Spectral Asymptotics of Some Functionals Arising in Statistical Inference for SPDEs

Sergey V. Lototsky Boris L. Rosovskii

Department of Mathematics, M.I.T., Room 2-247, 77 Massachusetts Avenue, Cambridge, MA 02139. E-mail: lototsky@math.mit.edu

Center for Applied Mathematical Sciences, University of Southern California, Los Angeles, CA 90089. E-mail: rozovski@cams-00.usc.edu

Published in *Stochastic Processes and Their Application*, volume 79, No. 1, pp. 69–94, 1999.

Abstract

A parameter estimation problem is considered for a stochastic evolution equation on a compact smooth manifold. Specifically, we concentrate on asymptotic properties of spectral estimates, i.e. estimates based on finite number of spatial Fourier coefficients of the solution. Under certain non-degeneracy assumptions the estimate is proved to be consistent, asymptotically normal and asymptotically efficient as the dimension of the projections increases. Unlike previous works on the subject, no commutativity is assumed between the operators in the equation.

Keywords. Asymptotic normality, Convergence of moments, Parameter estimation, Stochastic evolution equations.

1 Introduction

Asymptotic estimation theory for stochastic processes is a mature area with well developed methodology and substantial wealth of far reaching results, see e.g. [8, 13, 14]. Lately there has been a growing interest in extending the results and methods of this theory to statistical estimation of random fields, in particular, random fields driven by stochastic partial differential equations (SPDEs), see e.g. [1, 2, 9, 10, 21].

It turned out that such an extension is far from routine. The infinite dimensional nature of random fields poses substantial technical challenges and generates interesting new effects, uncharacteristic of inference for stochastic processes.

One of the most interesting new effects is that the amount of "information" recovered from the measurements is a natural asymptotic parameter in statistical inference for random fields. To clarify this rather obscure statement, let us consider two examples.

Let $u(t,x)$ be a solution of the following stochastic PDE

$$\begin{aligned} &du(t,x) = \theta\nabla^2 u(t,x)dt + \varepsilon dW(t,x), (t,x) \in (0,T] \times (0,1), \\ &u(0,x) = u_0(x), \\ &u(t,0) = u(t,1) = 0, \end{aligned} \tag{1.1}$$

where $\dot{W}(t,x)$ is a space time white noise, θ is an unknown parameter subject to estimation, and ε is the noise intensity. It was shown in [5, 6, 7] that, for fixed T and ε, the MLE for θ is super-efficient (i.e. θ can be reconstructed "exactly" from measurements of $u(t,x)$ on $(0,T] \times (0,1)$). More precisely, there exists a sequence of maximum likelihood estimators $\hat{\theta}_n$, based on partial information about the field $u(t,x)$; this sequence converges to θ with probability 1, as $n \to \infty$, or, equivalently, as the amount of information about $u(t,x)$ used to construct θ_n converges to the "total information" contained in the measurements of $u(t,x)$ for a.a. $(t,x) \in (0,T] \times (0,1)$.

In contrast, if $u(t)$ is a one dimensional Ornstein-Uhlenbeck process solving the Ito equation

$$\begin{aligned} &du(t) = \theta u(t)dt + \varepsilon dw(t), t \in (0,T], \\ &u(0) = u_0, \end{aligned} \tag{1.2}$$

then the MLE or any other estimate based on the whole trajectory of the process $u(t)$, $t \in [0,T]$ (i.e. utilizing the "total information" contained in the process $u(t)$, $t \leq T$) does not reconstruct θ exactly. Only if $T \to \infty$ or $\varepsilon \to 0$, does the MLE estimate converge to θ.

Parameter estimation for PDEs is a particular case of the inverse problem that arises when the solution of a certain equation is observed and conclusions must be made about the coefficients of the equation. In the deterministic setting, numerous examples of such problems in ecology, material sciences, biology, etc. are given in the book by Banks and Kunisch [3]. The stochastic term is usually introduced in the equation to take into account those components of the model that cannot be described exactly (see e.g. [21]).

The asymptotic properties of MLEs for parameters of SPDEs were studied first by Huebner, Khasminskii, Rozovskii [6] and further investigated by Huebner and Rozovskii [7], Huebner [5], Piterbarg, Rozovskii [22], Lototsky [17], etc. The first three papers deal with observations that are continuous in time, while the fourth paper is concerned with the discrete time case.

In [5, 7, 22] the parameter estimation problem was considered for the Dirichlet boundary value problem

$$\begin{aligned} &du(t,x) + (\mathcal{A}_0 + \theta\mathcal{A}_1)u(t,x)dt = \varepsilon dW(t,x), (t,x) \in (0,T] \times G, \\ &u(0,x) = u_0(x), \\ &u(t,x)_{|\partial G} = 0, \end{aligned} \tag{1.3}$$

where θ is the unknown parameter belonging to an open subset of the real line, $\mathcal{A}_0$,$\mathcal{A}_1$ are partial differential operators, and G is a domain in $\boldsymbol{R}^d$.

Since u is a random field indexed by an infinite set of points t, x, a computable estimate of θ must be based on some kind of finite-dimensional "projection" of u even if the whole trajectory is observed. In particular, in [5, 7, 22] it was assumed that the measurements are given in the spectral form, i.e. as a finite set of spatial Fourier coefficients of the field $u(t,x)$, $\Pi^K u(t,x) = ((u_1(t), ..., u_K(t))$, where $u_i(t) = \int_G u(t,x)e_i(x)dx$ and $(e_i(x))_{i\geq 1}$ is a complete orthonormal system in $L^2(G)$. This assumption is quite natural, because for many types of sensors the output is naturally presented in the spectral form (as Fourier modes). Even if the measurements are obtained in spatial scale, i.e. as measurements of $u(t,x)$ on some spatial grid $x_j, j = 1, 2, ...$, then one can approximate the Fourier coefficients of $u(t, x)$ using these measurements. The asymptotic properties of the MLE were studied as the dimension of those projections increases while the length T of the observation interval and the amplitude ε of the noise remain fixed.

The main technical assumption used in all those works was that the operators $\mathcal{A}_0$ and $\mathcal{A}_1$ in (1.3) are formally self-adjoint and have a common system of eigenfunctions (which, of course, implies that the operators $\mathcal{A}_0$ and $\mathcal{A}_1$ commute). These are very restrictive assumptions that essentially reduce the scope of applications to the operators with constant coefficients.

The objective of the current paper is to consider an estimate of θ for equation (1.3) in the non-commutative case, without assuming anything about the eigenfunctions of the operators in the equation. Some preliminary results in this direction were obtained in [16, 17]. For the sake of simplicity, the equation is considered on a compact smooth d - dimensional manifold so that no boundary conditions are involved. As in [5, 7, 22], it is assumed that the operator $\mathcal{A}_0 + \theta\mathcal{A}_1$ is elliptic for all admissible values of θ. In contrast to the case of commuting operators $\mathcal{A}_0$ and $\mathcal{A}_1$, in the general setting it is impossible to obtain an explicit form of the MLE. Instead, we are considering a quasi-MLE (QMLE), an explicitly computable estimate for θ that coincides with the MLE in the case of commuting operators.

We prove that the QMLE possesses essentially the same asymptotic properties that the MLE has when $\mathcal{A}_0$ and $\mathcal{A}_1$ commute. Specifically, we prove that if $\mathcal{A}_1$ is the leading operator, then the QMLE of θ is consistent and asymptotically normal, as the dimension K of the projections tends to infinity. On the other hand, if $\mathcal{A}_0$ is the leading operator, then the QMLE is consistent and asymptotically normal if

$$order(\mathcal{A}_1) \geq \frac{1}{2}\left(order(\mathcal{A}_0 + \theta\mathcal{A}_1) - d\right) \tag{1.4}$$

and the operator $\mathcal{A}_1$ satisfies a certain non-degeneracy property. In particular, condition (1.4) is necessary for consistency. It was shown in [7] that, in the case of the Dirichlet problem in a domain of $\boldsymbol{R}^d$, if the operators $\mathcal{A}_0$ and $\mathcal{A}_1$ are selfadjoint elliptic with a common system of eigenfunctions, then condition (1.4) is necessary and sufficient for consistency, asymptotic normality, and asymptotic efficiency of the estimate. When (1.4) does not hold, the asymptotic shift of the estimate is computed. We also establish the rate of convergence for the QMLE. The rate is the same as that of the MLE in the case of commuting operators $\mathcal{A}_0$ and $\mathcal{A}_1$. To characterize the asymptotic efficiency of the QMLE, we proved that the normalized difference between the QMLE and θ converges in some sense to a Gaussian random variable with zero mean and unit variance, as the dimension of the projection tends to infinity.

The detailed description of the setting is given in **Section 2** and the main results are presented in **Section 3**. The proof of the main theorem about the consistency and asymptotic normality is given in **Section 5**.

In **Section 4** an example is presented, illustrating how the results obtained can be applied to the estimation of either thermodiffusivity or the cooling coefficient in the heat balance equation with a variable velocity field.

2 The Setting

Let M be a d-dimensional compact orientable $\mathbf{C}^\infty$ manifold with a smooth positive measure dx. If $\mathcal{L}$ is an elliptic positive definite self-adjoint differential operator of order $2m$ on M, then the operator $\Lambda = \mathcal{L}^{1/(2m)}$ is elliptic of order 1 and generates the scale $\{\mathbf{H}^s\}_{s\in\boldsymbol{R}}$ of Sobolev spaces on M [12, 24]. All differential operators on M are assumed to be non-zero with real $\mathbf{C}^\infty(M)$ coefficients, and only real elements of $\mathbf{H}^s$ will be considered. The variable x will usually be omitted in the argument of functions defined on M.

In what follows, an alternative characterization of the spaces $\{\mathbf{H}^s\}$ will be used. By Theorem I.8.3 in [24], the operator $\mathcal{L}$ has a complete orthonormal system of eigenfunctions $\{e_k\}_{k\geq 1}$ in the space $L_2(M, dx)$ of square integrable functions on M. With no loss of generality, it can be assumed that each $e_k(x)$ is real. Then for every $f \in L_2(M, dx)$ the representation

$$f = \sum_{k\geq 1} \psi_k(f) e_k \tag{2.1}$$

holds, where

$$\psi_k(f) = \int_M f(x) e_k(x) dx. \tag{2.2}$$

If $l_k > 0$ is the eigenvalue of $\mathcal{L}$ corresponding to e_k and $\lambda_k := l_k^{1/(2m)}$, then, for $s \geq 0$, $\mathbf{H}^s = \{f \in L_2(M, dx) : \sum_{k\geq 1} \lambda_k^{2s} |\psi_k(f)|^2 < \infty\}$, and for $s < 0$, $\mathbf{H}^s$ is the closure of $L_2(M, dx)$ in the norm $\|f\|_s = \sqrt{\sum_{k\geq 1} \lambda_k^{2s} |\psi_k(f)|^2}$. As a result, every element f of the space $\mathbf{H}^s$, $s \in \boldsymbol{R}$, can be identified with a sequence $\{\psi_k(f)\}_{k\geq 1}$ such that $\sum_{k\geq 1} \lambda_k^{2s} |\psi_k(f)|^2 < \infty$. The space $\mathbf{H}^s$, equipped with the inner product

$$(f, g)_s = \sum_{k\geq 1} \lambda_k^{2s} \psi_k(f) \psi_k(g), \;\; f, g \in \mathbf{H}^s, \tag{2.3}$$

is a Hilbert space.

