PARAMETER ESTIMATION IN HYPERBOLIC MULTICHANNEL MODELS

W. LIU AND S. V. LOTOTSKY

Abstract. A multichannel model is considered, with each channel represented by a linear second-order stochastic equation with two unknown coefficients. The channels are interpreted as the Fourier coefficients of the solution of a stochastic hyperbolic equation with possibly unbounded damping. The maximum likelihood estimator of the coefficients is constructed using the information from a finite number of channels. Necessary and sufficient conditions are determined for the consistency of the estimator as the number of channels increases, while the observation time and noise intensity remain fixed.

1. Introduction

Consider the wave equation with zero initial and boundary conditions and the space-time Gaussian white noise $\dot{W}(t,x)$ as the driving force:

$$\begin{aligned} &u_{tt}(t,x) = \theta u_{xx}(t,x) + \dot{W}(t,x),\ 0 < t \leq T,\ 0 < x < 1,\\ &u(0,x) = u_t(0,x) = 0,\ u(t,0) = u(t,1) = 0; \end{aligned} \tag{1.1}$$

subscripts denote partial derivatives: $u_{tt} = \partial^2 u/\partial t^2$, etc. Is it possible to find a consistent estimate of the parameter $\theta > 0$ from the observations of $u(t,x),\ t \in [0,T],\ x \in [0,1]$?

A positive answer to this question is obtained in [18] as follows. Consider the Fourier series expansion of the solution of (1.1):

$$u(t,x) = \sqrt{\frac{2}{\pi}} \sum_{k \geq 1} u_k(t)\ \sin(\pi k x). \tag{1.2}$$

Then define $v_k(t) = du_k(t)/dt$ and

$$\widehat{\theta}_N = -\frac{\sum_{k=1}^N \int_0^T k^2 u_k(t) dv_k(t)}{\pi^2 \sum_{k=1}^N \int_0^T k^4 u_k^2(t) dt}. \tag{1.3}$$

It is proved in [18] that $\lim_{N\to\infty} \widehat{\theta}_N = \theta$ with probability one and

$$\lim_{N\to\infty} N^{3/2}(\widehat{\theta}_N - \theta) = \zeta$$

in distribution, where ζ is a Gaussian random variable with zero mean and variance $12\pi^2\theta/T^2$.

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

Key words and phrases. Cylindrical Brownian motion, Second-Order Stochastic Equations, Stochastic Hyperbolic Equations.

At a first glance, there is nothing unusual about the result: we get a consistent and asymptotically normal estimator of an unknown parameter. A closer look indicates that the result is, in fact, rather unusual and leads to an interesting class of singular statistical models. Note that the asymptotic parameter N is not a part of the original problem and is introduced as a part of the solution. In other words, a consistent estimation of the parameter is possible with both the observation time interval and the amplitude of the noise fixed. This suggests that the estimation model defined by (1.1) is not a regular statistical model, and indeed, one can see from [21, Corollary 1] that the measures in a suitable Hilbert space generated by the solutions of (1.1) for different values of θ are mutually singular. As a result, estimation of the *drift* parameter θ in (1.1) is somewhat similar to estimation of the *diffusion* parameter σ in the stochastic ordinary differential equation $dX(t) = aX(t)dt + \sigma dW(t)$.

Generalization of (1.1) and (1.3), which is the objective of the current work, relies on the observation that the Fourier coefficients u_k of the solution are independent processes satisfying

$$du_k(t) = v_k(t)dt,\ dv_k(t) = -\pi^2 k^2 \theta u_k(t)dt + dw_k(t),$$

or, less formally,

$$\ddot{u}_k(t) + \pi^2 k^2 \theta u_k(t) = \dot{w}_k, \tag{1.4}$$

where $\dot{u}_k = du_k/dt$ and $w_k,\ k \geq 1$ are independent standard Brownian motions. Another key observation is that

$$\lim_{k\to\infty} k^2 \mathbb{E} \int_0^T u_k^2(t)dt = \frac{T^2}{4\pi^2\theta}. \tag{1.5}$$

