

## STATISTICAL INFERENCE FOR STOCHASTIC PARABOLIC EQUATIONS: A SPECTRAL APPROACH

S. V. LOTOTSKY

*Abstract*

---

A parameter estimation problem is considered for a stochastic parabolic equation driven by additive Gaussian noise that is white in time and space. The estimator is of spectral type and utilizes a finite number of the spatial Fourier coefficients of the solution. The asymptotic properties of the estimator are studied as the number of the Fourier coefficients increases, while the observation time and the noise intensity are fixed. A necessary and sufficient condition for consistency and asymptotic normality of the estimator is derived in terms of the eigenvalues of the operators in the equation, and a detailed proof is provided. Other estimation problems are briefly surveyed.

---

### 1. Introduction

**1.1. Motivation: The one-dimensional stochastic heat equation.** Consider the following stochastic equation

$$(1.1) \quad du(t, x) = \theta u_{xx}(t, x) dt + dW(t, x), \quad 0 < t \leq T, \quad x \in (0, \pi),$$

with zero initial and boundary conditions, where  $\theta > 0$  is an unknown real number and  $dW(t, x)$  is the noise term. With precise definitions to come later, at this point we interpret  $dW$  as a formal sum

$$dW(t, x) = \sum_{k \geq 1} h_k(x) dw_k(t),$$

where  $h_k(x) = \sqrt{2/\pi} \sin(kx)$ ,  $k \geq 1$ , and  $w_k$  are independent standard Brownian motions. Let us look for the solution of (1.1) as a Fourier

---

2000 *Mathematics Subject Classification*. Primary: 60H15, 62F12; Secondary: 60G15, 60G30, 62M05.

*Key words*. Cylindrical Brownian motion, Ornstein-Uhlenbeck process, singular statistical models.

series

$$u(t, x) = \sum_{k \geq 1} u_k(t) h_k(x).$$

Substitution of this series in (1.1) suggests that each  $u_k$  should satisfy

$$(1.2) \quad du_k(t) = -k^2 \theta u_k(t) dt + dw_k(t), \quad 0 < t \leq T,$$

with initial condition  $u_k(0) = 0$ . If the trajectory of  $u_k(t)$  is observed for one fixed  $k$  and all  $0 < t < T$ , then the maximum likelihood estimator of  $\theta$  based on this observation is

$$(1.3) \quad \hat{\theta}^{(k)} = -\frac{\int_0^T k^2 u_k(t) du_k(t)}{\int_0^T k^4 u_k^2(t) dt} = -\frac{u_k^2(T) - T}{2 \int_0^T k^2 u_k^2(t) dt};$$

see, for example, Liptser and Shiryaev [47, Formulas 17.25 and 17.45]. It is known [47, Theorem 17.4] that this estimator is consistent in the limit  $T \rightarrow +\infty$ :

$$\lim_{T \rightarrow +\infty} \frac{-\int_0^T k^2 u_k(t) du_k(t)}{\int_0^T k^4 u_k^2(t) dt} = \theta$$

with probability one.

Let us now assume that the trajectories of  $u_k(t)$  are observed for all  $0 < t < T$  and all  $k = 1, \dots, N$ , and let us combine the estimators (1.3) for different  $k$  as follows:

$$(1.4) \quad \hat{\theta}_N = -\frac{\sum_{k=1}^N \int_0^T k^2 u_k(t) du_k(t)}{\sum_{k=1}^N \int_0^T k^4 u_k^2(t) dt}.$$

First suggested by Huebner, Khas'minskiĭ, and Rozovskiĭ in [26], (1.4) is, in fact, the maximum likelihood estimator of  $\theta$  based on the observations  $u_k(t)$ ,  $k = 1, \dots, N$ ,  $0 < t < T$ .

It follows from (1.2) and (1.4) that

$$(1.5) \quad \hat{\theta}_N - \theta = -\frac{\sum_{k=1}^N \int_0^T k^2 u_k(t) dw_k(t)}{\sum_{k=1}^N \int_0^T k^4 u_k^2(t) dt}.$$

Note that both the top and the bottom of the fraction on the right-hand side of (1.5) are sums of independent random variables, and the analysis of the properties of the estimator  $\hat{\theta}_N$  is thus reduced to the study of these sums.

By direct computation,

$$\int_0^T \mathbb{E} u_k^2(t) dt = \frac{1}{2k^2} \int_0^T (1 - e^{-2k^2 t}) dt.$$

Consequently, as  $N \rightarrow \infty$ ,

$$\sum_{k=1}^N \int_0^T k^4 \mathbb{E} u_k^2(t) dt \sim \frac{T}{2} \sum_{k=1}^N k^2 \sim \frac{N^3 T}{6},$$

where notation  $a_N \sim b_N$  means  $\lim_{N \rightarrow \infty} (a_N/b_N) = 1$ . Since  $\mathbb{E} \int_0^T k^2 u_k(t) dw_k(t) = 0$ , it is reasonable to conjecture that,

- by the law of large numbers,  $\lim_{N \rightarrow \infty} (\hat{\theta}_N - \theta) = 0$  with probability one;
- by the central limit theorem, the sequence of random variables  $\{N^{3/2}(\hat{\theta}_N - \theta), N \geq 1\}$  converges in distribution to a zero-mean Gaussian random variable.

It is also clear that the proof of these conjectures will require a closer look at the Ornstein-Uhlenbeck processes (1.2) (see Section 2.1).

In the rest of the introduction, we discuss how (1.1) fits in the general framework of statistical estimation.

**1.2. Statistical estimation.** In many models, the general form of the equation is given by the basic laws governing the underlying process, while the particular features of the equation, such as coefficients, initial or boundary conditions, etc., must be determined from the observations of the process. This *model validation* is often accomplished with the help of statistical estimation.

Stochastic parabolic equations are used in various economical and physical models, such as the term structure of interest rates for bonds with different maturities (Aihara and Bagchi [8], [9], Cont [16]), the temperature of the top layer of the ocean (Frankignoul [20], Piterbarg and Rozovskii [57]), evolution of the population in time and space (Dawson [17], De [18]), spread of pollutants (Serrano and Adomian [70], Serrano and Unny [71]), etc. Equations of the type (1.1) provide a useful toy model for understanding the possible effect of the infinite number of dimensions and for deriving the *bench-mark* results about the corresponding estimators. Diagonalizable stochastic parabolic equations of the type discussed below can also model statistical problems in which information is coming from many independent, but not identical, channels (Korostelev and Yin [38]).

In the classical statistical estimation problem, the starting point is a family  $\mathbf{P}^\theta$  of probability measures depending on the parameter  $\theta \in \Theta \subset \mathbb{R}$ . Each  $\mathbf{P}^\theta$  is the distribution of a random element. It is assumed that a realization of one random element corresponding to one particular value

of  $\theta$  of the parameter is observed, and the objective is to estimate the value of this parameter from the observations.

One approach is to select the value  $\theta$  corresponding to the random element that is *most likely* to produce the observations. More precisely, we say that the statistical model (or estimation problem)  $\mathbf{P}^\theta$ ,  $\theta \in \Theta$ , is absolutely continuous if there exists a probability measure  $\mathbf{Q}$  such that all measures  $\mathbf{P}^\theta$  are absolutely continuous with respect to  $\mathbf{Q}$ . Then the **maximum likelihood estimator**  $\hat{\theta}$  of the unknown parameter is constructed by maximizing with respect to  $\theta$  the density  $d\mathbf{P}^\theta/d\mathbf{Q}$ . As a rule,  $\hat{\theta} \neq \theta$ , but one can hope that  $\hat{\theta}$  approaches  $\theta$  as more and more information becomes available. The amount of information can be increased in one of two ways: (a) increasing the sample size, for example, the observation time interval (large sample asymptotic); (b) reducing the amplitude of noise (small noise asymptotic).

If the measures  $\mathbf{P}^\theta$  are mutually singular for different values of  $\theta$ , then the model is called singular, and the value of the parameter can often be determined “exactly”. In reality, a singular model is usually approximated by a sequence of absolutely continuous models, and the parameter is then computed as the limit of the corresponding maximum likelihood estimators. For parabolic equations driven by additive space-time white noise, this approach was first suggested by Huebner, Khas'minskiĭ, and Rozovskiĭ [26], and was further investigated by Huebner and Rozovskiĭ [30], where a necessary and sufficient condition for the convergence of the estimators was stated in terms of the orders of the operators in the equation.

When the observations are finite-dimensional diffusions, the necessary and sufficient conditions for absolute continuity of the corresponding measures are well-known (see, for example, Liptser and Shiryaev [46, Chapter 7]). Many of the results have been extended to infinite dimensions by Kozlov [40], [41], Loges [39], [48], Mikulevičius and Rozovskiĭ [54], [55] and others. For linear equations, such as (1.1), whose solutions are Gaussian processes, there is another useful result, originally discovered independently by Feldman [19] and Hájek [21], [22]: Two Gaussian measures are either mutually absolutely continuous or mutually singular. In particular, we will see later (Theorem 4.8) that the measures generated by the solutions of (1.1) in a suitable Hilbert space are mutually singular for different values of  $\theta$ , and this singularity allows us to get the exact value of the parameter  $\theta$ , corresponding to the

observations, as

$$\theta = - \frac{\lim_{N \rightarrow \infty} \sum_{k=1}^N \int_0^T k^2 u_k(t) du_k(t)}{\lim_{N \rightarrow \infty} \sum_{k=1}^N \int_0^T k^4 u_k^2(t) dt}.$$

Since both limits are infinite, the expression must be approximated by  $\widehat{\theta}_N$  from (1.4). The situation is somewhat similar to the problem of estimating the diffusion coefficient in a finite-dimensional diffusion, where the exact value is known from the quadratic variation but is computed approximately using time discretization (see (2.27) and (2.28) below).

Here is the main result of the paper. Let  $\{h_k, k \geq 1\}$  be an orthonormal basis in a Hilbert space  $H$  and let  $W(t) = \sum_{k \geq 1} w_k(t) h_k$  be a cylindrical Brownian motion on  $H$ . Consider the linear stochastic parabolic equation

$$(1.6) \quad du + (A_0 + \theta A_1)u dt = dW(t), \quad 0 < t < T, \quad u(0) = 0.$$

Assume that the operators  $A_0$  and  $A_1$  have a common system of eigenfunctions:

$$A_0 h_k = \rho_k h_k, \quad A_1 h_k = \nu_k h_k.$$

Define

$$du_k(t) = -(\rho_k + \theta \nu_k)u_k(t) dt + dw_k$$

and

$$\widehat{\theta}_N = - \frac{\sum_{k=1}^N \nu_k \int_0^T (u_k(t) du_k(t) + \rho_k u_k^2(t) dt)}{\sum_{k=1}^N \nu_k^2 \int_0^T u_k^2(t) dt}.$$

**Theorem 1.1.** *The divergence of the series  $\sum_k \nu_k^2 / (\rho_k + \theta \nu_k)$  is necessary and sufficient to have consistency and asymptotic normality of  $\widehat{\theta}_N$ . In particular, if the series  $\sum_k \nu_k^2 / (\rho_k + \theta \nu_k)$  diverges, then*

$$\lim_{N \rightarrow \infty} \widehat{\theta}_N = \theta$$

with probability one,

$$\lim_{N \rightarrow \infty} \left( \sum_{k \leq N} \frac{\nu_k^2}{\rho_k + \theta \nu_k} \right)^{1/2} (\widehat{\theta}_N - \theta) \stackrel{d}{=} \mathcal{N}(0, 2/T),$$

and the measures generated by the solutions of equation (1.6) are mutually singular for different values of  $\theta$ .

For the convenience of the reader, the following section summarizes the main notions and technical tools necessary to study the estimation

problem for stochastic parabolic equation and to prove the above theorem: Ornstein-Uhlenbeck process and its properties, Law of Large Numbers and the Central Limit Theorem for independent but not identically distributed random variables, and the cylindrical Brownian motion. Section 2.4 summarizes the main facts and presents some examples related to absolutely continuous and singular statistical models; the book by Ibragimov and Khas'minskiĭ [31] provides more information on the subject. Section 3 illustrates the main steps of the proof of Theorem 1.1 in the particular case of the stochastic heat equation (1.1); Theorem 1.1 itself is proved in Section 4. Finally, Section 5 discusses other statistical estimation problems for stochastic parabolic equations.

### Notations

Throughout the presentation below, we fix a stochastic basis  $\mathbb{F} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$  with the usual assumptions (completeness of  $\mathcal{F}_0$  and right-continuity of  $\mathcal{F}_t$ ). We also assume that  $\mathbb{F}$  is large enough to support countably many independent standard Brownian motions. For a random variable  $\xi$ ,  $\mathbb{E}\xi$  and  $\text{Var } \xi$  denote the expectation and the variance, respectively.  $\mathbb{R}^n$  is an  $n$ -dimensional Euclidean space,  $\mathcal{N}(m, \sigma^2)$  is a Gaussian random variable with mean  $m$  and variance  $\sigma^2$ ,  $B^\top$  is the adjoint of the operator  $B$ .

Notation  $a_n \sim b_n$  for two sequences  $\{a_n, n \geq 1\}$ ,  $\{b_n, n \geq 1\}$ , with  $a_n > 0$ ,  $b_n > 0$ , means

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = 1.$$

For example,  $n^2 - 2n \sim n^2$  and  $\sum_{k=1}^n k^2 \sim n^3/3$ .