Below, notation $a_k \asymp b_k$ means

$$0 < c_1 \leq \liminf_{k\to\infty} (a_k/b_k) \leq \limsup_{k\to\infty} (a_k/b_k) \leq c_2 < \infty. \tag{2.4}$$

A **cylindrical Brownian motion** $W = (W(t))_{0\leq t\leq T}$ on M is defined as follows: for every $t \in [0, T]$, $W(t)$ is the element of $\cup_s \mathbf{H}^s$ such that $\psi_k(W(t)) = w_k(t)$, where $\{w_k\}_{k\geq 1}$ is a collection of independent one dimensional Wiener processes on the given probability space $(\Omega, \mathcal{F}, I\!\!F, \mathbf{P})$ with a complete filtration $I\!\!F = \{\mathcal{F}_t\}_{0\leq t\leq T}$. Since by Theorem II.15.2 in [24] $\lambda_k \asymp k^{1/d}$, $k \to \infty$,it follows that $W(t) \in \mathbf{H}^s$ for every $s < -d/2$. Direct computations show that W is an $\mathbf{H}^s$ - valued Wiener process with the covariance operator Λ^{2s}. This definition of W agrees with the alternative definitions of the cylindrical Brownian motion [18, 19, 25].

Let $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{N}$ be differential operators on M of orders $order(\mathcal{A})$, $order(\mathcal{B})$, and $order(\mathcal{N})$ respectively. It is assumed that

$$\max(order(\mathcal{A}), order(\mathcal{B}), order(\mathcal{N})) < 2m. \tag{2.5}$$

Consider the random field u defined on M by the evolution equation

$$du(t) + [\theta_1(\mathcal{L} + \mathcal{A}) + \theta_2 \mathcal{B} + \mathcal{N}] u(t) dt = dW(t), \;\; 0 < t \leq T, \;\; u(0) = u_0. \tag{2.6}$$

Here $\theta_1 > 0$, $\theta_2 \in \boldsymbol{R}$, and the dependence of u and W on x and ω is suppressed.

If the trajectory $u(t)$, $0 \le t \le T$, is observed, then the following scalar parameter estimation problems can be stated:

1). estimate θ_1 assuming that θ_2 is known;

2). estimate θ_2 assuming that θ_1 is known.

Remark 2.1 *The general model*

$$du(t) + [\theta_1 \mathcal{A}_0 + \theta_2 \mathcal{A}_1 + \mathcal{N}]u(t)dt = dW(t),\ 0 < t \le T,\ u(0) = u_0 \tag{2.7}$$

is reduced to (2.6) if the operator $\theta_1\mathcal{A}_0 + \theta_2\mathcal{A}_1$ is elliptic of order $2m$ for all admissible values of parameters θ_1, θ_2 and $order(\mathcal{A}_0) \ne order(\mathcal{A}_1)$. For example, if $order(\mathcal{A}_1) = 2m$, then $\mathcal{L} = (\mathcal{A}_1 + \mathcal{A}_1^*)/2 + (c+1)I$, $\mathcal{A} = (\mathcal{A}_1 - \mathcal{A}_1^*)/2 - (c+1)I$, $\mathcal{B} = \mathcal{A}_0$, where c is the lower bound on eigenvalues of $(\mathcal{A}_1 + \mathcal{A}_1^*)/2$ and I is the identity operator. Indeed, by Corollary 2.1.1 in [12], if an operator $\mathcal{P}$ is of even order with real coefficients, then the operator $\mathcal{P} - \mathcal{P}^*$ is of lower order than $\mathcal{P}$. With obvious modifications, the results presented below are also valid when the operators $\mathcal{A}_0$, $\mathcal{A}_1$ have the same order under an additional assumption that $\mathcal{A}_i = \mathcal{L}_i + \mathcal{A}_i'$, $i = 1, 2$, where the operators $\mathcal{L}_i$ are elliptic of order $2m$ with a common system of eigenfunctions and $\mathcal{A}_i'$ are operators of lower order.

Before discussing possible solutions to the above parameter estimation problems, let us recall some analytical properties of the field u.

Theorem 2.2 *For every $s > d/2$, if u_0 belongs to $L_2(\Omega; \mathbf{H}^{-s})$ and is $\mathcal{F}_0$–measurable, then equation (2.6) has a unique $\mathcal{F}_t$-adapted solution $u = u(t)$ so that*

$$u \in L_2(\Omega \times [0,T]; \mathbf{H}^{-s+m}) \cap L_2(\Omega; \mathbf{C}([0,T]; \mathbf{H}^{-s})) \tag{2.8}$$

with

$$\mathbf{E} \sup_{t\in[0,T]} \|u(t)\|_{-s}^2 + \mathbf{E}\int_0^T \|u(t)\|_{-s+m}^2 dt \le CT \sum_{k\ge1} \lambda_k^{-2s} + C_1(T)\mathbf{E}\|u_0\|_{-s}^2 < \infty. \tag{2.9}$$

Proof. By assumption, $\max(order(\mathcal{A}), order(\mathcal{B}), order(\mathcal{N})) < 2m$ and $\theta_1 > 0$. Then ellipticity of the operator $\mathcal{L}$ implies that for every $s \in \boldsymbol{R}$ there exist positive constants C_1 and C_2 so that for every $f \in \mathbf{C}^\infty$

$$-((\theta_1(\mathcal{L} + \mathcal{A}) + \theta_2\mathcal{B} + \mathcal{N})f, f)_s \le -C_1\|f\|_{s+m}^2 + C_2\|f\|_s^2, \tag{2.10}$$

which means that the operator $-(\theta_1(\mathcal{L}+\mathcal{A}) + \theta_2\mathcal{B} + \mathcal{N})$ is coercive in every normal triple $\{\mathbf{H}^{s+m}, \mathbf{H}^d, \mathbf{H}^{s-m}\}$. The statement of the theorem now follows from Theorem 3.1.4 in [23].

3 The Estimate and Its Properties

Both parameter estimation problems for (2.6) can be stated as follows: estimate $\theta \in \Theta$ from the observations of

$$du^{\theta}(t) + (\mathcal{A}_0 + \theta\mathcal{A}_1)u^{\theta}(t)dt = dW(t). \tag{3.1}$$

Indeed, if θ_2 is known, then $\mathcal{A}_0 = \theta_2\mathcal{B} + \mathcal{N}$, $\theta = \theta_1$, $\Theta = (0, +\infty)$, $\mathcal{A}_1 = \mathcal{L} + \mathcal{A}$, and if θ_1 is known, then $\mathcal{A}_0 = \theta_1(\mathcal{L} + \mathcal{A}) + \mathcal{N}$, $\theta = \theta_2$, $\Theta = \boldsymbol{R}$, $\mathcal{A}_1 = \mathcal{B}$. All main results will be stated in terms of (2.6), and (3.1) will play an auxiliary role.

It is assumed that the observed field u satisfies (3.1) for some unknown but fixed value θ^0 of the parameter θ. Depending on the circumstances, θ^0 can correspond to either θ_1 or θ_2 in (2.6), the other parameter being fixed and known. Even though the whole random field $u^{\theta^0}(t,x)$ is observed, the estimate of θ^0 will be computed using only finite dimensional processes $\Pi^K u^{\theta^0}$, $\Pi^K \mathcal{A}_0 u^{\theta^0}$, and $\Pi^K \mathcal{A}_1 u^{\theta^0}$. The operator Π^K used to construct the estimate is defined as follows: for every $f = \{\psi_k(f)\}_{k\geq 1} \in \cup_s \mathbf{H}^s$,

$$\Pi^K f = \sum_{k=1}^{K} \psi_k(f) e_k. \tag{3.2}$$

By (3.1),

$$d\Pi^K u^{\theta}(t) + \Pi^K(\mathcal{A}_0 + \theta\mathcal{A}_1)u^{\theta}(t)dt = dW^K(t), \tag{3.3}$$

where $W^K(t) = \Pi^K W(t)$. The process $\Pi^K u^{\theta} = (\Pi^K u^{\theta}(t), \mathcal{F}_t)_{0\leq t\leq T}$ is finite dimensional, continuous in the mean square, and Gaussian, but not, in general, a diffusion process because the operators $\mathcal{A}_0$ and $\mathcal{A}_1$ need not commute with Π^K. Denote by $\mathbf{P}^{\theta,K}$ the measure in $\mathbf{C}([0,T];\Pi^K(\mathbf{H}^0))$, generated by the solution of (3.3). The measure $\mathbf{P}^{\theta,K}$ is absolutely continuous with respect to the measure $\mathbf{P}^{\theta^0,K}$ for all $\theta \in \Theta$ and $K \geq 1$. Indeed, denote by $\mathcal{F}_t^{K,\theta}$ the σ-algebra generated by $\Pi^K u^{\theta}(s)$, $0 \leq s \leq t$, and let $U_t^{\theta,K}(X)$ be the operator from $\mathbf{C}([0,T];\Pi^K(\mathbf{H}^0))$ to $\mathbf{C}([0,T];\Pi^K(\mathbf{H}^0))$ such that, for all $t \in [0,T]$ and $\theta \in \Theta$,

$$U_t^{\theta,K}(\Pi^K u^{\theta}) = \mathbf{E}\left(\Pi^K(\mathcal{A}_0 + \theta\mathcal{A}_1)u^{\theta}|\mathcal{F}_t^{K,\theta}\right) \qquad (\mathbf{P}\text{- a.s.}) \tag{3.4}$$

Then, by Theorem 7.12 in [15], the process $\Pi^K u^{\theta}$ satisfies

$$d\Pi^K u^{\theta}(t) = U_t^{\theta,K}(\Pi^K u^{\theta})dt + d\tilde{W}^{\theta,K}(t), \ \Pi^K u^{\theta}(0) = \Pi^K u_0, \tag{3.5}$$

where $\tilde{W}^{\theta,K}(t) = \sum_{k=1}^K \tilde{w}_k^{\theta}(t)e_k$ and $\tilde{w}_k^{\theta}(t)$, $k = 1,\dots,K$, are independent one dimensional standard Wiener processes, possibly different for different θ. Since $\left\{\Pi^K(\mathcal{A}_0 + \theta\mathcal{A}_1)u^{\theta}, W^K\right\}$ is a Gaussian system for every $\theta \in \Theta$, it follows from Theorem 7.16 and Lemma 4.10 in [15] that

$$\begin{aligned}\frac{d\mathbf{P}^{\theta,K}}{d\mathbf{P}^{\theta^0,K}}(\Pi^K u^{\theta^0}) = \exp\Bigg\{ &\int_0^T \left(U_t^{\theta,K}(\Pi^K u^{\theta^0}) - U_t^{\theta^0,K}(\Pi^K u^{\theta^0}), d\Pi^K u^{\theta^0}(t)\right)_0 - \\ &\frac{1}{2}\int_0^T \left(\|U_t^{\theta,K}(\Pi^K u^{\theta^0})\|_0^2 - \|U_t^{\theta^0,K}(\Pi^K u^{\theta^0})\|_0^2\right) dt\Bigg\}.\end{aligned} \tag{3.6}$$

By definition, the maximum likelihood estimate (MLE) of θ^0 is then equal to

$$\arg\max_{\theta} \left(\frac{d\mathbf{P}^{\theta,K}}{d\mathbf{P}^{\theta^0,K}}\right)(\Pi^K u^{\theta^0}), \tag{3.7}$$

but since, in general, the functional $U_t^{\theta,K}(X)$ is not known explicitly, this estimate cannot be computed. The situation is much simpler if the operators $\mathcal{A}_0$ and $\mathcal{A}_1$ commute with Π^K so that $\Pi^K\mathcal{A}_i = \Pi^K\mathcal{A}_i\Pi^K$, $i = 0, 1$, and $U_t^{\theta,K}(X) = \Pi^K(\mathcal{A}_0 + \theta\mathcal{A}_1)X(t)$; in this case, the MLE $\hat{\theta}^K$ of θ^0 is computable and, as shown in [7],

$$\hat{\theta}^K = \frac{\int_0^T (\Pi^K\mathcal{A}_1 u^{\theta^0}(t), d\Pi^K u^{\theta^0}(t) - \Pi^K\mathcal{A}_0 u^{\theta^0}(t)dt)_0}{\int_0^T \|\Pi^K\mathcal{A}_1 u^{\theta^0}(t)\|_0^2 dt} \tag{3.8}$$

with the convention $0/0 = 0$.