The fact that the model is defined by a stochastic partial differential is not as important as the fact that the Fourier coefficients of the solution are independent processes with the property (1.5). Accordingly, a natural generalization of (1.1) is a multichannel model with independent channels $\{u_k,\ k \geq 1\}$ defined by

$$\begin{aligned} &\ddot{u}_k(t) - \mu_k(\theta_2)\dot{u}(t) + \lambda_k(\theta_1)u_k(t) = \dot{w}_k(t),\ 0 < t \leq T, \\ &u_k(0) = \dot{u}_k(0) = 0, \end{aligned} \tag{1.6}$$

where λ_k and μ_k are two sequences of numbers depending on two unknown parameters θ_1 and θ_2. The objective is to construct maximum likelihood estimators of the parameters $\theta_1,\ \theta_2$ given $u_k(t),\ k = 1, \dots, N,\ t \in [0, T]$, and to study the asymptotic properties of the estimators as $N \to \infty$. With little loss of generality, we assume linear dependence of λ_k and μ_k on the unknown parameters:

$$\lambda_k = \kappa_k + \theta_1 \tau_k, \quad \mu_k = \rho_k + \theta_2 \nu_k. \tag{1.7}$$

The analysis of (1.6) in full generality cannot be a subject of a single paper because there are 11 different regimes for the solution of the equation $\ddot{y}(t) - b\dot{y}(t) + ay(t) = 0$, depending on the roots of the characteristic equation $r^2 - 2br + a = 0$. As a result, we restrict our analysis to what we call *hyperbolic* multichannel models, that is, the models in which the channels u_k can be interpreted as the Fourier coefficients of a (broadly defined) stochastic hyperbolic equation. Roughly speaking, this means that $\lambda_k \uparrow +\infty$ and μ_k do not go to $-\infty$ too quickly (alternatively, the free motion of (1.6) does not grow exponentially for all sufficiently large k). For a precise formulation, we need a short digression into the subject of partial differential equations.

A typical example of a parabolic equation is the heat equation

$$u_t = u_{xx};$$

a typical example of a hyperbolic equation is the wave equation

$$u_{tt} = u_{xx}.$$

In a more abstract setting, if $\mathcal{A}$ is linear operator such that $\dot{u} + \mathcal{A}u = 0$ is a parabolic equation, then

$$\ddot{u} + \mathcal{A}u = 0$$

is natural to call a hyperbolic equation; $\dot{u}$ and $\ddot{u}$ are the first and second time derivatives of u.

Damping in a hyperbolic equations is introduced via a term depending on the first time derivative of the solution. For example, a damped wave equation is

$$u_{tt} = u_{xx} - au_t,\ a > 0.$$

Indeed, if we define the total energy $E(t) = \int \big(u_t^2(t,x) + u_x^2(t,x)\big)dx$, then integration by parts shows that

$$\frac{d}{dt}E(t) = -a\int u_t^2(t,x)dx;$$

it also shows that $a < 0$ (negative damping) corresponds to *amplification.* More generally, we write a damped linear hyperbolic equation in an abstract form

$$\ddot{u} + \mathcal{A}u = \mathcal{B}\dot{u}, \tag{1.8}$$

where $\mathcal{A}$ and $\mathcal{B}$ are linear operators on a separable Hilbert space H; depending on the properties of the operator $\mathcal{B}$, the result can be either damping or amplification.

In this formulation, the multichannel model (1.6) becomes a stochastic version of (1.8):

$$\ddot{u} + \mathcal{A}(\theta_1)u = \mathcal{B}(\theta_2)\dot{u} + \dot{W}, \qquad 0 < t \le T, \tag{1.9}$$

where $\dot{W}$ is Gaussian space-time white noise on H and $\mathcal{A}(\theta_1)$ and $\mathcal{B}(\theta_2)$ are linear operators defined by

$$\mathcal{A}(\theta_1)h_k = \lambda_k(\theta_1)h_k,\ \mathcal{B}(\theta_2)h_k = \mu_k(\theta_2)h_k;$$