## 2. Some background from probability and statistics

**2.1. The Ornstein-Uhlenbeck process.** Let  $w = w(t)$ ,  $t \geq 0$ , be a standard Brownian motion. In the terminology of this paper, an *Ornstein-Uhlenbeck process*  $X = X(t; a)$ ,  $t \geq 0$ , with parameter  $a$  is the solution of the stochastic ordinary differential equation

$$(2.1) \quad dX(t; a) = -aX(t; a) dt + dw(t), \quad t > 0.$$

The process  $X$  is called *stable* if  $a > 0$ . Note that separation of variables in equation (1.1) led to stable Ornstein-Uhlenbeck processes (1.2) with parameters  $\theta k^2$ , and similarly, the analysis of the more general stochastic parabolic equation (1.6) also leads to (2.1) with  $a > 0$ .

It follows from (2.1) that

$$(2.2) \quad X(t; a) = X(0; a)e^{-at} + \int_0^t e^{-a(t-s)} dw(s).$$

**Theorem 2.1.** *Fix  $T > 0$  and assume that  $X(0; a) = 0$ .*

(1) *Define the random variable*

$$\xi(a) = \int_0^T X^2(t; a) dt.$$

Then

$$(2.3) \quad \lim_{a \rightarrow +\infty} a \mathbb{E}\xi(a) = \frac{T}{2},$$

$$(2.4) \quad \mathbb{E}\xi^n(a) \leq C(n, T)a^{-n}, \quad n = 1, 2, \dots,$$

$$(2.5) \quad \lim_{a \rightarrow +\infty} a^3 \text{Var} \xi(a) = \frac{T}{2}.$$

(2) *Denote by  $\mathbf{P}_T^a$  the measure generated by the process  $X(t; a)$ ,  $0 \leq t \leq T$ , in the space of continuous functions on  $[0, T]$ . Then the measures  $\mathbf{P}_T^a$  are equivalent (mutually absolutely continuous) for all  $a$  and*

$$(2.6) \quad \frac{d\mathbf{P}_T^b}{d\mathbf{P}_T^a}(X(\cdot; a)) \\ = \exp \left( -(b-a) \int_0^T X(t; a) dX(t; a) - \frac{b^2 - a^2}{2} \int_0^T X^2(t; a) dt \right).$$

*In particular,  $\mathbf{P}_T^0$  is the Wiener measure (the measure generated by the standard Brownian motion), and*

$$(2.7) \quad \frac{d\mathbf{P}_T^a}{d\mathbf{P}_T^0}(X(\cdot; a)) = \exp \left( -a \int_0^T X(t; a) dX(t; a) - \frac{a^2}{2} \int_0^T X^2(t; a) dt \right).$$

*Proof:* (1) Everything is proved by direct computation. For (2.3), the computations are easy:

$$(2.8) \quad \mathbb{E}X^2(t; a) = \int_0^t e^{-2a(t-s)} ds = \frac{1}{2a}(1 - e^{-2at}),$$

so

$$(2.9) \quad \mathbb{E}\xi(a) = \int_0^T \mathbb{E}X^2(t; a) dt = \frac{T}{2a} - \frac{1}{4a^2}(1 - e^{-2aT}),$$

and the result follows.

For (2.4), we use

- Jensen's inequality

$$\left( \int_0^T f^2(t) dt \right)^n = \left( T \int_0^T f^2(t) \frac{dt}{T} \right)^n \leq T^n \int_0^T f^{2n}(t) \frac{dt}{T};$$

- the fact that  $X(t; a)$  is a zero-mean Gaussian random variable, which implies

$$(2.10) \quad \mathbb{E}X^{2n}(t; a) = \left( \prod_{k=1}^n (2k-1) \right) (\mathbb{E}X^2(t; a))^n.$$

Then

$$\begin{aligned} \mathbb{E}\xi^n(a) &= \mathbb{E} \left( \int_0^T X^2(t; a) dt \right)^n \leq T^{n-1} \int_0^T \mathbb{E}X^{2n}(t; a) dt \\ &\leq T^{n-1} \left( \prod_{k=1}^n (2k-1) \right) \int_0^T (\mathbb{E}X^2(t; a))^n dt \leq C(T, n) a^{-n}, \end{aligned}$$

where the last inequality follows from (2.10) and (2.8).

For (2.5), it is necessary to find  $\mathbb{E}(\xi(a))^2$ , and the computations are more complicated. Here are two possible ways to approach the computations.

1. One way is to use the formula for the moment generating function  $\psi(y; a) = \mathbb{E} \exp(-y\xi(a))$ ,  $y > 0$ :

$$(2.11) \quad \begin{aligned} \psi(y; a) &= e^{aT/2} \left( \frac{2\varrho(y; a)}{(\varrho(y; a) - a)e^{-T\varrho(y; a)} + (\varrho(y; a) + a)e^{T\varrho(y; a)}} \right)^{1/2} \end{aligned}$$

$$(2.12) \quad = e^{(\varrho(y; a) + a)T/2} \left( \frac{2\varrho(y; a)}{(\varrho(y; a) + a)(e^{T\varrho(y; a)} - 1) + 2\varrho(y; a)} \right)^{1/2},$$

where  $\varrho(y; a) = (a^2 + 2y)^{1/2}$ ; see Liptser and Shiryaev [47, Lemma 17.3]. Then

$$\mathbb{E}(\xi(a))^2 = \lim_{y \searrow 0} \frac{\partial^2 \psi(y; a)}{\partial y^2},$$



and both differentiation and evaluation of the limit can be carried out with the help of a computer algebra system. The details are left to the reader (see also Cialenco et al. [15]).

2. Alternatively, it follows from the definition of  $\xi(a)$  that

$$(2.13) \quad \begin{aligned} \mathbb{E}(\xi(a))^2 &= \int_0^T \int_0^T \mathbb{E}(X^2(t; a)X^2(s, a)) ds dt \\ &= 2 \int_0^T \int_0^t \mathbb{E}(X^2(t; a)X^2(s, a)) ds dt, \end{aligned}$$

and, for each  $t > s$ , the random variables  $X(t; a)$ ,  $X(s; a)$  are jointly Gaussian with zero mean and correlation coefficient

$$\rho(t, s) = \frac{e^{-a(t-s)} - e^{-a(t+s)}}{((1 - e^{-2at})(1 - e^{-2as}))^{1/2}}.$$

Note that if  $\alpha$ ,  $\beta$  are jointly Gaussian, with zero mean, unit variance, and correlation  $\rho$ , then

$$\begin{aligned} \mathbb{E}(\alpha^2\beta^2) &= \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^2 y^2 \exp\left(-\frac{1}{2(1-\rho^2)}(x^2+y^2-2\rho xy)\right) dx dy \\ &= 1 + 2\rho^2. \end{aligned}$$

As a result,

$$\mathbb{E}(X^2(t; a)X^2(s, a)) = \frac{(1 - e^{-2at})(1 - e^{-2as})}{4a^2} (1 + 2\rho^2(t, s)),$$

and, by (2.13),

$$\begin{aligned} \mathbb{E}(\xi(a))^2 &= \frac{T^2}{4a^2} + \frac{1}{a^2} \int_0^T \int_0^t e^{-2a(t-s)} ds dt + o(a^{-3}) \\ &= (\mathbb{E}\xi(a))^2 + \frac{T}{2a^3} + o(a^{-3}), \end{aligned}$$

where  $\lim_{a \rightarrow +\infty} a^3 o(a^{-3}) = 0$ . This implies (2.5).

(2) For the proof of (2.6), see Lipster and Shiryaev [46, Theorem 7.19]. It is important to keep in mind that while the density  $d\mathbf{P}_T^b/d\mathbf{P}_T^a$  is a functional defined on all continuous functions, it has nice closed-form expressions only when evaluated on  $X(\cdot; b)$  or  $X(\cdot; a)$ ; each of these expressions defines a random variable on the original probability space  $\Omega$ .

Clearly,

$$\frac{d\mathbf{P}_T^a}{d\mathbf{P}_T^b}(X(\cdot; a)) = \left( \frac{d\mathbf{P}_T^b}{d\mathbf{P}_T^a}(X(\cdot; a)) \right)^{-1}.$$

Note also that (2.6) and (2.1) imply

$$\frac{d\mathbf{P}_T^b}{d\mathbf{P}_T^a}(X(\cdot; a)) = \exp \left( (a-b) \int_0^T X(t; a) dw(t) - \frac{(a-b)^2}{2} \int_0^T X^2(t; a) dt \right)$$

and so

$$\mathbb{E} \left( \frac{d\mathbf{P}_T^b}{d\mathbf{P}_T^a}(X(\cdot; a)) \right) = 1, \quad \mathbb{E} \left( \frac{d\mathbf{P}_T^a}{d\mathbf{P}_T^b}(X(\cdot; a)) \right) \neq 1.$$

Finally, let us point out that (2.7) is consistent with the Gisanov Theorem. Indeed, if

$$\begin{aligned} Z(a) &= \exp \left( a \int_0^T X(t; a) dX(t; a) + \frac{a^2}{2} \int_0^T X^2(t; a) dt \right) \\ &= \exp \left( a \int_0^T X(t; a) dw(t) - \frac{a^2}{2} \int_0^T X^2(t; a) dt \right) \end{aligned}$$

and  $\tilde{\mathbb{P}}$  is the probability measure on  $(\Omega, \mathcal{F}_T)$  such that  $d\tilde{\mathbb{P}} = Z(a) d\mathbb{P}$ , then, by the Girsanov Theorem [46, Theorem 6.3],  $X(t; a)$ ,  $0 \leq t \leq T$ , is a standard Brownian motion under  $\tilde{\mathbb{P}}$ . In particular,

$$\tilde{\mathbb{E}} \exp \left( -a \int_0^T X(t; a) dX(t; a) - \frac{a^2}{2} \int_0^T X^2(t; a) dt \right) = \tilde{\mathbb{E}}(1/Z(a)) = 1.$$

This completes the proof of Theorem 2.1.  $\square$

**2.2. LLN and CLT.** The proof of consistency and asymptotic normality of (1.4) and similar estimators relies on the Law of Large Numbers (LLN) and the Central Limit Theorem (CLT) for random variables that are independent but not identically distributed.

**Theorem 2.2** (The strong law of large numbers). *Let  $\xi_n$ ,  $n \geq 1$ , be a sequence of independent random variables and  $b_n$ ,  $n \geq 1$ , a sequence of positive numbers such that  $b_{n+1} \geq b_n$ ,  $\lim_{n \rightarrow \infty} b_n = +\infty$ , and*

$$\sum_{n \geq 1} \frac{\text{Var } \xi_n}{b_n^2} < \infty.$$

Then

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

with probability one.

*Proof:* See, for example, Shiryaev [72, Theorem IV.3.2].  $\square$

**Corollary 2.3.** *Let  $\xi_n$ ,  $n \geq 1$ , be independent random variables such that  $\xi_n \geq 0$ ,  $\sum_{k \geq 1} \mathbb{E}\xi_k = +\infty$ , and*

$$\sum_{n \geq 1} \frac{\text{Var } \xi_n}{(\sum_{k=1}^n \mathbb{E}\xi_k)^2} < \infty.$$

Then

$$\lim_{n \rightarrow \infty} \frac{\sum_{k=1}^n \xi_k}{\sum_{k=1}^n \mathbb{E}\xi_k} = 1$$

with probability one.

*Proof:* Take  $b_n = \sum_{k=1}^n \mathbb{E}\xi_k$  and apply Theorem 2.2.  $\square$

**Theorem 2.4** (Classical Central Limit Theorem). *Assume that  $\xi_n$ ,  $n \geq 1$ , are independent random variables with zero mean and variance  $\sigma_n^2 > 0$ , and assume that*

$$(2.14) \quad \sigma_n^2 \sim Cn^\alpha$$

and

$$(2.15) \quad \mathbb{E}\xi_n^4 \leq C_1\sigma_n^4$$

for some  $C > 0$ ,  $C_1 > 0$ , and  $\alpha \geq -1$ , all independent of  $n$ . Then, as  $n \rightarrow \infty$ , the sequence  $(\sum_{k=1}^n \xi_k) / (\sum_{k=1}^n \sigma_k^2)^{1/2}$  converges in distribution to the Gaussian random variable with zero mean and unit variance:

$$\lim_{n \rightarrow \infty} \frac{\sum_{k=1}^n \xi_k}{(\sum_{k=1}^n \sigma_k^2)^{1/2}} \stackrel{d}{=} \mathcal{N}(0, 1).$$

*Proof:* To simplify the notations, define

$$D_n = \sum_{k=1}^n \sigma_k^2.$$

We have to verify the classical condition of Lindeberg [72, Theorem III.4.1]:

$$(2.16) \quad \lim_{n \rightarrow \infty} \frac{\sum_{k=1}^n \mathbb{E}(\xi_k^2 I(|\xi_k| > \varepsilon \sqrt{D_n}))}{D_n} = 0$$

for every  $\varepsilon > 0$ , where  $I(A)$  is the indicator function of the set  $A$ . We have

$$\mathbb{E} \left( \xi_k^2 I(|\xi_k| > \varepsilon \sqrt{D_n}) \right) \leq (\mathbb{E} \xi_k^4)^{1/2} \left( \mathbb{P}(|\xi_k| > \varepsilon \sqrt{D_n}) \right)^{1/2} \leq \frac{\mathbb{E} \xi_k^4}{D_n \varepsilon^2},$$

where the first inequality follows from the Cauchy-Schwarz and the second, from the Chebychev inequality. By (2.14) and (2.15), the convergence in (2.16) will follow from

$$(2.17) \quad \lim_{n \rightarrow \infty} \frac{\sum_{k=1}^n k^{2\alpha}}{\left( \sum_{k=1}^n k^\alpha \right)^2} = 0.$$

If  $-1 \leq \alpha < -1/2$ , then (2.17) is obvious because the series on the top converges and the series on the bottom diverges. If  $\alpha \geq -1/2$ , then (2.17) follows from

$$\sum_{k=1}^n k^\beta \sim \frac{n^{1+\beta}}{1+\beta}, \quad \beta > -1. \quad \square$$

Another useful versions of the CLT comes from the theory of martingales.