Of course, expression (3.8) is well defined even when the operators $\mathcal{A}_0$ and $\mathcal{A}_1$ do not commute with Π^K. If $u^{\theta^0}(t)$ is observed on the interval $(0, T)$, then the right hand side of (3.8) is computable. Indeed, it is readily checked that

$$\frac{\int_0^T (\Pi^K\mathcal{A}_1 u^{\theta^0}(t), d\Pi^K u^{\theta^0}(t) - \Pi^K\mathcal{A}_0 u^{\theta^0}(t)dt)_0}{\int_0^T \|\Pi^K\mathcal{A}_1 u^{\theta^0}(t)\|_0^2 dt} \tag{3.9}$$

$$= \frac{\sum_{i=1}^K \int_0^T (u^{\theta^0}(t), \mathcal{A}_1^* e_i)_0 (du_i^{\theta^0}(t) - (u^{\theta^0}(t), \mathcal{A}_0^* e_i)_0 dt)}{\sum_{i=1}^K \int_0^T (u^{\theta^0}(t), \mathcal{A}_1^* e_i)_0^2 dt} \tag{3.10}$$

where $u_i^{\theta^0}(t) = (u^{\theta^0}(t), e_i)$, and $\mathcal{A}_j^*$ is the operator formally adjoint to $\mathcal{A}_j$.

Even though (3.8) is not, in general, the maximum likelihood estimate of θ^0, it is a natural estimate to consider. We will call it the *quasi maximum likelihood estimate (QMLE)* of θ^0. In what follows, it will be shown that the QMLE possesses essentially the same asymptotic properties that the MLE has when $\mathcal{A}_0$ and $\mathcal{A}_1$ commute.

To simplify the notations, the superscript θ^0 will be omitted wherever possible so that $u(t)$ is the solution of (2.6) or (3.1), corresponding to the true value of the unknown parameter. To study the properties of (3.8), note first of all that

$$\mathbf{P}\{\int_0^T \|\Pi^K\mathcal{A}_1 u(t)\|_0^2 dt > 0\} = 1 \tag{3.11}$$

for all sufficiently large K. Indeed, by assumption, the operator $\mathcal{A}_1$ is not identical zero and therefore $\left(\Pi^K\mathcal{A}_1 W_t\right)_{t\geq 0}$ is a continuous nonzero square integrable martingale, while

$$\left(\int_0^t \Pi^K\mathcal{A}_1[\mathcal{A}_0 + \theta^0\mathcal{A}_1]u(s)ds\right)_{t\geq 0} \tag{3.12}$$

is a continuous process with bounded variation.

It then follows from (3.8) and (3.11) that

$$\hat{\theta}^K = \theta^0 + \frac{\int_0^T (\Pi^K\mathcal{A}_1 u(t), dW^K(t))_0}{\int_0^T \|\Pi^K\mathcal{A}_1 u(t)\|_0^2 dt} \qquad (\mathbf{P}\text{- a.s.}) \tag{3.13}$$

Representation (3.13) will be used to study the asymptotic properties of $\hat{\theta}^K$ as $K \to \infty$. To get a consistent estimate, it is intuitively clear that $\int_0^T \|\Pi^K\mathcal{A}_1 u(t)\|_0^2 dt$ should tend to infinity as $K \to \infty$, and this requires certain non-degeneracy of the operator $\mathcal{A}_1$.

Definition 3.1 *A differential operator* $\mathcal{P}$ *of order* p *on* M *is called* **essentially non-degenerate** *if for every* $s \in \boldsymbol{R}$ *there exist positive constants* $\varepsilon,\ L,\ \delta$ *so that*

$$\|\mathcal{P}f\|_s^2 \geq \varepsilon\|f\|_{s+p}^2 - L\|f\|_{s+p-\delta}^2 \tag{3.14}$$

for all $f \in \mathbf{C}^\infty(M)$.

If the operator $\mathcal{P}^*\mathcal{P}$ is elliptic of order $2p$, then the operator $\mathcal{P}$ is essentially non-degenerate, because in this case the operator $\mathcal{P}^*\mathcal{P}$ is positive definite and self-adjoint so that the operator $(\mathcal{P}^*\mathcal{P})^{1/(2p)}$ generates an equivalent scale of Sobolev spaces on M. In particular, every elliptic operator satisfies (3.14). Since, by Corollary 2.1.2 in [12], for every differential operator $\mathcal{P}$ the operator $\mathcal{P}^*\mathcal{P} - \mathcal{P}\mathcal{P}^*$ is of order at most $2p-1$, the operator $\mathcal{P}$ is essentially non-degenerate if and only if $\mathcal{P}^*$ is.

Let us now formulate the main result concerning the properties of the estimate (3.13). Recall that the observed field u satisfies

$$du(t) + [\theta_1(\mathcal{L}+\mathcal{A}) + \theta_2\mathcal{B} + \mathcal{N}]u(t)dt = dW(t),\ 0 < t \leq T;\ u(0) = u_0, \tag{3.15}$$

with one of $\theta_2 = \theta_2^0$ or $\theta_1 = \theta_1^0$ known. According to (3.13), the estimate of the remaining parameter is given by

$$\hat{\theta}_1^K = \frac{\int_0^T (\Pi^K(\mathcal{L}+\mathcal{A})u(t), d\Pi^K du(t) - d\Pi^K(\theta_2^0\mathcal{B}+\mathcal{N})u(t))_0}{\int_0^T \|\Pi^K(\mathcal{L}+\mathcal{A})u(t)\|_0^2 dt}, \tag{3.16}$$

$$\hat{\theta}_2^K = \frac{\int_0^T (\Pi^K\mathcal{B}u(t), d\Pi^K du(t) - d\Pi^K(\theta_1^0(\mathcal{L}+\mathcal{A})+\mathcal{N})u(t))_0}{\int_0^T \|\Pi^K\mathcal{B}u(t)\|_0^2 dt}. \tag{3.17}$$

The following assumptions will be in force throughout the rest of the section.

H1. Equation (3.15) is considered on a compact d-dimensional smooth manifold M;

H2. $\theta_1^0 > 0$, $\theta_2^0 \in \boldsymbol{R}$;

H3. $\mathcal{L}$ is a positive definite self-adjoint elliptic operator of order $2m$;

H4. $\max(order(\mathcal{A}), order(\mathcal{B}), order(\mathcal{N})) < 2m$;

H5. u_0 is $\mathcal{F}_0$-measurable, $u_0 \in L_2(\Omega; \mathbf{H}^{-d/2})$, and u_0 is independent of W.

Theorem 3.2 *If* θ_2 *is known, then the estimate (3.16) of* θ_1^0 *is consistent and asymptotically normal:*

$$\begin{aligned} &\mathbf{P} - \lim_{K\to\infty} |\hat{\theta}_1^K - \theta_1^0| = 0;\\ &\Psi_{K,1}\cdot(\theta_1^0 - \hat{\theta}_1^K) \xrightarrow{d} \mathcal{N}(0,1), \end{aligned} \tag{3.18}$$

where $\Psi_{K,1} = \sqrt{(T/(2\theta_1^0))\sum_{n=1}^K l_n}$.

If θ_1 *is known, then the estimate (3.17) of* θ_2^0 *is consistent and asymptotically normal under an additional assumption that the operator* $\mathcal{B}$ *is essentially non-degenerate and* $order(\mathcal{B}) = b \geq m - d/2$. *In that case,*

$$\begin{aligned} &\mathbf{P} - \lim_{K\to\infty} |\hat{\theta}_2^K - \theta_2^0| = 0;\\ &\Psi_{K,2}\cdot(\theta_2^0 - \hat{\theta}_2^K) \xrightarrow{d} \mathcal{N}(0,1), \end{aligned} \tag{3.19}$$

where $\Psi_{K,2} \asymp \sqrt{\sum_{n=1}^K l_n^{(b-m)/m}}$.

This theorem is proved in Section 5.

Remark 3.3 *1. Since $l_k \asymp k^{2m/d}$, the rate of convergence for $\hat{\theta}_1^K$ is $\Psi_{K,1} \asymp K^{m/d+1/2}$, and for $\hat{\theta}_2^K$, it is*

$$\Psi_{K,2} \asymp \begin{cases} K^{(b-m)/d+1/2}, & \text{if } b > m - d/2, \\ \ln K, & \text{if } b = m - d/2. \end{cases} \tag{3.20}$$

2. Rephrasing Theorem 3.2 in terms of the operators $\mathcal{A}_0$ and $\mathcal{A}_1$ introduced in the beginning of the Section, we can say that if the assumption

H_0. The operator $\mathcal{A}_1$ is essentially non-degenerate and

$$order(\mathcal{A}_1) \geq \frac{1}{2}\left(order(\mathcal{A}_0 + \theta\mathcal{A}_1) - d\right) \tag{3.21}$$

holds true, then the QMLEs (3.16) and (3.17) are consistent and asymptotically normal. It can be shown (see Lemma 5.1 below) that the assumption H_0 yields that $P-a.e.$

$$\lim_{K\to\infty} \int_0^T \|\Pi^K \mathcal{A}_1 u\|_0^2 dt = +\infty. \tag{3.22}$$

On the other hand, it is known (see [19]) that (3.22) holds if and only if the distributions of u^θ for different values of θ are singular. It was shown in [7] that if $\mathcal{A}_1$ is an elliptic operator commuting with $\mathcal{A}_0$, and some other less important conditions are satisfied, then the distributions of u^θ for different values of θ are singular if and only if inequality (3.21) holds.

3. All the statements of the theorem remain true if, instead of differential operators, pseudo-differential operators of class $S^{\cdot}_{\rho,\delta}$ are considered with $\rho > \delta$ [12, 24].

Denote by Ξ the set of real valued non-negative functions $h = h(x)$, $x \in \boldsymbol{R}$, that are non-decreasing for $x > 0$ and satisfy $h(0) = 0$, $h(-x) = h(x)$.

