functionals arising in statistical inference for SPDEs, Stochastic Process. Appl. 79 (1) (1999) 69–94.
[28] S.V. Lototsky, B.L Rozovskii, Parameter estimation for stochastic evolution equations with non-commuting operators, in: V. Korolyuk, N. Portenko, H. Syta (Eds.), Skorohod's Ideas in Probability Theory, Institute of Mathematics of National Academy of Sciences of Ukraine, Kiev, Ukraine, 2000, pp. 271–280.
[29] G. Métivier, Valeurs propres d'opérateurs définis par la restriction de systèmes variationnels à des sous-espaces, J. Math. Pures Appl. (9) 57 (2) (1978) 133–156.
[30] R. Mikulevicius, B.L. Rozovskii, Stochastic Navier–Stokes equations for turbulent flows, SIAM J. Math. Anal. 35 (5) (2004) 1250–1310.
[31] C. Penland, A stochastic model of Indo-Pacific sea surface temperature anomalies, Physica D 96 (1996) 534–558.

[32] C. Penland, B.D. Ewald, On modelling physical systems with stochastic models: diffusion versus Lévy processes, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 366 (1875) (2008) 2457–2476.
[33] C. Penland, L. Matrosova, A balance condition for stochastic numerical models with application to the El Niño–southern oscillation, J. Climate 7 (1994) 1352–1372.
[34] C. Prévôt, M. Röckner, A concise course on stochastic partial differential equations, in: Lecture Notes in Mathematics, vol. 1905, Springer, Berlin, 2007.
[35] B.L. Rozovskii, Linear theory and applications to nonlinear filtering, in: Stochastic Evolution Systems, in: Mathematics and its Applications (Soviet Series), vol. 35, Kluwer Academic Publishers Group, Dordrecht, 1990.
[36] A.N. Shiryaev, Probability, 2nd ed., in: Graduate Texts in Mathematics, vol. 95, Springer-Verlag, New York, 1996.
[37] R. Temam, Navier–Stokes equations and nonlinear functional analysis, 2nd ed., in: CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 66, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1995.
[38] R. Temam, Navier–Stokes Equations: Theory and Numerical Analysis, AMS Chelsea Publishing, Providence, RI, 2001. Reprint of the 1984 edition.