**Theorem 2.5** (Martingale Central Limit Theorem). *Let  $M_n = M_n(t)$ ,  $t \geq 0$ ,  $n \geq 1$ , be a sequence of continuous square-integrable martingales with quadratic variations  $\langle M_n \rangle = \langle M_n \rangle(t)$ . If, for some  $T > 0$ ,*

$$\lim_{n \rightarrow \infty} \frac{\langle M_n \rangle(T)}{\mathbb{E} \langle M_n \rangle(T)} = 1 \text{ in probability,}$$

then

$$\lim_{n \rightarrow \infty} \frac{M_n(T)}{(\mathbb{E} \langle M_n \rangle(T))^{1/2}} \stackrel{d}{=} \mathcal{N}(0, 1).$$

*Proof:* This follows from a limit theorem for martingales: if  $X_n, X$  are continuous square-integrable martingales such that  $X$  is a Gaussian process and  $\lim_{n \rightarrow \infty} \langle X_n \rangle(T) = \langle X \rangle(T)$  in probability, then  $\lim_{n \rightarrow \infty} X_n(T) = X(T)$  in distribution; see, for example Jacod and Shiryaev [36, Theorem VIII.4.17] or Liptser and Shiryaev [45, Theorem 5.5.4(II)]. It now remains to take

$$X_n(t) = \frac{M_n(t)}{(\mathbb{E} \langle M_n \rangle(T))^{1/2}}, \quad X(t) = \frac{w(t)}{\sqrt{T}},$$

where  $w$  is a standard Brownian motion. □

**Corollary 2.6.** *Let  $w_k = w_k(t)$  be independent standard Brownian motions and let  $f_k = f_k(t)$  be adapted, continuous, square-integrable processes such that*

$$\lim_{n \rightarrow \infty} \frac{\sum_{k=1}^n \int_0^T f_k^2(t) dt}{\sum_{k=1}^n \mathbb{E} \int_0^T f_k^2(t) dt} = 1$$

*in probability. Then*

$$\lim_{n \rightarrow \infty} \frac{\sum_{k=1}^n \int_0^T f_k(t) dw_k(t)}{\left( \sum_{k=1}^n \mathbb{E} \int_0^T f_k^2(t) dt \right)^{1/2}} \stackrel{d}{=} \mathcal{N}(0, 1).$$

*Proof:* This follows from Theorem 2.5 with

$$M_n(t) = \sum_{k=1}^n \int_0^t f_k(s) dw_k(s). \quad \square$$

**2.3. Cylindrical Brownian Motion.** Let  $H$  be a separable Hilbert space with norm  $\|\cdot\|_H$  and inner product  $(\cdot, \cdot)_H$ .

**Definition 2.7.** A cylindrical Brownian motion  $W = W(t)$  on a Hilbert space  $H$  is a collection of zero-mean Gaussian random processes  $W_f = W_f(t)$ ,  $t \geq 0$ ,  $f \in H$ , such that, for every  $f, g \in H$  and  $t, s \geq 0$ ,

$$(2.18) \quad \mathbb{E}(W_f(t)W_g(s)) = (f, g)_H \min(t, s).$$

**Proposition 2.8.** *Let  $W$  be a cylindrical Brownian motion on a Hilbert space  $H$ .*

(a) *The mapping  $f \mapsto W_f$  is linear and therefore, for every  $t_i$ ,  $i = 1, \dots, m$ ,  $f_j$ ,  $j = 1, \dots, n$ , the collection of random variables  $\{W_{f_j}(t_i)$ ,  $i = 1, \dots, m$ ,  $j = 1, \dots, n\}$  is a Gaussian system.*

(b) *If  $h_1, h_2$  are the elements of  $H$  such that  $\|h_1\|_H = \|h_2\|_H = 1$  and  $(h_1, h_2)_H = 0$ , then the processes  $w_1 = w_1(t)$  and  $w_2 = w_2(t)$  defined by  $w_j(t) = W_{h_j}(t)$ ,  $j = 1, 2$ , are independent standard Brownian motions.*

*Proof:* (a) By direct computation, for  $a, b \in \mathbb{R}$  and  $f, g \in H$ ,

$$\mathbb{E}(W_{af+bg}(t) - aW_f(t) - bW_g(t))^2 = 0;$$

cf. Nualart [56, Definition 1.1.1].

(b) By definition of  $W$ , the pair  $(w_1, w_2)$  is a zero-mean Gaussian process, and  $\mathbb{E}w_i(t)w_j(s) = \min(t, s)(h_i, h_j)_H$ , which completes the proof.  $\square$

If  $f = f(t)$  is an adapted, square-integrable process with values in  $H$  (that is,  $\mathbb{E} \int_0^T \|f(t)\|_H^2 dt < \infty$ ), then we define the stochastic integral  $\int_0^T \langle f(t), dW(t) \rangle$  by the formula

$$(2.19) \quad \int_0^T \langle f(t), dW(t) \rangle = \sum_{k \geq 1} \int_0^T (f(t), h_k)_H dw_k(t),$$

where  $\{h_k, k \geq 1\}$  is an orthonormal basis in  $H$  and  $w_k(t) = W_{h_k}(t)$ ; of course, this definition does not depend on the choice of the basis in  $H$  (see, for example, Rozovskiĭ [68, Chapter 2] or Walsh [74, Chapter I]).

Proposition 2.8(b) suggests a representation of  $W(t)$  as a series

$$(2.20) \quad W(t) = \sum_{k \geq 1} w_k(t) h_k,$$

where  $\{h_k, k \geq 1\}$  is an orthonormal basis in  $H$  and  $w_k = W_{h_k}$ ,  $k \geq 1$ , are independent standard Brownian motions. While the series does not converge in  $H$ , it is possible to embed  $H$  in a bigger Hilbert space, where the series will converge. For example, define  $X$  as the closure of  $H$  in the norm

$$(2.21) \quad \|f\|_X = \left( \sum_{k \geq 1} k^{-2} (f, h_k)_H^2 \right)^{1/2}.$$

Direct computations show that  $W$  is a continuous  $X$ -valued square-integrable martingale, and

$$\mathbb{E} \|W(t)\|_X^2 = t \sum_{k \geq 1} k^{-2} = \frac{\pi^2 t}{6}.$$

There are many other spaces in which  $W$  becomes a continuous square-integrable martingale: it is enough to replace  $k^{-2}$  in (2.21) with  $k^{-\beta}$  with  $\beta > 1$ .

If  $W$  is a cylindrical Brownian motion on  $L_2((0, +\infty))$ , and  $\chi_x$  is the indicator function of the interval  $[0, x]$ , then, by direct computation,  $W(t, x) = W_{\chi_x}(t)$  is a Brownian sheet and, for every  $f \in L_2((0, +\infty))$ ,

$$W_f(t) = \int_0^t \int_0^{+\infty} f(y) W(ds, dy);$$

see, for example, Walsh [74, p. 284].

Definition 2.7 can be generalized to allow spatial covariance: given a non-negative, bounded, self-adjoint operator  $Q$  on  $H$ , we define the  $Q$ -cylindrical Brownian motion  $W^Q$  on  $H$  by replacing (2.18) with

$$(2.22) \quad \mathbb{E}\left(W_f^Q(t)W_g^Q(s)\right) = (Qf, g)_H \min(t, s).$$

If  $Q$  has a complete orthonormal system of eigenfunctions  $\{h_k, k \geq 1\}$  and  $Qh_k = q_k h_k, q_k > 0$ , then (2.20) becomes

$$W^Q(t) = \sum_{k \geq 1} q_k w_k(t) h_k,$$

where  $w_k = q_k^{-1} W_{h_k}^Q, k \geq 1$  are independent standard Brownian motions.

Recall that

- (a) an operator  $B$  from a separable Hilbert space  $H$  to a separable Hilbert space  $X$  is called Hilbert-Schmidt if

$$\sum_{k \geq 1} \|Bh_k\|_X^2 < \infty$$

for one (hence all) orthonormal basis  $\{h_k, k \geq 1\}$  in  $H$ ;

- (b) a non-negative self-adjoint operator  $Q$  on a separable Hilbert space  $H$  is called **trace class** if

$$\sum_{k \geq 1} (Qh_k, h_k)_H < \infty$$

for one (hence all) orthonormal basis  $\{h_k, k \geq 1\}$  in  $H$ .

**Proposition 2.9.** (a) Let  $W$  be a cylindrical Brownian motion on a separable Hilbert space  $H$  and  $X$ , a Hilbert space such that  $H$  is a dense sub-set of  $X$ . Then  $W$  is a continuous  $X$ -valued square-integrable martingale if and only if the embedding  $j: H \rightarrow X$  is a Hilbert-Schmidt operator; in this case,  $W$  naturally extends to a  $Q$ -cylindrical Brownian motion on  $X$  with  $Q = jj^\top$ , and  $Q$  is trace class.

(b) Let  $W^Q$  be a  $Q$ -cylindrical Brownian motion on a separable Hilbert space  $H$ . Then  $W^Q$  is a continuous  $H$ -valued square-integrable martingale if and only if the operator  $Q$  is trace class.

*Proof:* Below is an outline of the proof; the details are left to the reader.

- (a) We have

$$\mathbb{E}\|W(t) - W(s)\|_X^2 = (t - s) \sum_{k \geq 1} \|h_k\|_X^2,$$

and the series converges if and only if  $j$  is Hilbert-Schmidt. The continuity of  $W$  then follows by the Kolmogorov criterion (Kunita [42, Theorem 1.4.1]). Finally, for  $f \in X$ , we set

$$W_f^Q = W_{j^\top f};$$

the adjoint operator  $j^\top$  is defined on all of  $X$  (see, for example, Yosida [75, Theorem VII.2]). Note that  $(Qh_k, h_k)_H = \|j^\top h_k\|_H^2 = \|h_k\|_X^2$ .

(b) This follows from

$$\mathbb{E}\|W^Q(t) - W^Q(s)\|_H^2 = (t - s) \sum_{k \geq 1} (Qh_k, h_k)_H. \quad \square$$

**2.4. Statistical models.** A statistical model (or experiment) generated by random elements  $X(\theta)$  is a collection  $\mathcal{P} = \{\mathcal{X}, \mathfrak{X}, \mathbf{P}^\theta, \theta \in \Theta\}$ , where each  $\mathbf{P}^\theta$  is a probability measure on a measurable space  $(\mathcal{X}, \mathfrak{X})$  such that  $\mathbf{P}^\theta(A) = \mathbb{P}(X(\theta) \in A)$ ,  $A \in \mathfrak{X}$ . In the parametric models,  $\Theta$  is a subset of a finite-dimensional Euclidean space.

An estimator of  $\theta$  is a *random variable*  $\Psi(X(\theta))$ , where  $\Psi$  is a measurable mapping from  $\mathcal{X}$  to  $\Theta$ . The corresponding estimate of  $\theta$  is the *number*  $\Psi(X^\circ(\theta^*))$ , where  $X^\circ(\theta)$  is the observed realization of the random element  $X(\theta)$ .

In general, an estimate of  $\theta$ , being a realization of a random variable, is not equal to  $\theta$ . Accordingly, a family  $\{\mathcal{P}^N, N > 0\}$  of statistical models is introduced, with  $N$  characterizing the amount of information about  $\theta$  (the larger  $N$ , the more information is available to the observer). For example,  $\mathcal{P}^N$  can be a product of  $N$  independent copies of  $\mathcal{P}$ , which corresponds to observing  $N$  independent realizations of  $X$ .

Given  $\mathcal{P}^N$ , the corresponding family of estimators is then constructed and studied in the limit  $N \rightarrow \infty$ . One of the objectives is to establish consistency of the estimators (convergence to the true value of the parameter) as  $N \rightarrow \infty$ .

In *absolutely continuous* statistical models, maximum likelihood estimators are often used.

**Definition 2.10.** A statistical model  $\mathcal{P}$  is called **absolutely continuous** if there exists a probability measure  $\mathbf{Q}$  on  $(\mathcal{X}, \mathfrak{X})$  such that every  $\mathbf{P}^\theta$  is absolutely continuous with respect to  $\mathbf{Q}$ . The statistical model  $\mathcal{P}$  is called **singular** if the measures  $\mathbf{P}^{\theta_1}$  and  $\mathbf{P}^{\theta_2}$  are mutually singular for  $\theta_1 \neq \theta_2$ .



Let  $\mathcal{P}$  be an absolutely continuous model and consider the density

$$p(x, \theta) = \frac{d\mathbf{P}^\theta}{d\mathbf{Q}}(x, \theta), \quad x \in \mathcal{X}, \quad \theta \in \Theta.$$

The maximum likelihood estimator  $\hat{\theta}$  of  $\theta$  is defined by

$$(2.23) \quad \hat{\theta}(X) = \arg \max_{\theta \in \bar{\Theta}} p(x, \theta) \Big|_{x=X(\theta)},$$

where  $\bar{\Theta}$  is the closure of  $\Theta$ . Similarly, a collection  $\{\mathcal{P}^N, N > 0\}$  of absolutely continuous statistical models leads to a collection  $\hat{\theta}_N$  of maximum likelihood estimators.