Theorem 3.4 *Assume that $\mathbf{E}\|u_0\|_{-d/2}^q < \infty$ for all $q > 0$. Let $h \in \Xi$ be a function so that $|h(x)| \leq C \cdot (1 + |x|^\sigma)$ for some $C, \sigma > 0$. Denote by ξ_g a Gaussian random variable with zero mean and unit covariance.*

If θ_2 is known, then the estimate (3.16) of θ_1^0 satisfies

$$\lim_{K\to\infty} \mathbf{E}h\left(\Psi_{K,1} \cdot (\hat{\theta}_1^K - \theta_1^0)\right) = \mathbf{E}h(\xi_g). \tag{3.23}$$

If θ_1 is known, the operator $\mathcal{B}$ is essentially non-degenerate, and $order(\mathcal{B}) \geq m - d/2$, then the estimate (3.17) of θ_2^0 satisfies

$$\lim_{K\to\infty} \mathbf{E}h\left(\Psi_{K,2} \cdot (\hat{\theta}_2^K - \theta_2^0)\right) = \mathbf{E}h(\xi_g). \tag{3.24}$$

The proof of Theorem 3.4 is based on the following result to be proved later.

Lemma 3.5 *If $\mathcal{P}$ is an essentially non-degenerate operator of order $p > m - d/2$ and*

$$\Psi_K = \sqrt{\mathbf{E}\int_0^T \|\Pi^K \mathcal{P} u(t)\|_0^2 dt}, \tag{3.25}$$

then for every $q > 0$ there exists a $K_0 = K_0(q) > 0$ so that

$$\sup_{K \geq K_0} \mathbf{E} \left| \frac{\int_0^T \left(\Pi^K \mathcal{P} u(t), dW^K(t)\right)_0}{\int_0^T \|\Pi^K \mathcal{P} u(t)\|_0^2 dt} \cdot \Psi_K \right|^q < \infty. \tag{3.26}$$

Proof of Theorem 3.4. With no loss of generality, it can be assumed that the function h is continuous. Indeed, the monotonicity assumption implies that h has at most countably many discontinuities, while the random variables in question have densities with respect to the Lebesgue measure. After that, the statements of the theorem follow from Theorem 3.2, since Lemma 3.5 and equality (3.13) imply that the families of random variables $\left\{h\left(\Psi_{K,1} \cdot (\hat{\theta}_1^K - \theta_1^0)\right),\ K \geq K_0(\sigma+1)\right\}$ and $\left\{h\left(\Psi_{K,2} \cdot (\hat{\theta}_2^K - \theta_2^0)\right),\ K \geq K_0(\sigma+1)\right\}$ are uniformly integrable.

Remark 3.6 *Analysis of the proofs of Theorem 3.2 and Lemma 3.5 shows that the convergence in Theorem 3.4 is, in fact, uniform with respect to θ^0 on every compact set of the parameter space. It is shown in [8, Theorem III.1.3] that, under some additional conditions, this uniform convergence implies certain asymptotic efficiency of the estimate. The conditions in question do not hold for our general model (2.6), but do hold in the case of commuting operators [7].*

Theorem 3.7 *If θ_1^0 is known and order$(\mathcal{B}) = b < m - d/2$, then the measures generated in $\mathbf{C}([0,T];\mathbf{H}^s)$, $s < -d/2$, by the solutions of (3.15) are equivalent for all $\theta_2 \in \boldsymbol{R}$ and*

$$\mathbf{P} - \lim_{K \to \infty} \hat{\theta}_2^K = \theta_2^0 + \frac{\int_0^T (\mathcal{B}u(t), dW(t))_0}{\int_0^T \|\mathcal{B}u(t)\|_0^2 dt}. \tag{3.27}$$

Proof. By (2.9),

$$\mathbf{E}\int_0^T \|\mathcal{B}u(t)\|_0^2 dt < \infty \tag{3.28}$$

for all $\theta_2 \in \boldsymbol{R}$, and therefore the stochastic integral $\int_0^T (\mathcal{B}u(t), dW(t))_0$ is well defined [18, 19, 25]. Then (3.27) follows from (3.17) and the properties of the stochastic integral.

Next, denote by P^{θ_2} the measure generated in $\mathbf{C}([0,T];\mathbf{H}^s)$, $s < -d/2$, by the solution of (3.15) corresponding to the given value of θ_2. Inequality (3.28) implies that

$$\int_0^T \|\mathcal{B}u(t)\|_0^2 dt < \infty \quad (\mathbf{P}\text{- a.s.}) \tag{3.29}$$

and therefore by Corollary 1 in [18] the measures P^{θ_2} are equivalent for all $\theta_2 \in \boldsymbol{R}$ with the likelihood ratio

$$\begin{aligned} &\frac{dP^{\theta_2}}{dP^{\theta_2^0}}(u) = \\ &\exp\Big((\theta_2 - \theta_2^0)\int_0^T (\mathcal{B}u(t), dW(t))_0 - (1/2)(\theta_2 - \theta_2^0)^2 \int_0^T \|\mathcal{B}u(t)\|_0^2 dt\Big), \end{aligned} \tag{3.30}$$

where $u(t)$ is the solution of (3.15) corresponding to $\theta_2 = \theta_2^0$. Note that

$$\hat{\theta}_2 = \theta_2^0 + \frac{\int_0^T (\mathcal{B}u(t), dW(t))_0}{\int_0^T \|\mathcal{B}u(t)\|_0^2 dt} \tag{3.31}$$

maximizes the likelihood ration (3.30).

If the operators $\mathcal{A}$, $\mathcal{B}$, $\mathcal{N}$ have the same eigenfunctions as $\mathcal{L}$, then the coefficients $\psi_k(u(t))$ are independent (for different k) Ornstein-Uhlenbeck processes and $\Pi^K \mathcal{A}u(t) = \Pi^K \mathcal{A}\Pi^K u(t)$, with similar relations for $\mathcal{B}$ and $\mathcal{N}$. As a result, other properties of (3.16) and (3.17) can be established, including strong consistency and asymptotic efficiency [5, 7, 22], and, in the case of the continuous time observations, all estimates are computable explicitly in terms of $\psi_k(u(t)), k = 1, \ldots, K$.

In general, the computation of $\hat{\theta}_1^K$ and $\hat{\theta}_2^K$ using (3.16) and (3.17) respectively requires the knowledge of the whole field u rather than its projection. Still, the operators $\Pi^K(\mathcal{L} + \mathcal{A})$, $\Pi^K\mathcal{B}$, and $\Pi^K\mathcal{N}$ have finite dimensional range, which should make the computations feasible. Another option is to *replace* u by $\Pi^K u$. This can simplify the computations, but the result is, in some sense, even further from the maximum likelihood estimate, because some information is lost, and the asymptotic properties of the resulting estimate are more difficult to study. In general, the construction of the estimate depending *only* on the projection $\Pi^K u(t)$ is equivalent to the parameter estimation for a partially observed system with observations being given by (3.3). Without special assumptions on the operators $\mathcal{A}_0$ and $\mathcal{A}_1$, this problem is extremely difficult even in the finite dimensional setting.

4 An Example

Consider the following stochastic partial differential equation:

$$du(t,x) = (D\nabla^2 u(t,x) - (\vec{v}(x), \nabla)u(t,x) - \lambda u(t,x))dt + dW(t,x). \tag{4.1}$$

It is called the **heat balance equation** and describes the dynamics of the sea surface temperature anomalies [4]. In (4.1), $x = (x_1, x_2) \in \boldsymbol{R}^2$, $\vec{v}(x) = (v_1(x_1,x_2), v_2(x_1,x_2))$ is the velocity field of the top layer of the ocean (it is assumed to be known), D is thermodiffusivity, λ is the cooling coefficient. The equation is considered on a rectangle $|x_1| \leq a$; $|x_2| \leq c$ with periodic boundary conditions $u(t, -a, x_2) = u(t, a, x_2)$, $u(t, x_1, -c) = u(t, x_1, c)$ and zero initial condition. This reduces (4.1) to the general model (3.15) with M being a torus, $d = 2$, $\mathcal{L} = -\nabla^2 = -\partial^2/\partial x_1^2 - \partial^2/\partial x_2^2$, $\mathcal{A} = 0$, $\mathcal{B} = I$ (the identity operator), $\mathcal{N} = (\vec{v}, \nabla) = v_1(x_1,x_2)\partial/\partial x_1 + v_2(x_1,x_2)\partial/\partial x_2$, $\theta_1 = D$, $\theta_2 = \lambda$. Then $order(\mathcal{L}) = 2$ (so that $m = 1$), $order(\mathcal{A}) = 0$, $order(\mathcal{B}) = 0$ (so that $b = 0$), and $order(\mathcal{N}) = 1$. The basis $\{e_k\}_{k\geq 1}$ is the suitably ordered collection of real and imaginary parts of

$$g_{n_1,n_2}(x_1,x_2) = \frac{1}{\sqrt{4ac}} \exp\left\{\sqrt{-1}\pi(x_1 n_1/a + x_2 n_2/c)\right\}, \ n_1, n_2 \geq 0. \tag{4.2}$$

By Theorem 3.2, the estimate of D is consistent and asymptotically normal, the rate of convergence is $\Psi_{K,1} \asymp K$; the estimate of λ is also consistent and asymptotically normal with the rate of convergence $\Psi_{K,2} \asymp \sqrt{\ln K}$, since $b = 0 = m - d/2$ and (3.14) holds.

Unlike the case of the commuting operators, the proposed approach allows non-constant velocity field. Still, a significant limitation is that the value of $\vec{v}(x)$ must be known.

5 Proof of Theorem 3.2

Hereafter, $u(t)$ is the solution of (3.15) corresponding to the true value of the parameters (θ_1^0 and θ_2^0) and C is a generic constant with possibly different values in different places.

To prove the asymptotic normality of the estimate, the following version of the central limit theorem will be used. The proof can be found in [5].

Lemma 5.1 *If $\mathcal{P}$ is a differential operator on M and*

$$\mathbf{P}-\lim_{K\to\infty}\frac{\int_0^T\|\Pi^K\mathcal{P}u(t)\|_0^2dt}{\mathbf{E}\int_0^T\|\Pi^K\mathcal{P}u(t)\|_0^2dt}=1, \tag{5.1}$$

then

$$\lim_{K\to\infty}\frac{\int_0^T(\Pi^K\mathcal{P}u(t),dW^K(t))_0dt}{\sqrt{\mathbf{E}\int_0^T\|\Pi^K\mathcal{P}u(t)\|_0^2dt}}=\mathcal{N}(0,1) \tag{5.2}$$

in distribution.

Once (5.1) and (5.2) hold and

$$\lim_{K\to\infty}\mathbf{E}\int_0^T\|\Pi^K\mathcal{P}u(t)\|_0^2dt=+\infty, \tag{5.3}$$

the convergence

$$\mathbf{P}-\lim_{K\to\infty}\frac{\int_0^T(\Pi^K\mathcal{P}u(t),dW^K(t))_0dt}{\int_0^T\|\Pi^K\mathcal{P}u(t)\|_0^2dt}=0 \tag{5.4}$$

follows. Thus, it suffices to establish (5.1) and compute the asymptotics of $\mathbf{E}\displaystyle\int_0^T\|\Pi^K\mathcal{P}u(t)\|_0^2dt$ for a suitable operator $\mathcal{P}$.

If $\psi_k(t):=\psi_k(u(t))$, then (3.15) implies

$$d\psi_k(t)=-\theta_1^0l_k\psi_k(t)-\psi_k\Big((\theta_1^0\mathcal{A}+\theta_2^0\mathcal{B}+\mathcal{N})u(t)\Big)dt+dw_k(t),\ \psi_k(0)=\psi_k(u_0). \tag{5.5}$$

According to the variation of parameters formula, the solution of this equation is given by $\psi_k(t)=\xi_k(t)+\eta_k(t)$, where

$$\begin{aligned}
&\xi_k(t)=\int_0^t e^{-\theta_1^0l_k(t-s)}dw_k(s),\\
&\eta_k(t)=\psi_k(0)e^{-\theta_1^0l_kt}-\int_0^t e^{-\theta_1^0l_k(t-s)}\psi_k\Big((\theta_1^0\mathcal{A}+\theta_2\mathcal{B}+\mathcal{N})u(s)\Big)ds\\
&:=\eta_{0k}(t)+\eta_{1k}(t).
\end{aligned} \tag{5.6}$$

If $\xi(t)$ and $\eta(t)$ are the elements of $\cup_s\mathbf{H}^s$ defined by the sequences $\{\xi_k(t)\}_{k\geq1}$ and $\{\eta_k(t)\}_{k\geq1}$ respectively, then the solution of (3.15) can be written as $u(t)=\xi(t)+\eta(t)$.

The following technical result will be used in the future. The proof is given in Appendix.

Lemma 5.2 *If $a > 0$ and $f(t) \geq 0$, then*

$$\int_0^T \Big(\int_0^t e^{-a(t-s)} f(s)ds\Big)^2 dt \leq \frac{\int_0^T f^2(t)dt}{a^2}. \tag{5.7}$$

It is shown in the next lemma that, under certain conditions on the operator $\mathcal{P}$, the asymptotics of

$$\mathbf{E}\int_0^T \|\Pi^K \mathcal{P} u(t)\|_0^2 dt \tag{5.8}$$

is determined by the asymptotics of

$$\mathbf{E}\int_0^T \|\Pi^K \mathcal{P} \xi(t)\|_0^2 dt. \tag{5.9}$$

Lemma 5.3 *If $\mathcal{P}$ is an essentially non-degenerate operator of order p on M and $p \geq m - d/2$, then*

$$\mathbf{E}\int_0^T \|\Pi^K \mathcal{P} \xi(t)\|_0^2 dt \asymp \sum_{k=1}^N l_k^{(p-m)/m}, \ K \to \infty, \tag{5.10}$$

$$\lim_{K\to\infty} \frac{\mathbf{E}\int_0^T \|\Pi^K \mathcal{P} \eta(t)\|_0^2 dt}{\mathbf{E}\int_0^T \|\Pi^K \mathcal{P} \xi(t)\|_0^2 dt} = 0, \tag{5.11}$$

$$\mathbf{P} - \lim_{K\to\infty} \frac{\int_0^T \|\Pi^K \mathcal{P} \eta(t)\|_0^2 dt}{\mathbf{E}\int_0^T \|\Pi^K \mathcal{P} \xi(t)\|_0^2 dt} = 0, \tag{5.12}$$

$$\mathbf{P} - \lim_{K\to\infty} \frac{\int_0^T \|\Pi^K \mathcal{P} \xi(t)\|_0^2 dt}{\mathbf{E}\int_0^T \|\Pi^K \mathcal{P} \xi(t)\|_0^2 dt} = 1. \tag{5.13}$$

In the particular case $\mathcal{P} = \mathcal{L} + \mathcal{A}$, a stronger version of (5.10) *holds:*

$$\lim_{K\to\infty} \frac{\mathbf{E}\int_0^T \|\Pi^K \mathcal{P} \xi(t)\|_0^2 dt}{\frac{T}{2\theta_1^0}\sum_{k=1}^K l_k} = 1. \tag{5.14}$$

Proof.
Proof of (5.10) *and* (5.14). It follows from the independence of $\xi_k(t)$ for different k that

$$\begin{gathered}\mathbf{E}\sum_{k=1}^K |\psi_k(\mathcal{P}\xi(t))|^2 = \mathbf{E}\sum_{k=1}^K \Big|\sum_{n\geq 1} \xi_n(t)(e_n, \mathcal{P}^* e_k)_0\Big|^2 = \\ \sum_{k=1}^K \sum_{n\geq 1} \frac{1}{2\theta_1^0 l_n}(1 - e^{-2\theta_1^0 l_n t})|(e_n, \mathcal{P}^* e_k)_0|^2.\end{gathered} \tag{5.15}$$

Integration yields:

$$\mathbf{E}\int_0^T \|\Pi^K \mathcal{P} \xi(t)\|_0^2 dt = \sum_{k=1}^K \sum_{n\geq 1} \frac{1}{2\theta_1^0 l_n}\left(T - \frac{1}{2\theta_1^0 l_n}(1 - e^{-2\theta_1^0 l_n T})\right)|(e_n, \mathcal{P}^* e_k)_0|^2. \tag{5.16}$$

Since $l_k > 0$ and $\theta_1^0 > 0$, it follows that $1 - e^{-2\theta_1^0 l_k T} > 0$ for all k. Then the last inequality and the definition of the norm $\|\cdot\|_s$ imply

$$\begin{aligned} &\frac{T}{2\theta_1^0}\sum_{k=1}^K \|\mathcal{P}^* e_k\|_{-m}^2 - C\sum_{k=1}^K \|\mathcal{P}^* e_k\|_{-2m}^2 \le \mathbf{E}\int_0^T \|\Pi^K \mathcal{P}\xi(t)\|_0^2 dt \\ &\le \frac{T}{2\theta_1^0}\sum_{k=1}^K \|\mathcal{P}^* e_k\|_{-m}^2. \end{aligned} \tag{5.17}$$

Since $\mathcal{P}$ satisfies (3.14),

$$\|\mathcal{P}^* e_k\|_{-m}^2 \ge \varepsilon \|e_k\|_{p-m}^2 - K\|e_k\|_{p-m-\delta}^2 = \varepsilon \lambda_k^{2(p-m)}(1 - (K/\varepsilon)\lambda_k^{-2\delta}). \tag{5.18}$$

In addition, $\|\mathcal{P}^* e_k\|_r^2 \le C\|e_k\|_{r+p}^2$ and $\lambda_k = l_k^{1/(2m)}$. The result (5.10) follows.

To prove (5.14) note first of all that if $\mathcal{P} = \mathcal{L} + \mathcal{A}$, then the non-degeneracy condition (3.14) holds with $p = 2m,\ \varepsilon = 1,\ \delta = m - order(\mathcal{A})/2$, because

$$\|\mathcal{L}f\|_s = \|f\|_{s+2m}, \quad \|\mathcal{A}f\|_s \le C\|f\|_{s+2m-2\delta}, \tag{5.19}$$

and, since the order of the operator $\mathcal{A}^*\mathcal{L}$ is $4m - 2\delta$,

$$\begin{aligned} &(\mathcal{A}^*\mathcal{L}f, f)_s = (\Lambda^{-(2m-\delta)}\mathcal{A}^*\mathcal{L}f, \Lambda^{2m-\delta}f)_s \le \\ &\|\Lambda^{-(2m-\delta)}\mathcal{A}^*\mathcal{L}f\|_s\ \|\Lambda^{2m-\delta}f\|_s \le C\|f\|_{s+2m-\delta}^2. \end{aligned} \tag{5.20}$$

As a result, since $\|e_k\|_s^2 = l_k^{s/m}$, it follows that

$$l_k(1 - Cl_k^{-\delta/m}) \le \|\mathcal{P}^* e_k\|_{-m}^2 \le l_k(1 + Cl_k^{-\delta/m}), \tag{5.21}$$

and consequently (5.14) follows from (5.17) and (5.21).

Proof of (5.11). Consider first $\eta_0(t) = \{\eta_{0k}(t)\}$ (see (5.6)). With the notation $\gamma = 2(p-m)/d$,

$$\begin{aligned} &\mathbf{E}\int_0^T \|\Pi^N \mathcal{P}\eta_0(t)\|_0^2 dt \le C\sum_{k=1}^K \frac{1}{l_k}\mathbf{E}|\psi_k(\mathcal{P}u_0)|^2 \\ &\le C\sum_{k=1}^K k^{\gamma+1}\lambda_k^{-2(p+d/2)}\mathbf{E}|\psi_k(\mathcal{P}u_0)|^2. \end{aligned} \tag{5.22}$$

Note that

$$\sum_{k\ge 1}\lambda_k^{-2(p+d/2)}\mathbf{E}|\psi_k(\mathcal{P}u_0)|^2 \le C\mathbf{E}\|u_0\|_{-d/2}^2 < \infty. \tag{5.23}$$

If $\gamma = -1$, then

$$\lim_{K\to\infty}\frac{\mathbf{E}\int_0^T \|\Pi^K\mathcal{P}\eta_0(t)\|_0^2 dt}{\mathbf{E}\int_0^T \|\Pi^K\mathcal{P}\xi(t)\|_0^2 dt} \le \lim_{K\to\infty}\frac{C\mathbf{E}\|u_0\|_{-d/2}^2}{\ln K} = 0. \tag{5.24}$$

If $\gamma > -1$, then

$$\lim_{K\to\infty}\frac{\mathbf{E}\int_0^T \|\Pi^K\mathcal{P}\eta_0(t)\|_0^2 dt}{\mathbf{E}\int_0^T \|\Pi^K\mathcal{P}\xi(t)\|_0^2 dt} \le \lim_{K\to\infty}\frac{C\sum_{k=1}^K k^{\gamma+1}\lambda_k^{-2(p+d/2)}\mathbf{E}|\psi_k(\mathcal{P}u_0)|^2}{K^{\gamma+1}} = 0 \tag{5.25}$$

by (5.23) and the Kronecker lemma.