The parameter  $N$  does not have to be discrete. For example, consider a family of Ornstein-Uhlenbeck processes  $X = X(t; \theta)$  defined by

$$(2.24) \quad dX(t; \theta) = \theta X(t; \theta) dt + \sigma dw(t), \quad 0 < t < T,$$

with known  $\sigma > 0$ . For every fixed  $\sigma$  and  $T$ , we get an absolutely continuous statistical model in which  $\mathcal{X}$  is the set of continuous real-valued functions on  $[0, T]$ ,  $\mathfrak{X}$  is the Borel sigma-algebra on  $\mathcal{X}$ , and  $\mathbf{Q}$  is the Wiener measure (the measure on  $(\mathcal{X}, \mathfrak{X})$  generated by the Brownian motion  $w$ ); see Theorem 2.1(2). Then

$$(2.25) \quad \hat{\theta} = \frac{\int_0^T X(t; \theta) dX(t; \theta)}{\int_0^T X^2(t; \theta) dt};$$

see Liptser and Shiryaev [47, Formula 17.45]. There are at least three ways to achieve consistency:

- (1) Keeping  $T$  and  $\sigma$  fixed, consider  $N$  independent copies  $X_k$  of  $X$ .  
Then

$$\hat{\theta}_N = \frac{\sum_{k=1}^N \int_0^T X_k(t; \theta) dX_k(t; \theta)}{\sum_{k=1}^N \int_0^T X_k^2(t; \theta) dt}.$$

- (2) Keeping  $\sigma$  fixed, let  $T \rightarrow \infty$  in (2.25), so that  $N = T$  (large time asymptotic).
- (3) Keeping  $T$  fixed and assuming  $X(0; \theta) \neq 0$ , let  $\sigma \rightarrow 0$ , so that  $N = 1/\sigma$  (small noise asymptotic).

It is also clear that the same three methods can be used to achieve consistency in any absolutely continuous model generated by a stochastic evolution equation. For a detailed analysis of the models generated by stochastic ordinary differential equations, see the books [43], [44] by Yu. A. Kutoyants.

**Definition 2.11.** The estimator  $\Psi_N$  of  $\theta$  is said to converge to  $\theta$  with the **rate of convergence**  $N^\alpha$ ,  $\alpha > 0$ , if the sequence  $\{N^\alpha(\Psi_N - \theta), N > 0\}$  converges in distribution to a non-degenerate random variable  $\zeta$  (that is,  $\text{Var} \zeta > 0$ ). If  $\zeta$  is a Gaussian random variable, then  $\Psi_N$  is called **asymptotically normal**.

In general, given a collection  $\{\mathcal{P}^N, N > 0\}$  of absolutely continuous statistical models, the asymptotic properties of the maximum likelihood estimator  $\hat{\theta}_N$ , such as consistency and asymptotic normality, can be described in terms of the properties of the corresponding densities; see, for example, Ibragimov and Khas'minskiĭ [31, Theorem III.1.1].

If the statistical model  $\mathcal{P}$  is singular, then it is often possible (at least in theory) to get the true value of the parameter without introducing the family  $\mathcal{P}^N$ . A well-known example is estimation of the diffusion coefficient  $\theta$  from the observations of

$$(2.26) \quad X(t) = \theta w(t), \quad \theta > 0,$$

where  $w = w(t)$  is a standard Brownian motion: since  $X$  is a square-integrable martingale with quadratic variation  $\langle X \rangle(t) = \theta^2 t$ , it follows that

$$(2.27) \quad \theta = \sqrt{\frac{\langle X \rangle(T)}{T}}$$

for every  $T > 0$ . Note that, since the quadratic variation is not available directly, a *computable* form of (2.27) is

$$(2.28) \quad \theta = \left( \frac{1}{T} \lim_{N \rightarrow \infty} \sum_{k=1}^N \left( X(Tk/N) - X(T(k-1)/N) \right)^2 \right)^{1/2}.$$

Similar ideas can be used to estimate the diffusion coefficient in more general Itô equations  $dX = b(t, X(t)) dt + \theta \sigma(t, X(t)) dw(t)$ . For more details, see the survey by Ait-Sahalia [11] and references therein.

Another example of a singular model is in the paper by Khas'minskiĭ et al. [37]: if the observations are a two-dimensional diffusion process  $(X, Y)$  with

$$dX(t) = Y(t) dt, \quad dY(t) = X(t) dt + \theta Y(t) dt + X(t) dw(t)$$

with a standard Brownian motion  $w = w(t)$  and initial conditions  $X(0) = 0$ ,  $Y(0) = 1$ , then, by direct computations,

$$\theta = \lim_{t \rightarrow 0} \frac{1}{t} \int_0^t \frac{dY(s)}{\sqrt{X^2(s) + Y^2(s)}};$$

note that the special choice of the initial conditions is essential.

We will see in Section 4 that stochastic parabolic equations give rise to a large class of singular models.

A systematic way to study a singular model is to approximate it with a family of absolutely continuous models. For (2.26), this family comes from the time discretizations (2.28), and for many stochastic parabolic equations, from the space discretization. We illustrate this idea in the next section using the stochastic heat equation on the interval.

### 3. Analysis of the stochastic heat equation on the interval

Let  $W = W(t)$  be a cylindrical Brownian motion on  $L_2((0, \pi))$  and  $\theta > 0$ . Consider the following stochastic heat equation

$$(3.1) \quad du(t, x) = \theta u_{xx}(t, x) dt + dW, \quad 0 < t \leq T, \quad x \in (0, \pi),$$

with zero initial and boundary conditions, where  $\theta > 0$  is a real number. To simplify the notations, we do not indicate explicitly the dependence of  $u$  on  $\theta$ .

**Definition 3.1.** A solution of equation (3.1) is a random element with values in  $L_2((0, T) \times (0, \pi))$ , such that, for every twice continuously differentiable on  $[0, \pi]$  function  $v = v(x)$  satisfying  $v(0) = v(\pi) = 0$  and every  $t \in [0, T]$ , the following equality holds with probability one:

$$(3.2) \quad (u, v)_{L_2((0, \pi))}(t) = \theta \int_0^t (u, v_{xx})_{L_2((0, \pi))}(s) ds + W_v(t).$$

**Proposition 3.2.** *There exists a unique solution of equation (3.1), and*

$$(3.3) \quad \mathbb{E} \|u\|_{L_2((0, \pi))}^2(t) \leq \frac{\pi^2}{12\theta}, \quad t \geq 0.$$

*Proof:* We solve equation (3.1) using the classical method of separation of variables. Let  $h_k(x) = \sqrt{2/\pi} \sin(kx)$ ,  $k \geq 1$ . Taking  $v = h_k$  in (3.2), we find that  $u_k(t) = (u, h_k)_{L_2((0, \pi))}(t)$  satisfies

$$(3.4) \quad u_k(t) = -\theta k^2 \int_0^t u_k(s) ds + w_k(t),$$

where  $w_k$  is a standard Brownian motion. Therefore,

$$\mathbb{E} u_k^2(t) = \int_0^t e^{-2\theta k^2(t-s)} ds \leq \frac{1}{2\theta k^2}.$$

Since  $h_k$ ,  $k \geq 1$ , is an orthonormal basis in  $L_2((0, \pi))$ , we conclude that  $u = \sum_{k \geq 1} u_k h_k$  is a solution of (3.1) and

$$\mathbb{E} \|u\|_{L_2((0, \pi))}^2(t) = \sum_{k \geq 1} \mathbb{E} u_k^2(t) \leq \sum_{k \geq 1} \frac{1}{2\theta k^2} = \frac{\pi^2}{12\theta}.$$

Uniqueness of solution of (3.1) follows from the uniqueness of solution of (3.4) for every  $k$ . Proposition 3.2 is proved.  $\square$

Note that since the Brownian motions  $w_k$ ,  $k \geq 1$ , are independent, the Ornstein-Uhlenbeck processes  $u_k$ ,  $k \geq 1$ , are also independent.

Let us now consider the problem of estimating the number  $\theta$  from the observations of the solution of (3.2). One can show that the solution generates a Gaussian measure in space of continuous processes with values in a suitable Hilbert space, and the measures are *singular* for different values of  $\theta$  (see Theorem 4.8 below). Using the terminology of Section 2.4, we have a singular statistical model, and we will approximate it with a sequence of absolutely continuous models by discretizing the space.

Assume that the observations of  $u_k(t)$  are available for  $t \in [0, T]$  and  $k = 1, \dots, N$ . For each  $\theta$  and each  $k$ , the Ornstein-Uhlenbeck process  $u_k$  generates the measure  $\mathbf{P}_T^{\theta, k}$  in the space of continuous real-valued functions on  $[0, T]$ , and, by Theorem 2.1(2), the measures are equivalent for different values of  $\theta$ . Similarly, the vector  $u_{(N, \theta)} = (u_k, k = 1, \dots, N)$ , generates a probability measure  $\mathbf{P}_{N, T}^\theta$  on the space of continuous  $\mathbb{R}^N$ -valued functions on  $[0, T]$ . Since the random processes  $u_k$  are independent for different  $k$ ,  $\mathbf{P}_{N, T}^\theta$  is a product measure:  $\mathbf{P}_{N, T}^\theta = \prod_{k=1}^N \mathbf{P}_T^{\theta, k}$ , and thus the measures  $\mathbf{P}_{N, T}^\theta$  are equivalent for different values of  $\theta$ . In particular, by (2.7),

$$(3.5) \quad \frac{d\mathbf{P}_{N, T}^\theta}{d\mathbf{P}_{N, T}^0}(u_{(N, \theta)}) \\ = \exp \left( \sum_{k=1}^N \left( -\theta k^2 \int_0^T u_k(t) du_k(t) - \frac{\theta^2 k^4}{2} \int_0^T u_k^2(t) dt \right) \right).$$

Maximizing the right-hand side of (3.5) with respect to  $\theta$ , we get the following expression for the maximum likelihood estimator  $\hat{\theta}_N$  of  $\theta$ , based on the observations  $u_k(t)$ ,  $k = 1, \dots, N$ ,  $t \in [0, T]$ :

$$(3.6) \quad \hat{\theta}_N = - \frac{\sum_{k=1}^N k^2 \int_0^T u_k(t) du_k(t)}{\sum_{k=1}^N k^4 \int_0^T u_k^2(t) dt}.$$

**Theorem 3.3.** *Estimator (3.6) is strongly consistent and asymptotically normal in the limit  $N \rightarrow \infty$ :*

$$(3.7) \quad \lim_{N \rightarrow \infty} \widehat{\theta}_N = \theta \quad \text{with probability one,}$$

$$(3.8) \quad \lim_{N \rightarrow \infty} N^{3/2}(\widehat{\theta}_N - \theta) = \zeta \quad \text{in distribution,}$$

where  $\zeta$  is a Gaussian random variable with zero mean and variance  $6\theta/T$ .

*Proof:* It follows from (3.4) that

$$(3.9) \quad \widehat{\theta}_N - \theta = -\frac{\sum_{k=1}^N k^2 \int_0^T u_k(t) dw_k(t)}{\sum_{k=1}^N k^4 \int_0^T u_k^2(t) dt}.$$

To prove consistency, we use Corollary 2.3. Note that each  $u_k$  is a stable Ornstein-Uhlenbeck process with parameter  $a = k^2\theta$ . By Theorem 2.1(1),

$$(3.10) \quad \mathbb{E} \int_0^T u_k^2(t) dt \sim \frac{T}{2\theta k^2}, \quad \text{Var} \int_0^T u_k^2(t) dt \sim \frac{T}{2\theta^3 k^6},$$

and so

$$\mathbb{E} \left( k^4 \int_0^T u_k^2(t) dt \right) \sim \frac{Tk^2}{2\theta}, \quad \text{Var} \left( k^4 \int_0^T u_k^2(t) dt \right) \sim \frac{Tk^2}{2\theta^3}.$$

By Corollary 2.3,

$$(3.11) \quad \lim_{N \rightarrow \infty} \frac{\sum_{k=1}^N k^4 \int_0^T u_k^2(t) dt}{\sum_{k=1}^N k^4 \int_0^T \mathbb{E} u_k^2(t) dt} = 1.$$

Next, we apply Theorem 2.2 with  $\xi_n = \int_0^T n^2 u_n(t) dw_n(t)$  and  $b_n = \sum_{k=1}^n \mathbb{E} \xi_k^2 \sim Tn^3/(6\theta)$  to conclude that

$$\lim_{N \rightarrow \infty} \frac{\sum_{k=1}^N k^2 \int_0^T u_k(t) dw_k(t)}{\sum_{k=1}^N k^4 \int_0^T \mathbb{E} u_k^2(t) dt} = 0$$

with probability one, and then (3.7) follows from (3.9) and (3.11).

Asymptotic normality (3.8) now follows from (3.11) and Corollary 2.6. Alternatively, since  $u_k(t)$  is a Gaussian random variable, (3.8) can be derived from Theorem 2.4 with

$$\sigma_k^2 = \mathbb{E} \int_0^T k^4 u_k^2(t) dt \sim \frac{Tk^2}{2\theta}. \quad \square$$

#### 4. Diagonalizable stochastic parabolic equations

The ideas used to study the stochastic heat equation on the interval (3.1) extend with little or no modification to equations such as

$$\begin{aligned} du - \theta \Delta u dt &= dW, \quad \theta > 0, \\ du - (\Delta u + \theta u) dt &= dW, \\ du + (\Delta^2 u + \theta \Delta u) dt &= dW, \end{aligned}$$

where  $\Delta$  is the Laplace operator, and to an abstract parabolic equation

$$(4.1) \quad du + (A_0 + \theta A_1)u dt = dW$$

under suitable assumptions about the operators  $A_0, A_1$ . The key feature of equation (3.1) is the possibility to write the solution using separation of variables; in what follows, we generalize this feature to (4.1) using the notion of a diagonalizable equation.

In general, if  $W$  is a cylindrical Brownian motion on a Hilbert space  $H$ , the solution of (4.1) is not an element of  $H$  for  $t > 0$ . There are two main approaches to circumvent this difficulty:

- (1) To introduce spatial covariance in the noise and consider  $W^Q$  instead of  $W$ .
- (2) To consider the equation in a bigger Hilbert space.

By Proposition 2.9, the two approaches are essentially equivalent, and we will use the second one. Later on, we will see that many equations driven by  $W^Q$  can be reduced to equations driven by  $W$ .

**4.1. Existence and regularity of solution.** Introduce the following objects:

- (1)  $H$ , a separable Hilbert space with an orthonormal basis  $\{h_k, k \geq 1\}$ ;
- (2)  $W = W(t)$ , a cylindrical Brownian motion on  $H$ ;
- (3)  $A_0, A_1$ , linear operators on  $H$ ;
- (4)  $\Theta = [a, b]$ , a closed bounded interval in  $\mathbb{R}$ .

For  $\theta \in \Theta$ , consider the following equation:

$$(4.2) \quad du(t) + (A_0 + \theta A_1)u(t) dt = dW(t), \quad 0 < t \leq T,$$

with zero initial condition  $u(0) = 0$  and fixed non-random  $T > 0$ .

**Definition 4.1.** Equation (4.2) is called **diagonalizable** if the operators  $A_0$  and  $A_1$  have point spectrum and a common system of eigenfunctions  $\{h_k, k \geq 1\}$ .

Denote by  $\rho_k$ ,  $\nu_k$ , and  $\mu_k(\theta)$  the eigenvalues of the operators  $A_0$ ,  $A_1$ , and  $A_0 + \theta A_1$ :

$$(4.3) \quad A_0 h_k = \rho_k h_k, \quad A_1 h_k = \nu_k h_k, \quad \mu_k(\theta) = \rho_k + \theta \nu_k.$$

**Definition 4.2.** A diagonalizable equation (4.2) is called **parabolic** if there exist positive numbers  $C^*$ ,  $c_1$ ,  $c_2$  such that  $\{\mu_k(\theta) + C^*, k \geq 1\}$  is a positive, non-decreasing, and unbounded sequence for all  $\theta \in \Theta$  and

$$(4.4) \quad c_1 \leq \frac{\mu_k(\theta_1) + C^*}{\mu_k(\theta_2) + C^*} \leq c_2$$

for all  $\theta_1, \theta_2 \in \Theta$ .

**Proposition 4.3.** *If equation (4.2) is diagonalizable and parabolic, then*

$$(4.5) \quad \lim_{k \rightarrow \infty} \mu_k(\theta) = +\infty$$

uniformly in  $\theta \in \Theta$ , and there exists an index  $J \geq 1$  and a number  $c_0$  such that, for all  $k \geq J$  and  $\theta \in \Theta$ ,

$$(4.6) \quad \mu_k(\theta) > 1,$$

$$(4.7) \quad \frac{|\nu_k|}{\mu_k(\theta)} \leq c_0.$$

*Proof:* Since  $\{\mu_k(\theta) + C^*, k \geq 1\}$  is a positive, non-decreasing, and unbounded sequence for all  $\theta \in \Theta$  and (4.4) holds, we have (4.5), and then (4.6) follows.

To prove (4.7), we argue by contradiction. Assume that the sequence  $\{|\nu_k| \mu_k^{-1}(\theta), k \geq 1\}$  is not uniformly bounded. Then there is a sequence  $\{|\nu_{k_j}| \mu_{k_j}^{-1}(\theta_j), j \geq 1\}$  such that

$$(4.8) \quad \lim_{j \rightarrow \infty} \frac{|\nu_{k_j}|}{\theta_j \nu_{k_j} + \rho_{k_j}} = +\infty.$$

With no loss of generality, assume that  $\nu_{k_j} > 0$ , and, since  $\Theta$  is compact, we also assume that  $\lim_{j \rightarrow \infty} \theta_j = \theta^\circ \in \Theta$  (if not, extract a further subsequence).

Then (4.8) implies

$$(4.9) \quad \lim_{j \rightarrow \infty} \frac{\rho_{k_j}}{\nu_{k_j}} = -\theta^\circ.$$

Note that  $\lim_{j \rightarrow \infty} |\nu_{k_j}| = +\infty$ , because  $\lim_{j \rightarrow \infty} (\theta^\circ \nu_{k_j} + \rho_{k_j}) = +\infty$ . Consequently,

$$\lim_{j \rightarrow \infty} \frac{\mu_{k_j}(\theta) + C^*}{\mu_{k_j}(\theta^\circ) + C^*} = \frac{\theta - \theta^\circ}{\theta^\circ + \lim_{j \rightarrow \infty} (\rho_{k_j} / \nu_{k_j})} = \infty, \quad \theta \neq \theta^\circ.$$

As a result, if (4.7) fails, then so does (4.4) for  $\theta_1 \neq \theta^\circ$ ,  $\theta_2 = \theta^\circ$ .  $\square$

**Example 4.4.** Let  $G$  be a smooth bounded domain in  $\mathbb{R}^d$  or a smooth compact  $d$ -dimensional manifold with a smooth measure,  $H = L_2(G)$ , and let  $\Delta$  be the Laplace operator on  $G$  (with zero boundary conditions if  $G$  is a domain). It is known (see, for example, Safarov and Vassiliev [69] or Shubin [73]) that

- (1)  $\Delta$  has a complete orthonormal system of eigenfunctions in  $H$ ;
- (2) the corresponding eigenvalues  $\lambda_k$  are negative, can be arranged in decreasing order, and there is a positive number  $c_0$  such that

$$(4.10) \quad |\lambda_k| \sim c_0 k^{2/d}.$$

The reader can verify that each of the following equations is diagonalizable and parabolic:

$$\begin{aligned} du - \theta \Delta u dt &= dW, & 0 < a \leq \theta \leq b, \\ du - (\Delta u + \theta u) dt &= dW, & a \leq \theta \leq b, \\ du + (\Delta^2 u + \theta \Delta u) dt &= dW, & a \leq \theta \leq b. \end{aligned}$$

From now on, we assume that equation (4.2) is diagonalizable and parabolic, and the eigenvalues of the operators  $A_0, A_1$  are enumerated so that  $\{\mu_k(\theta), k \geq 1\}$  is a non-decreasing sequence and (4.6) holds.

Let  $X$  be the closure of  $H$  in the norm

$$(4.11) \quad \|f\|_X = \left( \sum_{k \geq 1} k^{-2} (f, h_k)_H^2 \right)^{1/2}.$$

Then every element  $f$  of  $X$  is represented by a Fourier series  $f = \sum_{k \geq 1} f_k h_k$  and

$$\|f\|_X^2 = \sum_{k \geq 1} k^{-2} f_k^2.$$

Recall that the cylindrical Brownian motion  $W = W(t)$  is a continuous square-integrable martingale with values in  $X$  (see Proposition 2.9).

**Definition 4.5.** A solution of equation (4.2) is a continuous  $X$ -valued random process  $u = u(t)$  such that

$$(4.12) \quad u(t) = \sum_{k \geq 1} u_k(t) h_k$$

and

$$(4.13) \quad u_k(t) = -\mu_k(\theta) \int_0^t u_k(s) ds + W_{h_k}(t).$$



**Theorem 4.6.** *Assume that equation (4.2) is diagonalizable and parabolic. Then there exists a unique solution  $u = u(t)$  of (4.2).*

*Proof:* Uniqueness of solution follows from the uniqueness of solution of (4.13) for every  $k$ :

$$u_k(t) = \int_0^t e^{-\mu_k(\theta)(t-s)} dw_k(s).$$

It remains to show that the process  $u$  defined by (4.12) is a continuous  $X$ -valued process.

For  $0 \leq s < t \leq T$ ,

$$\begin{aligned} u_k(t) - u_k(s) &= \int_0^s \left( e^{-\mu_k(\theta)(t-r)} - e^{-\mu_k(\theta)(s-r)} \right) dw_k(r) \\ &\quad + \int_s^t e^{-\mu_k(\theta)(t-r)} dw(r), \end{aligned}$$

and so

$$(4.14) \quad \mathbb{E}(u_k(t) - u_k(s))^2 = \int_0^s \left( e^{-\mu_k(\theta)(t-r)} - e^{-\mu_k(\theta)(s-r)} \right)^2 dr \\ + \int_s^t e^{-2\mu_k(\theta)(t-r)} dr;$$

note that

$$\begin{aligned} &\int_0^s \left( e^{-\mu_k(\theta)(t-r)} - e^{-\mu_k(\theta)(s-r)} \right)^2 dr \\ &= \left( 1 - e^{-2\mu_k(\theta)s} \right) \left( 1 - e^{-\mu_k(\theta)(t-s)} \right) \frac{1 - e^{-\mu_k(\theta)(t-s)}}{2\mu_k(\theta)}. \end{aligned}$$

By (4.5), there exists a  $C > 1$  such that  $-2\mu_k(\theta) \leq C$  for all  $k \geq 1$  and  $\theta \in \Theta$ . Then (4.14) implies

$$(4.15) \quad \mathbb{E}(u_k(t) - u_k(s))^2 \leq C_0 |t - s|$$

with a suitable constant  $C_0$ , for example,  $C_0 = e^{CT}((1 + e^{CT})^2 + 1)$ . As a result,

$$\mathbb{E}\|u(t)\|_X^2 \leq C_0 T \sum_{k \geq 1} k^{-2} = \frac{C_0 T \pi^2}{6},$$

which implies  $u(t) \in L_2(\Omega; X)$  for all  $0 \leq t \leq T$ . Next, by the Cauchy-Schwarz inequality,

$$\begin{aligned} \mathbb{E}\|u(t) - u(s)\|_X^4 &= \mathbb{E} \left( \sum_{k \geq 1} k^{-2} (u_k(t) - u_k(s))^2 \right)^2 \\ &\leq \frac{\pi}{6} \sum_{k \geq 1} k^{-2} \mathbb{E}(u_k(t) - u_k(s))^4. \end{aligned}$$

Since each  $u_k$  is a zero-mean Gaussian process, (4.15) implies

$$\mathbb{E}(u_k(t) - u_k(s))^4 \leq 3C_0|t - s|^2,$$

and the continuity of  $u$  follows from the Kolmogorov criterion (see, for example, Kunita [42, Theorem 1.4.1]).  $\square$

*Remark 4.7.* Since the solution is defined by its Fourier coefficients, the space  $X$  is not an essential part of the definition and is only necessary to represent  $u$  as a process. The reader can check that Theorem 4.6 holds for any Hilbert space  $X$  such that  $H$  is a dense sub-set of  $X$  and the inclusion  $j: H \rightarrow X$  is a Hilber-Schmidt operator, so that  $\sum_{k \geq 1} \|h_k\|_X^2 < \infty$ .

**4.2. Parameter estimation.** Consider the diagonalizable parabolic equation

$$(4.16) \quad du(t) + (A_0 + \theta A_1)u(t) dt = dW(t), \quad 0 < t \leq T, \quad u(0) = 0,$$

driven by a cylindrical Brownian motion on a Hilbert space  $H$ . Let  $X$  be a Hilbert space such that  $H$  is a dense subset of  $X$  and  $W = W(t)$  is an  $X$ -valued continuous square-integrable martingale (for example, we can define  $X$  by (4.11)). According to Theorem 4.6, the solution  $u = u(t)$  of this equation is a continuous  $X$ -valued process

$$u(t) = \sum_{k \geq 1} u_k(t) h_k,$$

with

$$(4.17) \quad du_k(t) = -\mu_k(\theta)u_k(t) dt + dw_k(t), \quad u_k(0) = 0,$$

where  $\{h_k, k \geq 1\}$  is an orthonormal basis in  $H$ ,  $w_k = W_{h_k}$ ,  $A_0 h_k = \rho_k h_k$ ,  $A_1 h_k = \nu_k h_k$ , and  $\mu_k(\theta) = \rho_k + \theta \nu_k$ .

Assume that the observations of  $u_k(t)$  are available for  $t \in [0, T]$  and  $k = 1, \dots, N$ . For each  $\theta$  and each  $k$ , the Ornstein-Uhlenbeck process  $u_k$  generates the measure  $\mathbf{P}_T^{\theta, k}$  in the space of continuous real-valued functions on  $[0, T]$ , and, by Theorem 2.1(2), the measures are equivalent for different values of  $\theta$ . Similarly, the vector  $u_{(N, \theta)} = (u_k, k = 1, \dots, N)$ ,

generates a probability measure  $\mathbf{P}_{N,T}^\theta$  on the space of continuous  $\mathbb{R}^N$ -valued functions on  $[0, T]$ . Since the random processes  $u_k$  are independent for different  $k$ ,  $\mathbf{P}_{N,T}^\theta$  is a product measure:  $\mathbf{P}_{N,T}^\theta = \prod_{k=1}^N \mathbf{P}_T^{\theta,k}$ , and thus the measures  $\mathbf{P}_{N,T}^\theta$  are equivalent for different values of  $\theta$ . In particular, by (2.7),

$$(4.18) \quad \frac{d\mathbf{P}_{N,T}^\theta}{d\mathbf{P}_{N,T}^0}(u_{(N,\theta)}) \\ = \exp\left(\sum_{k=1}^N \left(-\mu_k(\theta) \int_0^T u_k(t) du_k(t) - \frac{(\mu_k(\theta))^2}{2} \int_0^T u_k^2(t) dt\right)\right).$$

Maximizing the expression on the right-hand side of (4.18) with respect to  $\theta$ , we get the following expression for the maximum likelihood estimator  $\hat{\theta}_N$  of  $\theta$  based on the observations  $u_k(t)$ ,  $k = 1, \dots, N$ ,  $t \in [0, T]$ :

$$(4.19) \quad \hat{\theta}_N = -\frac{\sum_{k=1}^N \nu_k \int_0^T (u_k(t) du_k(t) + \rho_k u_k^2 dt)}{\sum_{k=1}^N \nu_k^2 \int_0^T u_k^2(t) dt}.$$

Define  $J = \min\{k : \mu_n(\theta) > 0 \text{ for all } n \geq k \text{ and } \theta \in \Theta\}$ ; see (4.6).