Next consider $\eta_1(t) = \{\eta_{1k}(t)\}$. By assumptions,

$$c := \max(order(\mathcal{A}), order(\mathcal{B}), order(\mathcal{N})) < 2m. \tag{5.26}$$

By Lemma 5.2,

$$\int_0^T |\eta_{1n}(t)|^2 dt \leq \frac{1}{(\theta_1^0 l_n)^2} \int_0^T |\psi_n\Big((\theta_1^0\mathcal{A} + \theta_2^0\mathcal{B} + \mathcal{N})u(t)\Big)|^2 dt, \tag{5.27}$$

which implies that, for every $r \in \boldsymbol{R}$,

$$\begin{aligned} &\sum_{n\geq 1} \lambda_n^{2r} \int_0^T |\psi_n(\mathcal{P}\eta_1(t))|^2 dt \equiv \int_0^T \|\mathcal{P}\eta_1(t)\|_r^2 dt \leq C \int_0^T \|\eta_1(t)\|_{r+p}^2 dt \equiv \\ &C \sum_n \lambda_n^{2(r+p)} \int_0^T |\eta_{1n}(t)|^2 dt \leq C \int_0^T \|u(t)\|_{r-2m+c+p}^2 dt. \end{aligned} \tag{5.28}$$

If $c_1 := 2m - c > 0$ and $r = -x$, where $x = \max(0, d/2 + c_1/2 + p + c - 3m)$, then $-x - 2m + c + p = m - d/2 - c_1/2$ and, by (2.8), $\mathbf{E}\int_0^T \|u(t)\|_{-x-2m+c+p}^2 < \infty$. As a result, since $\lambda_k \asymp k^{1/d}$,

$$\begin{aligned} &\frac{\mathbf{E}\int_0^T \|\Pi^K\mathcal{P}\eta_1(t)\|_0^2 dt}{\mathbf{E}\int_0^T \|\Pi^K\mathcal{P}\xi(t)\|_0^2 dt} = \frac{\sum_{n=1}^K \lambda_n^{-2x}\lambda_n^{2x}\mathbf{E}\int_0^T |\psi_n(\mathcal{P}\eta_1(t))|^2 dt}{\mathbf{E}\int_0^T \|\Pi^K\mathcal{P}\xi(t)\|_0^2 dt} \leq \\ &\frac{CK^{2x/d}\sum_{n\geq 1}\lambda_n^{-2x}\mathbf{E}\int_0^T |\psi_n(\mathcal{P}\eta_1(t))|^2 dt}{\mathbf{E}\int_0^T \|\Pi^K\mathcal{P}\xi(t)\|_0^2 dt} \leq \frac{CK^{2x/d}}{\sum_{k=1}^K \lambda_k^{2(p-m)}} \to 0 \text{ as } K \to \infty, \end{aligned} \tag{5.29}$$

because if $p - m = -d/2$, then $d/2 + c_1/2 + p + c - 3m = -c_1/2 < 0$ so that $x = 0$, while for $p - m > -d/2$ the sum $\sum_{k=1}^K \lambda_k^{2(p-m)}$ is of order $K^{2(p-m)/d+1}$ and $2(p-m)/d + 1 > (d + 2(p-m) - c_1/2) = 2x/d$. Equality (5.11) is proved. Then (5.12) follows from (5.11) and the Chebychev inequality.

Proof of (5.13). There are two steps in the proof. Writing $X_K(t) := \|\Pi^K\mathcal{P}\xi(t)\|_0^2$, the first step is to show that

$$var(X_K(t)) \leq C \sum_{k=1}^K \lambda_k^{4(p-m)} \tag{5.30}$$

for all $t \in [0, T]$. The second step is to show that (5.30) implies

$$\mathbf{P} - \lim_{K\to\infty} \frac{\int_0^T X_K(t)dt}{\mathbf{E}\int_0^T X_K(t)dt} = 1. \tag{5.31}$$

1). If $X_K^M(t) := \sum_{k=1}^K |\sum_{n=1}^M \xi_n(t)(e_n, \mathcal{P}^*e_k)_0|^2$, then $X_K^M(t)$ is a quadratic form of the Gaussian vector $(\xi_1(t), \ldots, \xi_M(t))$. The matrix of the quadratic form is $A = [A_{nn'}]_{n,n'=1,\ldots,M}$ with

$$A_{nn'} = \sum_{k=1}^K (e_n, \mathcal{P}^*e_k)_0(e_{n'}, \mathcal{P}^*e_k)_0, \tag{5.32}$$

and the covariance matrix of the Gaussian vector is

$$R = diag\left(\frac{1 - e^{-2\theta_1^0 l_n t}}{2\theta_1^0 l_k},\ n = 1, \ldots, M\right). \tag{5.33}$$

Direct computations yield

$$\mathbf{E}X_K^M(t) = \sum_{k=1}^{K}\sum_{n=1}^{M} \frac{1}{2\theta_1^0 l_n}(1-e^{-2\theta_1^0 l_n t})|(e_n, \mathcal{P}^* e_k)_0|^2 = trace(AR). \tag{5.34}$$

Analysis of the proof of (5.10) shows that, for every $t \in [0,T]$ and $k = 1,\ldots,K$, the series $\sum_{n\geq 1} \xi_n(t)(e_n, \mathcal{P}^* e_k)_0$ converges with probability one and in the mean square. Consequently,

$$\begin{aligned} &\lim_{M\to\infty} X_K^M(t) = X_K(t) \qquad (\mathbf{P}\text{- a.s.});\\ &\lim_{M\to\infty} \mathbf{E}X_K^M(t) = \sum_{k=1}^{K}\sum_{n\geq 1} \mathbf{E}|\xi_n(t)|^2\ |(e_n, \mathcal{P}^* e_k)_0|^2 = \mathbf{E}X_K(t). \end{aligned} \tag{5.35}$$

Next,

$$\begin{aligned} &var(X_K^M(t)) = 2trace((AR)^2) \leq C\sum_{n,n'} \frac{1}{l_n l_{n'}} A_{nn'}^2 = \\ &\sum_{k,k'=1}^{K} |(\tilde{\mathcal{P}} e_k, e_{k'})_0|^2 \lambda_k^{4(p-m)} \leq \sum_{k=1}^{K} \|\tilde{\mathcal{P}} e_k\|_0^2 \lambda_k^{4(p-m)} \leq C\sum_{k=1}^{K} \lambda_k^{4(p-m)}, \end{aligned} \tag{5.36}$$

where $\tilde{\mathcal{P}} := \mathcal{P}\Lambda^{-2m}\mathcal{P}^*\Lambda^{2(m-p)}$ is a bounded operator in $\mathbf{H}^0$. After that, inequality (5.30) follows from (5.35) and the Fatou lemma:

$$\begin{aligned} &var(X_K(t)) = \mathbf{E}\lim_{M\to\infty}|X_K^M(t)|^2 - |\mathbf{E}\lim_{M\to\infty} X_K^M(t)|^2 = \\ &\mathbf{E}\lim_{M\to\infty}|X_K^M(t)|^2 - \lim_{M\to\infty}|\mathbf{E}X_K^M(t)|^2 \leq \liminf_{M\to\infty}\mathbf{E}|X_K^M(t)|^2 - \lim_{M\to\infty}|\mathbf{E}X_K^M(t)|^2 \\ &\leq \liminf_{M\to\infty} var(X_K^M(t)) \leq C\sum_{k=1}^{K}\lambda_k^{4(p-m)}. \end{aligned} \tag{5.37}$$

2). If $Y_K := \int_0^T (X_K(t) - \mathbf{E}X_K(t))dt / \mathbf{E}\int_0^T X_K(t)dt$, then

$$\frac{\int_0^T X_K(t)dt}{\mathbf{E}\int_0^T X_K(t)dt} = 1 + Y_K \tag{5.38}$$

and

$$\mathbf{E}Y_K^2 \leq \frac{T\int_0^T (var(X_K(t))dt}{\left(\mathbf{E}\int_0^T X_K(t)dt\right)^2} \leq C\frac{\sum_{k=1}^K \lambda_k^{4(p-m)}}{\left(\sum_{k=1}^K \lambda_k^{2(p-m)}\right)^2} \to 0 \quad \text{as } K\to\infty. \tag{5.39}$$

By the Chebychev inequality, $\mathbf{P}-\lim_{K\to\infty} Y_K = 0$, which implies (5.13).

Corollary 5.4 *If $\mathcal{P}$ is an essentially non-degenerate operator of order p on M and $p \geq m - d/2$, then*

$$\mathbf{E}\int_0^T \|\Pi^K \mathcal{P}u(t)\|^2 dt \asymp \sum_{k=1}^{K} l_k^{(p-m)/m},\ K\to\infty, \tag{5.40}$$

and

$$\mathbf{P} - \lim_{K\to\infty} \frac{\int_0^T \|\Pi^K\mathcal{P}u(t)\|_0^2 dt}{\mathbf{E}\int_0^T \|\Pi^K\mathcal{P}u(t)\|_0^2 dt} = 1. \tag{5.41}$$

In the particular case $\mathcal{P} = \mathcal{L} + \mathcal{A}$, a stronger version of (5.40) *holds:*

$$\lim_{K\to\infty} \frac{\mathbf{E}\int_0^T \|\Pi^K \mathcal{P}u(t)\|_0^2 dt}{\frac{T}{2\theta_1^0}\sum_{k=1}^K l_k} = 1. \tag{5.42}$$

Proof. By the inequality $|2xy| \le \epsilon x^2 + \epsilon^{-1}y^2$, which holds for every $\epsilon > 0$ and every real x, y,

$$\begin{aligned}
&(1-\epsilon)\mathbf{E}\int_0^T \|\Pi^K\mathcal{P}\xi(t)\|_0^2 dt + (1-\frac{1}{\epsilon})\mathbf{E}\int_0^T \|\Pi^K\mathcal{P}\eta(t)\|_0^2 dt \le \\
&\mathbf{E}\int_0^T \|\Pi^K\mathcal{P}u(t)\|_0^2 dt \le \\
&(1+\epsilon)\mathbf{E}\int_0^T \|\Pi^K\mathcal{P}\xi(t)\|_0^2 dt + (1+\frac{1}{\epsilon})\mathbf{E}\int_0^T \|\Pi^K\mathcal{P}\eta(t)\|_0^2 dt.
\end{aligned} \tag{5.43}$$

Since ϵ is arbitrary, (5.40) follows from (5.11) and (5.10). After that, (5.41) follows from (5.13). Similarly, (5.42) follows from (5.11) and (5.14).

To prove the first part of Theorem 3.2, note that in this case $\mathcal{P} = \mathcal{L} + \mathcal{A}$, and it remains to use Lemma 5.1 and equations (5.41) and (5.42) from Corollary 5.4.

Similarly, the second part of the theorem follows with $\mathcal{P} = \mathcal{B}$; now (3.14) is assumed. Analysis of the proof shows that

$$\lim_{K\to\infty} \frac{\Psi_{K,2}^2}{\sum_{k=1}^K l_k^{(b-m)/m}} \ge \frac{\varepsilon T}{2\theta_1^0}. \tag{5.44}$$

6 Proof of Lemma 3.5

The following notation will be used:

$$\gamma = 2(p-m)/d \ge -1. \tag{6.1}$$

Since by Lemma 5.3

$$\Psi_K^2 \asymp \sum_{k=1}^K k^\gamma, \tag{6.2}$$

it follows that it is sufficient to prove the inequalities

$$\mathbf{E}\left|\int_0^T \left(\Pi^K\mathcal{P}u(t), dW^K(t)\right)_0\right|^q \le C\cdot\left(\sum_{k=1}^K k^\gamma\right)^{q/2} \tag{6.3}$$

and

$$\mathbf{E}\left(\int_0^T \|\Pi^K\mathcal{P}u(t)\|_0^2 dt\right)^{-q} \le C\cdot\left(\sum_{k=1}^K k^\gamma\right)^{-q} \tag{6.4}$$

for all $q > 0$ and all sufficiently large K. The numbers C in the above inequalities do not depend on K but can depend on everything else, including q and T.