**Theorem 4.8.** *Assume that equation (4.16) is diagonalizable and parabolic.*

(a) *The following conditions are equivalent:*

(1)

$$(4.20) \quad \sum_{k=J}^{\infty} \frac{\nu_k^2}{\mu_k(\theta)} = +\infty \text{ for all } \theta \in \Theta;$$

(2)

$$(4.21) \quad \lim_{N \rightarrow \infty} \hat{\theta}_N = \theta \text{ with probability one for all } \theta \in \Theta;$$

(3) *the measures  $\{\bar{\mathbf{P}}_T^\theta, \theta \in \Theta\}$  generated by the solutions of (4.16) in the space of continuous  $X$ -valued processes are mutually singular for different  $\theta$  (as in Theorem 4.6,  $X$  is a Hilbert space such that the embedding  $H \rightarrow X$  is Hilbert-Schmidt).*

(b) *If (4.20) holds, then*

$$(4.22) \quad \lim_{N \rightarrow \infty} \sqrt{\sum_{k=J}^N \frac{\nu_k^2}{\mu_k(\theta)}} (\hat{\theta}_N - \theta) \stackrel{d}{=} \mathcal{N}(0, 2/T).$$

*Proof:* (a) First, we show that (4.20) is equivalent to (4.21). By (4.17),

$$(4.23) \quad \widehat{\theta}_N - \theta = -\frac{\sum_{k=1}^N \int_0^T \nu_k u_k(s) dw_k(s)}{\sum_{k=1}^N \int_0^T \nu_k^2 u_k^2(s) ds};$$

both the top and the bottom on the right-hand side of (4.23) are sums of independent random variables. Next,

$$(4.24) \quad \mathbb{E} \int_0^T \nu_k^2 u_k^2(t) dt \sim \frac{T \nu_k^2}{2\mu_k(\theta)}$$

(see (2.3)), and

$$(4.25) \quad \text{Var} \left( \int_0^T \nu_k^2 u_k^2(t) dt \right) \sim \frac{T \nu_k^4}{2\mu_k^3(\theta)};$$

(see (2.5)). If (4.20) does not hold, then (4.24) implies

$$\sum_{k \geq 1} \mathbb{E} \int_0^T \nu_k^2 u_k^2(s) ds < \infty$$

and so the series

$$\sum_{k \geq 1} \int_0^T \nu_k^2 u_k^2(s) ds$$

converges with probability one. Therefore,

$$(4.26) \quad \lim_{N \rightarrow \infty} (\widehat{\theta}_N - \theta) = \frac{\int_0^T \langle \langle A_1 u(t), dW(t) \rangle \rangle dt}{\int_0^T \|A_1 u(t)\|_H^2 dt} \neq 0;$$

see (4.23) and (2.19).

On the other hand, if (4.20) holds, then

$$(4.27) \quad \sum_{n \geq J} \frac{\nu_n^2 \mu_n^{-1}(\theta)}{\left( \sum_{k=J}^n \nu_k^2 \mu_k^{-1}(\theta) \right)^2} < \infty.$$

Indeed, setting  $a_n = \nu_n^2 \mu_n^{-1}(\theta)$  and  $A_n = \sum_{k=J}^n a_k$ , we notice that

$$\sum_{n \geq J} \frac{a_n}{A_n^2} \leq \sum_{n \geq J+1} \left( \frac{1}{A_{n-1}} - \frac{1}{A_n} \right) = \frac{1}{A_J}.$$

Then the strong law of large numbers (Theorem 2.2), together with the equality

$$\mathbb{E} \int_0^T u_k(s) dw_k(s) = 0, \quad k \geq 1,$$

implies

$$\lim_{N \rightarrow \infty} \frac{\sum_{k=1}^N \int_0^T \nu_k u_k(s) dw_k(s)}{\sum_{k=1}^N \mathbb{E} \int_0^T \nu_k^2 u_k^2(s) ds} = 0 \quad \text{with probability one.}$$

Next, it follows from (4.27) and (4.6) that

$$(4.28) \quad \sum_{n \geq J} \frac{\nu_n^4 \mu_n^{-3}(\theta)}{\left( \sum_{k=J}^n \nu_k^2 \mu_k^{-1}(\theta) \right)^2} < \infty,$$

because, by Proposition 4.3,  $|\nu_k|/\mu_k(\theta) < c_0$  for  $k \geq J$ . Then another application of the strong law of large numbers (Corollary 2.3) shows that

$$(4.29) \quad \lim_{N \rightarrow \infty} \frac{\sum_{k=1}^N \int_0^T \nu_k^2 u_k^2(s) ds}{\sum_{k=1}^N \mathbb{E} \int_0^T \nu_k^2 u_k^2(s) ds} = 1$$

with probability one, and (4.21) follows. This completes the proof that (4.20) is equivalent to (4.21).

Next, we show that (4.20) is equivalent to singularity of measures  $\bar{\mathbf{P}}_T^\theta$ . Since  $u$  is a Gaussian process, the measures are either mutually absolutely continuous or mutually singular (Feldman [19] or Hájek [21], [22]), and, by a result of Koski and Loges [39, Proposition 1], the measures are mutually absolutely continuous if and only if the series  $\sum \nu_k^2/\mu_k(\theta)$  converges (see also Mikulevičius and Rozovskiĭ [54, Corollary 1] for a more general result about absolute continuity of measures).

(b) To prove (4.22), use (4.29) and apply Corollary 2.6 with  $f_k(t) = \nu_k u_k(t)$ . An interested reader can also verify that, in general, condition (4.20) is not enough to apply the classical Central Limit Theorem (Theorem 2.4).

This completes the proof of Theorem 4.8.  $\square$

**4.3. Discussion and examples.** First of all, let us formulate condition (4.20) in terms of the orders of the operators in the equation. Let  $A_0, A_1$  be elliptic differential or pseudo-differential operators, either on a smooth bounded domain in  $\mathbb{R}^d$  or on a smooth compact  $d$ -dimensional manifold, and let  $m_0, m_1$ , be the orders of  $A_0, A_1$  respectively, so that  $2m = \max(m_0, m_1)$ . Then, under rather general conditions we have

$$(4.30) \quad |\nu_k| \sim c_1 k^{m_1/d}, \quad \mu_k(\theta) \sim c(\theta) k^{2m/d}$$

for some positive numbers  $c_1, c(\theta)$ ; see, for example, Safarov and Vasiliev [69].

If (4.30) holds, then condition (4.20) becomes  $(2m_1 - 2m)/d \geq -1$  or

$$(4.31) \quad m_1 \geq m - \frac{d}{2},$$

which was established by Huebner and Rozovskiĭ [30]. On the other hand, Theorem 4.8 covers operators with more exotic eigenvalues, such as  $\nu_k = k \ln k$  or  $\nu_k = e^k$ ; such eigenvalues can appear in the problems of statistical inference based on information from many independent but not identical channels [38].

The reader can verify that the additional assumption  $\nu_k^2/\mu_k(\theta) \sim k^\beta$  for some  $\beta \geq -1$  simplifies the proof of Theorem 4.8 in at least two ways:

- (1) Relation (4.25) can be replaced with a less delicate bound using (2.4):

$$\text{Var} \left( \int_0^T \nu_k^2 u_k^2(t) dt \right) \leq \mathbb{E} \left( \int_0^T \nu_k^2 u_k^2(t) dt \right)^2 \leq \frac{C\nu_k^4}{\mu_k^2(\theta)}.$$

- (2) The classical Central Limit Theorem (Theorem 2.4) can be applied instead of a more sophisticated martingale version.

Next, we consider the effects of a non-zero initial condition. Even though it was assumed everywhere that  $u(0) = 0$ , Theorem 4.8 extends to nonzero initial condition  $u(0) = \phi$  as long as  $\phi$  is non-random and belongs to  $H$ . Indeed, the Fourier coefficients of the solution satisfy

$$u_k(t) = \phi_k e^{-\mu_k(\theta)t} + \int_0^t e^{-\mu_k(\theta)(t-s)} dw_k(s)$$

and, for  $k \geq J$ ,

$$\mathbb{E} \int_0^T u_k^2(t) dt = \frac{\phi_k^2}{2\mu_k(\theta)} (1 - e^{-2\mu_k(\theta)T}) + \frac{T}{2\mu_k(\theta)} - \frac{1}{4\mu_k^2(\theta)} (1 - e^{-2\mu_k(\theta)T}).$$

As a result, if

$$(4.32) \quad \sum_{k \geq 1} \phi_k^2 < \infty,$$

then (4.24) and (4.25) hold. The computations also show that

- (1) it is important to have  $\phi$  non-random: otherwise, the processes  $u_k$  will no longer be independent, and the analysis will become much more complicated;
- (2) condition (4.32) can be further relaxed, although the specifics will depend on the rate of growth of  $\nu_k$  and  $\mu_k(\theta)$ ;
- (3) if the series  $\sum_{k \geq 1} \phi_k^2$  diverges fast enough, then a consistent estimator is possible even if (4.20) does not hold.

The details are left to an interested reader (see also Huebner [23]).

Next, we discuss how the presence of the spatial covariance in the noise term affects the model.

Let us consider the equation

$$du + (A_0 + \theta A_1)u dt = dW^Q(t),$$

where  $Q$  is a positive linear self-adjoint operator. Then we can write  $Q = BB^T$  for some operator  $B$ , and the equation becomes

$$du + (A_0 + \theta A_1)u dt = B dW(t).$$

If  $B^{-1}$  exists, then we get back to the original model (4.16) by considering  $v = B^{-1}u$ ,  $\tilde{A}_0 = B^{-1}A_0B$ ,  $\tilde{A}_1 = B^{-1}A_1B$ :

$$dv(t) + (\tilde{A}_0 + \theta\tilde{A}_1)v(t) dt = dW(t),$$

provided this equation is diagonalizable and parabolic in the sense of Definitions 4.1 and 4.2.

If  $B^{-1}$  does not exist, there are two possibilities:

- (1)  $(u_0, h_i)_0 = 0$  for every  $i$  such that  $Bh_i = 0$ . In this case,  $u_i(t) = 0$  for all  $t > 0$ , so that we can factor out the kernel of  $B$  and reduce the problem to invertible  $B$ .
- (2)  $(u_0, h_i)_0 \neq 0$  for some  $i$  such that  $Bh_i = 0$ . In this case,  $u_i(t) = u_i(0)e^{-\rho_i t - \nu_i \theta t}$  and  $\theta$  is determined exactly from the observations of  $u_i(t)$ :

$$\theta = \frac{1}{\nu_i(t-s)} \ln \frac{u_i(s)}{u_i(t)} - \frac{\rho_i}{\nu_i}, \quad t \neq s.$$

Let us now look at some concrete examples of (4.16).

- (1) Consider equation

$$du - \theta u_{xx} dt = dW, \quad 0 < t < T, \quad x \in (0, 1), \quad \theta > 0,$$

with zero initial and boundary conditions.

Clearly,  $\nu_k = \pi^2 k^2$ ,  $\rho_k = 0$ ,  $\mu_k(\theta) = \theta \pi^2 k^2$ , and (4.20) holds. Then

$$u_k(t) = \int_0^T e^{-\theta \pi^2 k^2 (t-s)} dw_k,$$

$$\hat{\theta}_N = - \frac{\sum_{k=1}^N k^2 \int_0^T u_k(t) du_k(t)}{\sum_{k=1}^N \int_0^T \pi^2 k^4 u_k(t) dt},$$

$$\lim_{N \rightarrow \infty} N^{3/2} (\hat{\theta}_N - \theta) \stackrel{d}{=} \mathcal{N}(0, 6\theta / (\pi^2 T)),$$

where  $\mathcal{N}(a, \sigma^2)$  is a normal random variable with mean  $a$  and variance  $\sigma^2$ , and the convergence is in distribution.

More generally, if equation

$$du + \theta A_1 u dt = dW, \quad \theta > 0,$$

is diagonalizable and parabolic and  $\nu_k > 0$ , then  $\nu_k^2/\mu_k(\theta) = \nu_k/\theta$  and condition (4.20) is satisfied, so that  $\hat{\theta}_N$  is consistent and asymptotically normal:

$$\begin{aligned} u_k(t) &= \int_0^T e^{-\theta \nu_k(t-s)} dw_k, \\ \hat{\theta}_N &= -\frac{\sum_{k=1}^N \nu_k^2 \int_0^T u_k(t) du_k(t)}{\sum_{k=1}^N \int_0^T \nu_k^4 u_k(t) dt}, \\ \lim_{N \rightarrow \infty} \left( \sum_{k=1}^N \nu_k \right)^{1/2} (\hat{\theta}_N - \theta) &\stackrel{d}{=} \mathcal{N}(0, 2\theta/T). \end{aligned}$$

(2) Consider equation

$$du - (\Delta u + \theta u) dt = dW, \quad 0 < t < T, \quad x \in G \subset \mathbb{R}^d,$$

with zero initial and boundary conditions and  $d \geq 2$ . Denote by  $\lambda_k$ ,  $k \geq 1$ , the eigenvalues of the Laplace operator  $\Delta$ ; recall that  $\lambda_k < 0$ . Clearly,  $\nu_k = -1$ ,  $\rho_k = -\lambda_k$ ,  $\mu_k(\theta) = -\lambda_k - \theta$ . Then

$$\begin{aligned} u_k(t) &= \int_0^T e^{(\lambda_k + \theta)(t-s)} dw_k, \\ \hat{\theta}_N &= -\frac{\sum_{k=1}^N \int_0^T u_k(t) (du_k(t) - \lambda_k u_k(t) dt)}{\sum_{k=1}^N \int_0^T u_k^2(t) dt}, \\ \lim_{N \rightarrow \infty} \Phi_d(N) (\hat{\theta}_N - \theta) &\stackrel{d}{=} \mathcal{N}(0, \sigma_d^2), \end{aligned}$$

where

$$\Phi_d(N) = \begin{cases} \sqrt{\ln N}, & \text{if } d = 2, \\ N^{(d-2)/(2d)}, & \text{if } d > 2, \end{cases} \quad \sigma_d^2 = \begin{cases} \frac{2c_\circ}{T}, & \text{if } d = 2, \\ \frac{2c_\circ}{T} \left(1 - \frac{2}{d}\right), & \text{if } d > 2, \end{cases}$$

and  $c_\circ$  is from (4.10).



If  $d = 1$ , then (4.31) does not hold ( $m_1 = 0$ ,  $m = 1$ ) and  $\widehat{\theta}_N$  is not a consistent estimator of  $\theta$ .

(3) Consider the equation

$$du + (\Delta^2 u + \theta \Delta u) dt = dW, \quad 0 < t < T, \quad x \in G \subset \mathbb{R}^d,$$

with zero initial and boundary conditions:

$$u|_{t=0} = u|_{\partial G} = \Delta u|_{\partial G} = 0.$$

As before, denote by  $\lambda_k$  the eigenvalues of the Laplacian  $\Delta$ . Clearly,  $\nu_k = \lambda_k$ ,  $\rho_k = \lambda_k^2$ ,  $\mu_k(\theta) = \lambda_k^2 + \theta \lambda_k$ . Then, for every  $d \geq 1$ ,

$$u_k(t) = \int_0^T e^{-(\lambda_k^2 + \theta \lambda_k)(t-s)} dw_k,$$

$$\widehat{\theta}_N = - \frac{\sum_{k=1}^N \int_0^T \lambda_k u_k(t) (du_k(t) - \lambda_k^2 u_k(t) dt)}{\sum_{k=1}^N \int_0^T \lambda_k^2 u_k^2(t) dt},$$

$$\lim_{N \rightarrow \infty} \sqrt{N} (\widehat{\theta}_N - \theta) \stackrel{d}{=} \mathcal{N}(0, 2/T).$$

## 5. Further directions

The proof of Theorem 4.8 is the main objective of the current paper. There are certainly many other statistical problems that have been studied for stochastic parabolic equations, and below is a (partial) list of these problems. Surveys by Prakasa Rao [63], [65] can provide more details on some of the topics.

**5.1. Diagonalizable equations.** The bottom line is that any problem of statistical inference for a stable Ornstein-Uhlenbeck process has a potential for an extension to diagonalizable stochastic parabolic equations. On the other hand, note that the numerous interesting problems for the unstable Ornstein-Uhlenbeck process do not have similar extensions to stochastic parabolic equations because of the parabolicity condition: only finitely many Fourier coefficients of the solution can be unstable processes.

Here is a (partial) list of the corresponding results and references.

- (1) **Several unknown parameters:** Huebner [23], [24].
- (2) **Bayesian estimators and hypothesis testing:** Bishwal [12], [13], Prakasa Rao [59].
- (3) **Time-dependent drift:** Huebner and Lototsky [27], [28].

- (4) **Observations in discrete time and/or space:** Huebner [23], Markussen [53], Piterbarg and Rozovskiĭ [58], Prakasa Rao [60], [62], [64], [66].
- (5) **Optimal linear filtering of a time-dependent drift:** Lototsky [50].

Somewhat different types of diagonalizable equations have also been considered.

- (1) **Equations with multiplicative noise:** Cialenco and Lototsky [14].
- (2) **Equations driven by the cylindrical fractional Brownian motion:** Cialenco et al. [15], Prakasa Rao [67].

Equations with multiplicative noise are diagonalizable only if the noise has no spatial structure. The simplest example is

$$du = \theta u_{xx} dt + u dw(t), \quad 0 < x < \pi,$$

with zero boundary conditions, where  $w$  is a standard Brownian motion. Of course, it is no longer possible to assume that  $u(0, x) = 0$ . The Fourier coefficients  $u_k$ ,  $k \geq 1$ , are now Geometric Brownian motions driven by the *same* Brownian motion  $w$ , making the problem *extremely* singular from the statistical point of view. In particular, it was shown in [14] that the parameter  $\theta$  can be determined exactly and in *closed form* from just two Fourier coefficients. For example, if  $u_1(0) \neq 0$  and  $u_2(0) \neq 0$ , then

$$\theta = \frac{1}{3T} \ln \frac{u_1(T)u_1(0)}{u_2(T)u_1(0)}$$

for every  $T > 0$ .

To get a cylindrical fractional Brownian motion, the usual Brownian motions in (2.20) are replaced with independent fractional Brownian motions  $w^H$  having the same Hurst parameter  $H$ . When  $H > 1/2$ , many of the results about the resulting maximum likelihood estimator turn out the same. In particular, [67] studies the corresponding modification of equation (1.1), and [15] establishes an analogue of Theorem 4.8.

**5.2. General equations.** If equation is not diagonalizable, that is, the operators  $A_0$  and  $A_1$  do not have a common system of eigenfunctions, then the analysis of the estimation problem becomes substantially more complicated. While the result about singularity of measures is still valid under the condition (4.31), there is no natural family of regular models to consider. In fact, there are at least two possibilities:

- (1) **Galerkin approximations:** Huebner [23], [24], Huebner et al. [29].
- (2) **Finite-dimensional projections:** Lototsky and Rozovskiĭ [51], [52], Lototsky [49].

Both [29] and [51] consider the equation

$$(5.1) \quad du + (A_0 + \theta A_1)u dt = dW,$$

fix an orthonormal basis  $\{h_k, k \geq 1\}$  in  $H$ , and define  $\Pi^N$ , the orthogonal projection on the span of  $h_1, \dots, h_N$ , and  $W^N = (W_{h_1}, \dots, W_{h_N})$ . The equation is assumed parabolic in the usual sense of partial differential equations, which implies existence, uniqueness, and regularity of the solution similar to Theorem 4.6 (see, for example, [68, Chapter 3]; our Definition 4.2 of parabolicity is a particular case of the general definition applied to diagonalizable equations). According to Mikulevičius and Rozovskiĭ [54, Corollary 1], the measures generated by the solutions for different  $\theta$  are mutually absolutely continuous if and only if

$$\mathbb{E} \int_0^T \|A_1 u(t)\|_H^2 dt < \infty;$$

in the diagonalizable case, this is equivalent to

$$\sum_{k \geq 1} \int_0^T \nu_k^2 \mathbb{E} u_k^2(t) dt < \infty,$$

or (see (4.24))  $\sum_{k \geq J} \nu_k^2 / \mu_k(\theta) < \infty$ ,  $\theta \in \Theta$ .

If  $u$  is a solution of (5.1), the Galerkin approximation  $u^N$  of  $u$  is the solution of the equation

$$(5.2) \quad du^N + (\Pi^N A_0 + \theta \Pi^N A_1)u^N dt = dW^N,$$

while the projection  $\Pi^N u$  satisfies

$$d\Pi^N u(t) + (\Pi^N A_0 + \theta \Pi^N A_1)u dt = dW^N,$$

which is not an equation for  $\Pi^N u$  unless  $\Pi^N$  commutes with  $A_0$  and  $A_1$ . Note that  $u^N = \Pi^N u$  in the diagonalizable case if the basis  $\{h_k, k \geq 1\}$  is the common system of eigenfunctions of  $A_0, A_1$ . The estimator suggested in [29] is

$$(5.3) \quad \hat{\theta}^N = \frac{\int_0^T \left( \Pi^N A_1 u^N(t), du^N(t) - \Pi^N A_0 u^N(t) dt \right)_H}{\int_0^T \|\Pi^N A_1 u^N(t)\|_H^2 dt},$$

which is the maximum likelihood estimator for the absolutely continuous statistical model generated by (5.2). The multi-parameter case is considered in [24].

The estimator suggested in [51] is

$$(5.4) \quad \hat{\theta}^N = \frac{\int_0^T (\Pi^N A_1 u(t), d\Pi^N u(t) - \Pi^N A_0 u(t) dt)_H}{\int_0^T \|\Pi^N A_1 u(t)\|_H^2 dt},$$

which is not a maximum likelihood estimator. Under the order condition (4.31), the infinite-dimensional model (5.1) is singular, and, as  $N \rightarrow \infty$ , consistency (in probability) and asymptotic normality hold for both (5.3) and (5.4). The paper [52] is a shorter version of [51], and [49] studies estimators of the type (5.4) for the two-parameter estimation problem

$$du + (\theta_0 A_0 + \theta_1 A_1)u dt = dW.$$

**5.3. Non-spectral methods.** Similar to finite-dimensional models, parameter estimation in stochastic parabolic equations can be studied in long-time or small-noise asymptotics. The interesting situation is when the infinite-dimensional problem is absolutely continuous, and in this situation

- (1) Estimators in the large-time asymptotic were studied by Loges [39], [48].
- (2) Estimators in the small-noise asymptotic were studied by Huebner [25], Ibragimov and Khas'minskiĭ [32], [33], [34], [35], and Prakasa Rao [61].

A different class of problems is a combination of filtering and estimation, when the observations are

$$(5.5) \quad y(t) = \int_0^s B u(s) ds + v(t),$$

where  $B$  is a operator with a fixed finite-dimensional range  $\mathbb{R}^n$  and  $v$  is a  $\mathbb{R}^n$ -valued Brownian motion independent of  $W$ . From the statistical point of view, this problem is always absolutely continuous, as the measures generated in  $\mathcal{C}((0, T); \mathbb{R}^n)$  by  $y$  are absolutely continuous with respect to the Wiener measure (Liptser and Shiryaev [46, Theorem 7.4]).

A number of papers by Aihara [1], [2], [4], [5], [6], [10] investigate the non-parametric estimation problem with observations (5.5), when  $u$  is a solution of

$$du = \sum_{i,j=1}^d \frac{\partial}{\partial x_i} \left( \theta(x) \frac{\partial u}{\partial x_j} \right) dt + dW^Q(t),$$

under various smoothness assumptions on  $\theta$ . A similar problem was also studied for hyperbolic equations [3], [7].

### Acknowledgements

The work was partially supported by the NSF Grant DMS-0803378. The author is grateful to Igor Cialenco for reading an earlier version of the paper and making many useful comments.

### References

- [1] S. I. AIHARA, Parameter identification for stochastic parabolic systems, in: “*Systems and control*”, Mita, Tokyo, 1991, pp. 1–12.
- [2] S. I. AIHARA, Regularized maximum likelihood estimate for an infinite-dimensional parameter in stochastic parabolic systems, *SIAM J. Control Optim.* **30(4)** (1992), 745–764.
- [3] S. I. AIHARA, Maximum likelihood estimate for discontinuous parameter in stochastic hyperbolic systems. White noise models and stochastic systems, (Enschede, 1992), *Acta Appl. Math.* **35(1–2)** (1994), 131–151.
- [4] S. I. AIHARA, Consistency property of extended least-squares parameter estimation for stochastic diffusion equation, *Systems Control Lett.* **34(5)** (1998), 249–256.
- [5] S. I. AIHARA, Identification of a discontinuous parameter in stochastic parabolic systems, *Appl. Math. Optim.* **37(1)** (1998), 43–69.
- [6] S. I. AIHARA AND A. BAGCHI, Infinite-dimensional parameter identification for stochastic parabolic systems, *Statist. Probab. Lett.* **8(3)** (1989), 279–287.
- [7] S. I. AIHARA AND A. BAGCHI, Parameter identification for hyperbolic stochastic systems, *J. Math. Anal. Appl.* **160(2)** (1991), 485–499.
- [8] S. I. AIHARA AND A. BAGCHI, Stochastic hyperbolic dynamics for infinite-dimensional forward rates and option pricing, *Math. Finance* **15(1)** (2005), 27–47.
- [9] S. I. AIHARA AND A. BAGCHI, Parameter estimation of parabolic type factor model and empirical study of US treasury bonds, in: “*System modeling and optimization*”, IFIP Int. Fed. Inf. Process. **199**, Springer, New York, 2006, pp. 207–217.
- [10] S. I. AIHARA AND Y. SUNAHARA, Identification of an infinite-dimensional parameter for stochastic diffusion equations, *SIAM J. Control Optim.* **26(5)** (1988), 1062–1075.