By definition,

$$\int_0^T \Big(\Pi^K \mathcal{P}u(t), dW^K(t)\Big)_0 = \sum_{k=1}^K \int_0^T \psi_k\Big(\mathcal{P}u(t)\Big)dw_k(t), \tag{6.5}$$

$$\|\Pi^K \mathcal{P}u(t)\|^2 = \sum_{k=1}^K |\psi_k\Big(\mathcal{P}u(t)\Big)|^2, \tag{6.6}$$

and for each t the coefficients $\psi_k\Big(\mathcal{P}u(t)\Big)$ are Gaussian random variables. Indeed, denote by $P_t f$ the solution of the equation

$$\begin{aligned} &dv(t) + (\theta_1^0(\mathcal{L}+\mathcal{A}) + \theta_2^0\mathcal{B} + \mathcal{N})\,dt = 0,\ 0 < t \le T;\\ &v(0) = f, \end{aligned} \tag{6.7}$$

The solution of (3.15) can then be written as

$$u(t) = P_t u_0 + \int_0^t P_{t-s} dW(s) := u_1(t) + u_2(t), \tag{6.8}$$

and the properties of the stochastic integral [23, Chapter 2] imply that $\psi_k\Big(\mathcal{P}u_2(t)\Big)$ are Gaussian random variables with zero mean and covariance

$$\mathbf{E}\psi_k\Big(\mathcal{P}u_2(t)\Big)\psi_m\Big(\mathcal{P}u_2(t)\Big) = \int_0^t (P_s^*\mathcal{P}^*e_k, P_s^*\mathcal{P}^*e_m)_0\, ds := A_{km}(t). \tag{6.9}$$

Remark 6.1 *For integers K_0 and $K > K_0$, denote by $a_k(K_0, K; t)$, $1 \le k \le K - K_0 + 1$, the eigenvalues of the matrix $[A_{km}(t),\ K_0 \le k, m \le K]$. If ζ_k are independent standard Gaussian random variables, then the random variable $\sum_{k=K_0}^K |\psi_k\Big(\mathcal{P}u_2(t)\Big)|^2$ has the same distribution as $\sum_{k=1}^{K-K_0+1} a_k(K_0, K; t)\zeta_k^2$. This follows from the general properties of Gaussian random vectors.*

Proof of (6.3). With no loss of generality, it will be assumed that $q = 2n$ is an even integer. By the Burkholder-Davis-Gandy inequality [11, Theorem IV.4.1],

$$\begin{aligned} &\mathbf{E}\left|\sum_{k=1}^K \int_0^T \psi_k\Big(\mathcal{P}u(t)\Big)dw_k(t)\right|^{2n} \le C\mathbf{E}\left(\int_0^T \sum_{k=1}^K |\psi_k\Big(\mathcal{P}u(t)\Big)|^2 dt\right)^n\\ &\le C\cdot\left(\mathbf{E}\left(\int_0^T \sum_{k=1}^K |\psi_k\Big(\mathcal{P}u_1(t)\Big)|^2 dt\right)^n + \mathbf{E}\left(\int_0^T \sum_{k=1}^K |\psi_k\Big(\mathcal{P}u_2(t)\Big)|^2 dt\right)^n\right). \end{aligned} \tag{6.10}$$

The properties of the operator P_t imply that

$$\begin{aligned} &\mathbf{E}\left(\int_0^T \sum_{k=1}^K |\psi_k\Big(\mathcal{P}u_1(t)\Big)|^2 dt\right)^n \le CK^{n(\gamma+1)}\mathbf{E}\left(\int_0^T \|P_t\mathcal{P}u_0\|^2_{m-p-d/2}dt\right)^n\\ &C\cdot K^{n(\gamma+1)}\mathbf{E}\|u_0\|^q_{-d/2} \le C\cdot\left(\sum_{k=1}^K k^\gamma\right)^{q/2}. \end{aligned} \tag{6.11}$$

Next, by the Hölder inequality,

$$\mathbf{E}\left(\int_0^T \sum_{k=1}^K |\psi_k\Big(\mathcal{P}u_2(t)\Big)|^2 dt\right)^n \leq C\int_0^T \mathbf{E}\left(\sum_{k=1}^K |\psi_k\Big(\mathcal{P}u_2(t)\Big)|^2\right)^n. \tag{6.12}$$

By Remark 6.1 and the multinomial expansion formula,

$$\begin{aligned}
&\mathbf{E}\left(\sum_{k=1}^K |\psi_k\Big(\mathcal{P}u_2(t)\Big)|^2\right)^n = \mathbf{E}\left(\sum_{k=1}^K a_k(1,K;t)\zeta_k^2\right)^n \\
&= \sum_{m_1+\cdots+m_K=n} \frac{n!}{m_1!\cdots m_K!} a_1^{m_1}(1,K;t)\cdots a_K^{m_K}(1,K;t)\mathbf{E}\zeta_1^{2m_1}\cdots\zeta_K^{2m_K} \\
&\leq (2n-1)!!\left(\sum_{k=1}^K a_k(1,K;t)\right)^n = (2n-1)!!\left(\sum_{k=1}^K \int_0^t \|P_s^*\mathcal{P}^* e_k\|_0^2 ds\right)^n \\
&\leq C\cdot\left(\sum_{k=1}^K \|e_k\|_{p-m}^2\right)^{q/2},
\end{aligned} \tag{6.13}$$

where the last inequality is a consequence of (A.4). Since $\|e_k\|_{p-m}^2 = \lambda_k^{2(p-m)} \asymp k^\gamma$, inequality (6.3) follows.

Proof of (6.4). Note first of all that the Jensen inequality implies

$$\begin{aligned}
&\mathbf{E}\left(\int_0^T \sum_{k=1}^K |\psi_k\Big(\mathcal{P}u(t)\Big)|^2 dt\right)^{-q} \leq \mathbf{E}\left(\int_{T/2}^T \sum_{k=K_0}^K |\psi_k\Big(\mathcal{P}u(t)\Big)|^2 dt\right)^{-q} \\
&\leq C\int_{T/2}^T \mathbf{E}\left(\sum_{k=K_0}^K |\psi_k\Big(\mathcal{P}u(t)\Big)|^2\right)^{-q} dt \\
&= \int_{T/2}^T \mathbf{E}\left(\sum_{k=K_0}^K |\psi_k\Big(\mathcal{P}u_1(t)\Big) + \psi_k\Big(\mathcal{P}u_2(t)\Big)|^2\right)^{-q} dt,
\end{aligned} \tag{6.14}$$

and then, in view of Lemma A.2, it is sufficient to consider the case $u_0 = 0$.

According to Remark 6.1, if $u_0 = 0$, then inequality (6.4) will follow from

$$\mathbf{E}\left(\sum_{k=1}^{K-K_0+1} a_k(K_0,K;t)\zeta_k^2\right)^{-q} \leq C\cdot(F_\gamma(K))^{-q}, \quad T/2 \leq t \leq T, \tag{6.15}$$

where

$$F_\gamma(K) = \begin{cases} \ln K, & \text{if } \gamma = -1 \\ K^{1+\gamma}, & \text{if } \gamma > -1. \end{cases} \tag{6.16}$$

Assume for the moment that, when ordered appropriately, the numbers $a_k(K_0,K;t)$ have the following property: there exist an integer K_0 and a real number $C > 0$ so that, for all $K > K_0$, $1 \leq k \leq K - K_0 + 1$, and $T/2 \leq t \leq T$,

$$a_k(K_0,K;t) \geq C\cdot(k+K_0)^\gamma. \tag{6.17}$$

If (6.17) holds, then for all sufficiently large K

$$\mathbf{E}\left(\sum_{k=1}^{K-K_0+1} a_k(K_0,K;t)\zeta_k^2\right)^{-q} \leq C\mathbf{E}\left(\sum_{k=1}^{K/2} k^\gamma\zeta_k^2\right)^{-q}, \tag{6.18}$$

and it remains to estimate the right hand side of the last inequality.

Since for every non-negative random variable ζ and every $q > 0$

$$\mathbf{E}\zeta^{-q} = \frac{1}{\Gamma(q)} \int_0^\infty t^{q-1} \mathbf{E} e^{-\zeta t} dt, \quad \Gamma(\cdot) \text{ is the Gamma function,} \tag{6.19}$$

it follows that

$$\begin{aligned} \mathbf{E} \left(\sum_{k=1}^K k^\gamma \zeta_k^2 \right)^{-q} &\le C \int_0^\infty t^{q-1} \prod_{k=1}^K \frac{1}{\sqrt{1 + 2tk^\gamma}} dt \\ &= C \int_0^\infty t^{q-1} \exp \left(-\frac{1}{2} \sum_{k=1}^K \ln(1 + 2tk^\gamma) \right) dt. \end{aligned} \tag{6.20}$$

If $\gamma = -1$, then

$$\begin{aligned} &\sum_{k=1}^K \ln(1 + 2t/k) \ge \sum_{1<l<4q+1} \quad \sum_{K^{l/(4q+1)}<k<K^{(l+1)/(4q+1)}} \ln(1 + 2t/k) \\ &\ge \sum_{1<l<4q+1} \ln \left(1 + 2t \sum_{K^{l/(4q+1)}<k<K^{(l+1)/(4q+1)}} 1/k \right) \ge 4q \ln(c_1 + c_2 t \ln K) \end{aligned} \tag{6.21}$$

so that

$$\mathbf{E} \left(\sum_{k=1}^K k^\gamma \zeta_k^2 \right)^{-q} \le C \int_0^\infty \frac{t^{q-1}}{(c_1 + c_2 t \ln K)^{2q}} dt \le C (\ln K)^{-q}. \tag{6.22}$$

If $\gamma > -1$, then

$$\begin{aligned} &\sum_{k=1}^K \ln(1 + 2tk^\gamma) \ge \sum_{1<l<4q+1} \quad \sum_{\frac{Kl}{4q+1}<k<\frac{K(l+1)}{4q+1}} \ln(1 + 2tk^\gamma) \\ &\ge \sum_{1<l<4q+1} \ln \left(1 + 2t \sum_{\frac{Kl}{4q+1}<k<\frac{K(l+1)}{4q+1}} k^\gamma \right) \ge 4q \ln(1 + CtK^{\gamma+1}) \end{aligned} \tag{6.23}$$

so that

$$\mathbf{E} \left(\sum_{k=1}^K k^\gamma \zeta_k^2 \right)^{-q} dt \le C_1 \int_0^\infty \frac{t^{q-1}}{(1 + C_2 t K^{\gamma+1})^{2q}} dt \le C \cdot (K^{\gamma+1})^{-q}. \tag{6.24}$$

To complete the proof of the lemma, it remains to verify (6.17). Direct computations show that if $y_k,\ K_0 \le k \le K$, are real numbers and $T/2 \le t \le T$, then

$$\begin{aligned} \sum_{k,m=K_0}^K A_{km}(t) y_k y_m &= \int_0^t \| \sum_{k=K_0}^K P_s^* \mathcal{P}^* y_k e_k \|_0^2 ds \\ &\ge C_1 t \| \sum_{k=K_0}^K e_k y_k \|_{p-m}^2 - C_2 \| \sum_{k=K_0}^K e_k y_k \|_{p-m-\delta_0}^2 \\ &\ge \sum_{k=K_0}^K y_k^2 k^\gamma (C_1 - C_2 k^{-\delta_0}), \end{aligned} \tag{6.25}$$

where $\delta_0 = \min(\delta, 2m - order(\mathcal{A}+\mathcal{B}+\mathcal{N})) > 0$ with δ from (3.14), and the first inequality follows from (A.5) in Appendix and essential non-degeneracy of $\mathcal{P}$. If K_0 is chosen so that $C_1 - C_2 K_0^{-\delta_0} \geq C_1/2$, then there exists $C > 0$ for which the matrix

$$[A_{km}(t) - Ck^{\gamma}\delta_{km}, K_0 \leq k, m \leq K] \tag{6.26}$$

is non-negative definite, and then (6.17) follows from Theorem 13.5.4 in [20].