- [11] Y. AĪT-SAHALIA, Likelihood inference for diffusions: a survey, in: “*Frontiers in statistics*”, Imp. Coll. Press, London, 2006, pp. 369–405.
- [12] J. P. N. BISHWAL, Bayes and sequential estimation in Hilbert space valued stochastic differential equations, *J. Korean Statist. Soc.* **28(1)** (1999), 93–106.
- [13] J. P. N. BISHWAL, The Bernstein-von Mises theorem and spectral asymptotics of Bayes estimators for parabolic SPDEs, *J. Aust. Math. Soc.* **72(2)** (2002), 287–298.
- [14] IG. CIALENCO AND S. V. LOTOTSKY, Parameter estimation in diagonalizable bilinear stochastic parabolic equations, <http://arxiv.org/abs/0709.1135>, *Stat. Inference Stoch. Process.* (to appear).
- [15] IG. CIALENCO, S. V. LOTOTSKY, AND J. POSPÍŠIL, Asymptotic properties of the maximum likelihood estimator for stochastic parabolic equations with additive fractional Brownian motion, <http://arxiv.org/abs/0804.0407>, *Stoch. Dyn.* (to appear).
- [16] R. CONT, Modeling term structure dynamics: an infinite dimensional approach, *Int. J. Theor. Appl. Finance* **8(3)** (2005), 357–380.
- [17] D. A. DAWSON, Qualitative behavior of geostochastic systems, *Stochastic Process. Appl.* **10(1)** (1980), 1–31.
- [18] S. S. DE, Stochastic model of population growth and spread, *Bull. Math. Biol.* **49(1)** (1987), 1–11.
- [19] J. FELDMAN, Equivalence and perpendicularity of Gaussian processes, *Pacific J. Math.* **8** (1958), 699–708.
- [20] C. FRANKIGNOUL, SST anomalies, planetary waves and RC in the middle rectitudes, *Rev. Geophys.* **23(4)** (1985), 357–390.
- [21] J. HÁJEK, On a property of normal distribution of any stochastic process, (Russian), *Czechoslovak Math. J.* **8(83)** (1958), 610–618.
- [22] J. HÁJEK, A property of  $J$ -divergences of marginal probability distributions, *Czechoslovak Math. J.* **8(83)** (1958), 460–463.
- [23] M. HUEBNER, Parameter Estimation for SPDE’s, PhD thesis, University of Southern California, Los Angeles, CA (1993).
- [24] M. HUEBNER, A characterization of asymptotic behaviour of maximum likelihood estimators for stochastic PDE’s, *Math. Methods Statist.* **6(4)** (1997), 395–415.
- [25] M. HUEBNER, Asymptotic properties of the maximum likelihood estimator for stochastic PDEs disturbed by small noise, *Stat. Inference Stoch. Process.* **2(1)** (1999), 57–68 (2000).
- [26] M. HUEBNER, R. Z. KHAS’MINSKIĀ, AND B. L. ROZOVSKIĀ, Two examples of parameter estimation for stochastic partial differential

- equations, in: “*Stochastic processes*”, Springer, New York, 1993, pp. 149–160.
- [27] M. HUEBNER AND S. V. LOTOTSKY, Asymptotic analysis of the sieve estimator for a class of parabolic SPDEs, *Scand. J. Statist.* **27(2)** (2000), 353–370.
- [28] M. HUEBNER AND S. V. LOTOTSKY, Asymptotic analysis of a kernel estimator for parabolic SPDE’s with time-dependent coefficients, *Ann. Appl. Probab.* **10(4)** (2000), 1246–1258.
- [29] M. HUEBNER, S. V. LOTOTSKY, AND B. L. ROZOVSKIĬ, Asymptotic properties of an approximate maximum likelihood estimator for stochastic PDEs, in: “*Statistics and control of stochastic processes*” (Moscow, 1995/1996), World Sci. Publ., River Edge, NJ, 1997, pp. 139–155.
- [30] M. HUEBNER AND B. L. ROZOVSKIĬ, On asymptotic properties of maximum likelihood estimators for parabolic stochastic PDE’s, *Probab. Theory Related Fields* **103(2)** (1995), 143–163.
- [31] I. A. IBRAGIMOV AND R. Z. KHAS’MINSKIĬ, “*Statistical estimation. Asymptotic theory*”, Translated from the Russian by Samuel Kotz, Applications of Mathematics **16**, Springer-Verlag, New York-Berlin, 1981.
- [32] I. A. IBRAGIMOV AND R. Z. KHAS’MINSKIĬ, Some estimation problems in infinite-dimensional Gaussian white noise, in: “*Festschrift for Lucien Le Cam*”, Springer, New York, 1997, pp. 259–274.
- [33] I. A. IBRAGIMOV AND R. Z. KHAS’MINSKIĬ, Problems of estimating the coefficients of stochastic partial differential equations. I, (Russian), *Teor. Veroyatnost. i Primenen.* **43(3)** (1998), 417–438; translation in: *Theory Probab. Appl.* **43(3)** (1999), 370–387.
- [34] I. A. IBRAGIMOV AND R. Z. KHAS’MINSKIĬ, Problems of estimating the coefficients of stochastic partial differential equations. II, (Russian), *Teor. Veroyatnost. i Primenen.* **44(3)** (1999), 526–554; translation in: *Theory Probab. Appl.* **44(3)** (2000), 469–494.
- [35] I. A. IBRAGIMOV AND R. Z. KHAS’MINSKIĬ, Problems of estimating the coefficients of stochastic partial differential equations. III, (Russian), *Teor. Veroyatnost. i Primenen.* **45(2)** (2000), 209–235; translation in: *Theory Probab. Appl.* **45(2)** (2001), 210–232.
- [36] J. JACOD AND A. N. SHIRYAEV, “*Limit theorems for stochastic processes*”, second edition, Grundlehren der Mathematischen Wissenschaften **288**, Springer-Verlag, Berlin, 2003.
- [37] R. Z. KHAS’MINSKIĬ, N. KRYLOV, AND N. MOSHCHUK, On the estimation of parameters for linear stochastic differential equations, *Probab. Theory Related Fields* **113(3)** (1999), 443–472.

- [38] A. KOROSTELEV AND G. YIN, Estimation of jump points in high-dimensional diffusion modulated by a hidden Markov chain, *Math. Methods Statist.* **15(1)** (2006), 88–102.
- [39] T. KOSKI AND W. LOGES, Asymptotic statistical inference for a stochastic heat flow problem, *Statist. Probab. Lett.* **3(4)** (1985), 185–189.
- [40] S. M. KOZLOV, Equivalence of measures in Itô’s linear partial differential equations, (Russian), *Vestnik Moskov. Univ. Ser. I Mat. Meh.* **4** (1977), 47–52.
- [41] S. M. KOZLOV, Some questions of stochastic partial differential equations, (Russian), *Trudy Sem. Petrowsk.* **4** (1978), 147–172.
- [42] H. KUNITA, “*Stochastic flows and stochastic differential equations*”, Cambridge Studies in Advanced Mathematics **24**, Cambridge University Press, Cambridge, 1990.
- [43] YU. A. KUTOYANTS, “*Identification of dynamical systems with small noise*”, Mathematics and its Applications **300**, Kluwer Academic Publishers Group, Dordrecht, 1994.
- [44] YU. A. KUTOYANTS, “*Statistical inference for ergodic diffusion processes*” Springer Series in Statistics, Springer-Verlag London, Ltd., London, 2004.
- [45] R. SH. LIPTSER AND A. N. SHIRYAEV, “*Theory of martingales*”, Translated from the Russian by K. Dzjaparidze, Mathematics and its Applications (Soviet Series) **49**, Kluwer Academic Publishers Group, Dordrecht, 1989.
- [46] R. SH. LIPTSER AND A. N. SHIRYAEV, “*Statistics of random processes. I. General theory*”, Translated from the 1974 Russian original by A. B. Aries, second, revised and expanded edition, Applications of Mathematics (New York) **5**, Stochastic Modelling and Applied Probability, Springer-Verlag, Berlin, 2001.
- [47] R. SH. LIPTSER AND A. N. SHIRYAEV, “*Statistics of random processes. II. Applications*”, Translated from the 1974 Russian original by A. B. Aries, second, revised and expanded edition, Applications of Mathematics (New York) **6**, Stochastic Modelling and Applied Probability, Springer-Verlag, Berlin, 2001.
- [48] W. LOGES, Girsanov’s theorem in Hilbert space and an application to the statistics of Hilbert space-valued stochastic differential equations, *Stochastic Process. Appl.* **17(2)** (1984), 243–263.
- [49] S. V. LOTOTSKY, Parameter estimation for stochastic parabolic equations: asymptotic properties of a two-dimensional projection-based estimator, *Stat. Inference Stoch. Process.* **6(1)** (2003), 65–87.



- [50] S. V. LOTOTSKY, Optimal filtering of stochastic parabolic equations, in: “*Recent developments in stochastic analysis and related topics*”, World Sci. Publ., Hackensack, NJ, 2004, pp. 330–353.
- [51] S. V. LOTOTSKY AND B. L. ROZOVSKIĬ, Spectral asymptotics of some functionals arising in statistical inference for SPDEs, *Stochastic Process. Appl.* **79**(1) (1999), 69–94.
- [52] S. V. LOTOTSKY AND B. L. ROZOVSKIĬ, Parameter estimation for stochastic evolution equations with non-commuting operators, in: “*Skorokhod’s Ideas in Probability Theory*” (V. Korolyuk, N. Portenko, and H. Syta, eds.), Institute of Mathematics of the National Academy of Sciences of Ukraine, Kiev, Ukraine, 2000, pp. 271–280.
- [53] B. MARKUSSEN, Likelihood inference for a discretely observed stochastic partial differential equation, *Bernoulli* **9**(5) (2003), 745–762.
- [54] R. MIKULEVIČIUS AND B. L. ROZOVSKIĬ, Uniqueness and absolute continuity of weak solutions for parabolic SPDEs. White noise models and stochastic systems, (Enschede, 1992), *Acta Appl. Math.* **35**(1–2) (1994), 179–192.
- [55] R. MIKULEVIČIUS AND B. L. ROZOVSKIĬ, Martingale problems for stochastic PDE’s, in: “*Stochastic partial differential equations: six perspectives*”, Math. Surveys Monogr. **64**, Amer. Math. Soc., Providence, RI, 1999, pp. 243–325.
- [56] D. NUALART, “*The Malliavin calculus and related topics*”, second edition, Probability and its Applications (New York), Springer-Verlag, Berlin, 2006.
- [57] L. PITERBARG AND B. L. ROZOVSKIĬ, Maximum likelihood estimators in the equations of physical oceanography, in: “*Stochastic modelling in physical oceanography*”, Progr. Probab. **39**, Birkhäuser Boston, Boston, MA, 1996, pp. 397–421.
- [58] L. PITERBARG AND B. L. ROZOVSKIĬ, On asymptotic problems of parameter estimation in stochastic PDE’s: discrete time sampling, *Math. Methods Statist.* **6**(2) (1997), 200–223.
- [59] B. L. S. PRAKASA RAO, Bayes estimation for some stochastic partial differential equations, Prague Workshop on Perspectives in Modern Statistical Inference: Parametrics, Semi-parametrics, Non-parametrics (1998), *J. Statist. Plann. Inference* **91**(2) (2000), 511–524.
- [60] B. L. S. PRAKASA RAO, Estimation for some stochastic partial differential equations based on discrete observations, *Calcutta Statist. Assoc. Bull.* **50**(199–200) (2000), 193–206.

- [61] B. L. S. PRAKASA RAO, Nonparametric inference for a class of stochastic partial differential equations. II, *Stat. Inference Stoch. Process.* **4(1)** (2001), 41–52.
- [62] B. L. S. PRAKASA RAO, Nonparametric inference for parabolic stochastic partial differential equations, *Random Oper. Stochastic Equations* **9(4)** (2001), 329–338.
- [63] B. L. S. PRAKASA RAO, Statistical inference for stochastic partial differential equations, in: “*Selected Proceedings of the Symposium on Inference for Stochastic Processes*” (Athens, GA, 2000), IMS Lecture Notes Monogr. Ser. **37**, Inst. Math. Statist., Beachwood, OH, 2001, pp. 47–70.
- [64] B. L. S. PRAKASA RAO, Nonparametric inference for a class of stochastic partial differential equations based on discrete observations, *Sankhyā Ser. A* **64(1)** (2002), 1–15.
- [65] B. L. S. PRAKASA RAO, On some problems of estimation for some stochastic partial differential equations, in: “*Uncertainty and optimality*”, World Sci. Publ., River Edge, NJ, 2002, pp. 71–153.
- [66] B. L. S. PRAKASA RAO, Estimation for some stochastic partial differential equations based on discrete observations. II, *Calcutta Statist. Assoc. Bull.* **54(215–216)** (2003), 129–141.
- [67] B. L. S. PRAKASA RAO, Parameter estimation for some stochastic partial differential equations driven by infinite dimensional fractional Brownian motion, *Theory Stoch. Process.* **10(3–4)** (2004), 116–125.
- [68] B. L. ROZOVSKIĬ, “*Stochastic evolution systems. Linear theory and applications to nonlinear filtering*”, Translated from the Russian by A. Yarkho, Mathematics and its Applications (Soviet Series) **35**, Kluwer Academic Publishers Group, Dordrecht, 1990.
- [69] YU. SAFAROV AND D. VASSILIEV, “*The asymptotic distribution of eigenvalues of partial differential operators*”, Translated from the Russian manuscript by the authors, Translations of Mathematical Monographs **155**, American Mathematical Society, Providence, RI, 1997.
- [70] S. E. SERRANO AND G. ADOMIAN, New contributions to the solution of transport equations in porous media, *Math. Comput. Modelling* **24(4)** (1996), 15–25.
- [71] S. E. SERRANO AND T. E. UNNY, Random evolution equations in hydrology, *Appl. Math. Comput.* **38(3)** (1990), 201–226.
- [72] A. N. SHIRYAEV, “*Probability*”, Translated from the first (1980) Russian edition by R. P. Boas, second edition, Graduate Texts in Mathematics **95**, Springer-Verlag, New York, 1996.

- [73] M. A. SHUBIN, “*Pseudodifferential operators and spectral theory*”, Translated from the 1978 Russian original by Stig I. Andersson, second edition, Springer-Verlag, Berlin, 2001.
- [74] J. B. WALSH, An introduction to stochastic partial differential equations, in: “*École d’été de probabilités de Saint-Flour, XIV-1984*”, Lecture Notes in Math. **1180**, Springer, Berlin, 1986, pp. 265–439.
- [75] K. YOSIDA, “*Functional analysis*”, sixth edition, Grundlehren der Mathematischen Wissenschaften **123**, Springer-Verlag, Berlin-New York, 1980.

*Current address:*

Department of Mathematics  
University of Southern California  
Los Angeles, CA 90089  
USA

*E-mail address:* [lototsky@math.usc.edu](mailto:lototsky@math.usc.edu)

*URL:* <http://www-rcf.usc.edu/~lototsky>

Rebut el 25 de juliol de 2008.