Acknowledgments

The work of the first author was partially supported by ONR Grant #N00014-95-1-0229 and by the NSF through the Institute for Mathematics and its Applications. The work of the second author was partially supported by ONR Grant #N00014-95-1-0229 and ARO Grant DAAH 04-95-1-0164.

Appendix

Proof of Lemma 5.2. Note that

$$\Big(\int_0^t e^{as} f(s)ds\Big)^2 = 2\int_0^t \int_0^s e^{as} e^{au} f(u) f(s) du ds. \tag{A.1}$$

If $U := \int_0^T \Big(\int_0^t e^{-a(t-s)} f(s)ds\Big)^2 dt$, then direct computations yield:

$$\begin{aligned} U = 2\int_0^T \int_0^t \int_0^s e^{-a(2t-s-u)} f(u) f(s) du ds dt \leq \\ a^{-1}\int_0^T \Big(\int_0^s e^{-a(s-u)} f(u) du\Big) f(s) ds \leq \\ a^{-1}\Big(\int_0^T f^2(s) ds\Big)^{1/2}\Big(\int_0^T \Big(\int_0^s e^{-a(s-u)} f(u) du\Big)^2 ds\Big)^{1/2}, \end{aligned} \tag{A.2}$$

and the result follows.

Lemma A.1 *Assume that $\mathcal{A}$ is an order $a < 2m$ differential operator on M. Denote by $P_t f$, $f \in C^{\infty}(M)$, the solution of the equation*

$$du(t) + (\mathcal{L}+\mathcal{A})u(t)dt = 0, \; 0 < t \leq T, \quad u(0) = f. \tag{A.3}$$

Then

$$\begin{aligned} &\int_0^T \|P_t f\|_{r+m}^2 dt \leq C(r,T)\|f\|_r^2, \quad \int_0^T \|\int_0^t P_{t-s} g(s) ds\|_{r+2m}^2 dt \\ &\leq C(r,T)\int_0^T \|g(s)\|_r^2 ds, \end{aligned} \tag{A.4}$$

and, as long as $T/2 \leq t \leq T$,

$$\int_0^t \|P_s f\|_{r+m}^2 ds \geq C_1(T)\|f\|_r^2 - C_2(r,T)\|f\|_{r+a-2m}^2. \tag{A.5}$$

Proof. Both inequalities in (A.4) follow from Theorem 3.1.4 in [23]. To prove (A.5), denote $P_t f$ by $V(t)$. By uniqueness, $V(t) = U(t)$, where $U = U(t)$ satisfies

$$dU(t) + (\mathcal{L}U(t) + \mathcal{A}V(t))dt = 0, \quad 0 < t \le T, \quad U(0) = f. \tag{A.6}$$

Denote by $\tilde{P}_t$ the semi-group generated by $-\mathcal{L}$. Then

$$U(t) = \tilde{P}_t f + \int_0^t \tilde{P}_{t-s}\mathcal{A}V(s)ds \tag{A.7}$$

and

$$\int_0^t \|U(s)\|^2_{r+m} \ge \frac{1}{2}\int_0^t \|\tilde{P}_s f\|^2_{r+m} ds - 2\int_0^t \|\int_0^s \tilde{P}_{s-\tau}\mathcal{A}V(\tau)d\tau\|^2_{r+m} ds. \tag{A.8}$$

Since for $T/2 \le t \le T$

$$\int_0^t \|\tilde{P}_s f\|^2_{r+m} ds = \int_0^t \sum_{k\ge 1} e^{-\lambda_k^{2m}s}|\psi_k(f)|^2 \lambda_k^{2(r+m)} ds \ge C(T)\|f\|^2_r \tag{A.9}$$

and

$$\begin{aligned}\int_0^t \|\int_0^s \tilde{P}_{s-\tau}\mathcal{A}V(\tau)d\tau\|^2_{r+m} ds &\le C(r,T)\int_0^T \|\mathcal{A}U(t)\|^2_{r-m} dt \\ \le C(r,T)\int_0^T \|P_t f\|^2_{r+a-m} dt &\le C(r,T)\|f\|^2_{r+a-2m},\end{aligned} \tag{A.10}$$

the result follows. Note that since $\mathcal{L}$ is self-adjoint, inequalities (A.4) and (A.5) hold if the operator P_s is replaced by its adjoint P_s^*.

Lemma A.2 *Assume that the components of the vector $\xi = \{\xi_k,\ k = 1, \dots, N\}$ are independent Gaussian random variables with zero mean and variance a_k, the vector $\eta = \{\eta_k,\ k = 1, \dots, N\}$ is independent of ξ, $q > 0$ is a real number, and $U \in \boldsymbol{R}^{N\times N}$ is an orthogonal matrix. Then*

$$\mathbf{E}\left(\sum_{k=1}^N ((U\xi)_k + \eta_k)^2\right)^{-q} \le \mathbf{E}\left(\sum_{k=1}^N |\xi_k|^2\right)^{-q}. \tag{A.11}$$

Proof. Denote by $\mathbf{E}'$ the conditional expectation given the σ-algebra generated by $\{\eta_k,\ k = 1, \dots, N\}$. Then

$$\begin{aligned}
&\mathbf{E}'\exp\left(-t\sum_{k=1}^N ((U\xi)_k + \eta_k)^2\right) = \\
&\prod_{k=1}^N \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left(-(1+2ta_k)x^2/2 - 2ta_k x(U^*\eta)_k - t\eta_k^2\right)dx \\
&= \exp\left(-t\sum_{k=1}^N (\eta_k^2 - \frac{2ta_k}{1+2ta_k}|(U^*\eta)_k|^2)\right) \\
&\times \prod_{k=1}^N \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left(-(1+2ta_k)(x - \frac{2t(U^*\eta)_k}{1+2ta_k})^2/2\right)dx \\
&\le \mathbf{E}\exp\left(-t\sum_{k=1}^N |\xi_k|^2\right),
\end{aligned} \tag{A.12}$$

and it remains to take the expectation $\mathbf{E}$ and use the relation

$$\mathbf{E}\zeta^{-q} = \frac{1}{\Gamma(q)}\int_0^\infty t^{q-1}\mathbf{E}e^{-\zeta t}dt, \quad \Gamma(\cdot) \text{ is the Gamma function.} \tag{A.13}$$

The lemma is proved.

References

[1] S. I. Aihara. Regularized Maximum Likelihood Estimate for an Infinite Dimensional Parameter in Stochastic Parabolic Systems. *SIAM Journal on Control and Optimization*, 30(4):745–764, 1992.

[2] A. Bagchi and V. Borkar. Parameter Identification in Infinite Dimensional Linear Systems. *Stochastics*, 12:201–213, 1984.

[3] H. T. Banks and K. Kunisch. *Estimation Techniques for Distributed Parameter Systems.* Birkhäuser, Boston, 1989.

[4] C. Frankignoul. SST Anomalies, Planetary Waves and RC in the Middle Rectitudes. *Reviews of Geophysics*, 23(4):357–390, 1985.

[5] M. Huebner. *Parameter Estimation for SPDE's.* PhD thesis, University of Southern California, Los Angeles, CA, 90089, June 1993.

[6] M. Huebner, R. Khasminskii, and B. L. Rozovskii. Two Examples of Parameter Estimation. In Cambanis, Ghosh, Karandikar, and Sen, editors, *Stochastic Processes*, Springer, New York, 1992.

[7] M. Huebner and B. Rozovskii. On Asymptotic Properties of Maximum Likelihood Estimators for Parabolic Stochastic PDE's. *Probability Theory and Related Fields*, 103:143–163, 1995.

[8] I. A. Ibragimov and R. Z. Khasminskii. *Statistical Estimation (Asymptotic Theory).* Springer, New York, 1981.

[9] I. A. Ibragimov and R. Z. Khasminskii. Some Estimation Problems in Infinite Dimensional White Noise. In D.Pollard, E.Torgersen, and G.Yang, editors, *Festschrift for Lucien Le Cam*, pages 259–274. Springer, Berlin, 1997.

[10] I. A. Ibragimov and R. Z. Khasminskii. Some Nonparametric Estimation Problems for Parabolic SPDE. Technical Report #31, Wayne State University, Department of Mathematics, 1997.

[11] N. V. Krylov. *Introduction to the Theory of Diffusion Processes.* American Mathematical Society, Providence, RI, 1995.

[12] H. Kumano-go. *Pseudo-Differential Operators.* The MIT Press, Cambridge, Massachusetts, 1981.

[13] Yu. A. Kutoyants. *Identification of Dynamical Systems With Small Noise.* Kluwer Academic Publisher, Dordrecht, 1994.

[14] L. LeCam. *Asymptotic Methods in Statistical Decision Theory.* Springer, New York, 1986.

[15] R. Sh. Liptser and A. N. Shiryayev. *Statistics of Random Processes.* Springer, New York, 1992.

[16] S. Lototsky and B. L. Rozovskii. Parameter Estimation for Stochastic Evolution Equations with Non-commuting Operators. In *Volume in honor of A. V. Skorokhod*. To appear.

[17] S. V. Lototsky. *Problems in Statistics of Stochastic Differential Equations.* PhD thesis, University of Southern California, Los Angeles, CA 90089, August 1996.

[18] R. Mikulevicius and B.L. Rozovskii. Uniqueness and Absolute Continuity of Weak Solutions for Parabolic SPDE's. *Acta Applicandae Mathematicae*, 35:179–192, 1994.

[19] R. Mikulevicius and B.L. Rozovskii. Martingale problems for nonlinear SPDE's. In R. Carmona and B. L. Rozovskii, editors, *Stochastic Partial Differential Equations. Six Perspectives.* AMS, Providence, RI, 1998. To appear.

[20] W. Nef. *Linear Algebra.* McGraw–Hill, New York, 1967.

[21] L. Piterbarg and B. Rozovskii. Maximum Likelihood Estimators in the Equations of Physical Oceanography. In R. J. Adler et al., editor, *Stochastic Modeling in Physical Oceanography*, pages 397–421. Birkhäuser, Boston, 1996.

[22] L. Piterbarg and B. Rozovskii. On Asymptotic Problems of Parameter Estimation in Stochastic PDE's: Discrete Time Sampling. *Mathematical Methods of Statistics*, 6(2):200–223, 1997.

[23] B. L. Rozovskii. *Stochastic Evolution Systems.* Kluwer Academic Publishers, 1990.

[24] M.A. Shubin. *Pseudodifferential Operators and Spectral Theory.* Springer, New York, 1987.

[25] J. B. Walsh. An Introduction to Stochastic Partial Differential Equations. In P. L. Hennequin, editor, *Ecole d'été de Probabilités de Saint-Flour, XIV*, volume 1180, pages 265–439, Springer, Berlin, 1984.