$\{h_k,\ k \ge 1\}$ is an orthonormal basis in H, $u = \sum_{k\ge1} u_k h_k$. Following (1.7), we further write $\mathcal{A}(\theta_1) = \mathcal{A}_0 + \theta_1\mathcal{A}_1$, $\mathcal{B}(\theta_2) = \mathcal{B}_0 + \theta_2\mathcal{B}_1$, where

$$\mathcal{A}_0 h_k = \kappa_k h_k, \quad \mathcal{B}_0 h_k = \rho_k h_k, \quad \mathcal{A}_1 h_k = \tau_k h_k, \quad \mathcal{B}_1 h_k = \nu_k h_k. \tag{1.10}$$

The eigenvalues of partial differential operators have *algebraic* asymptotic (see [23]), and the multi-channel formulation could be a more natural starting point than (1.9) when, for example, $\tau_k = e^k$ or $\nu_k = \ln\ln(k+2)$. The equivalent operator formulation (1.9) could still be useful as it represents the multichannel model as an abstract *linear infinite-dimensional stochastic evolution system.* The system generates a Gaussian measure in a suitable Hilbert space, and then identifiability of the parameters is equivalent to singularity of the corresponding infinite-dimensional statistical model.

It is this formulation of the estimation problem in the language of stochastic partial differential equations (SPDEs) that was first suggested in [6] and further investigated in [7, 20] for stochastic parabolic equations with one unknown parameter. Estimation of several parameters in parabolic equations has also been studied [5, 19].

When the underlying infinite-dimensional statistical model is regular, various estimation problems have been studied in both SPDE [3, 9, 10, 11] and multichannel [4, 12, 13, 22] settings. Finite-dimensional singular statistical problems have been studied in [8, 14, 15]. Nevertheless, analysis of (1.5) requires completely different tools and leads to new results.

Let us summarize the main results of the current paper. Consider a separable Hilbert space H with an orthonormal basis $\{h_k,\ k \geq 1\}$. Let X be a Hilbert space such that $H \subset X$ and the embedding operator $\jmath:\ H \to X$ is Hilbert-Schmidt: $\sum_{k\geq 1} \|h_k\|_X^2 < \infty$. Then the *cylindrical Brownian motion* $W(t) = \sum_{k\geq 1} w_k(t)h_k$ is an X-valued process. We say that the multi-channel model (1.6) is hyperbolic if, for each t, the infinite-dimensional process $u(t) = \sum_{k\geq 1} u_k(t)h_k$ is an element of X; this is known to be the case when (1.9) is a hyperbolic equation in the usual sense (see, for example, [2, Section 6.8]). In the parametric setting, it is important to have hyperbolicity uniformly for $\theta_1 \in \Theta_1$, $\theta_2 \in \Theta_2$ on compact sub-sets Θ_1, Θ_2 of $\mathbb{R}$. The main result of Section 2 is that (1.6) is uniformly hyperbolic if and only if

(1) there exist positive numbers C^*, c_1, c_2 such that $\{\kappa_k + \theta\tau_k + C^*,\ k \geq 1\}$ is a positive, non-decreasing, and unbounded sequence for all $\theta \in \Theta_1$ and

$$c_1 \leq \frac{\kappa_k + \theta\tau_k + C^*}{\kappa_k + \theta'\tau_k + C^*} \leq c_2 \tag{1.11}$$

for all $\theta, \theta' \in \Theta_1$;

(2) there exist positive numbers C, J such that, for all $k \geq J$ and all $\theta_1 \in \Theta_1$, $\theta_2 \in \Theta_2$,

$$T(\rho_k + \theta_2\nu_k) \leq \ln(\kappa_k + \theta_1\tau_k) + C. \tag{1.12}$$

The maximum likelihood estimators of θ_1, θ_2 based on u_k and $\dot{u}_k$, $k = 1, \ldots, N$ are constructed explicitly in Section 3, but the corresponding formulas are too large to present in the Introduction. Analysis of these estimators in the limit $N \to \infty$ is carried out in Sections 4 and 5. Here is the main result when $\mathcal{A}_i, \mathcal{B}_i$ in the equivalent formulation (1.9) are differential or pseudo-differential elliptic operators. Every such operator has a well-defined order; for a partial differential operator, it is the order of the highest derivative.

Theorem 1.1. *Assume that $\mathcal{A}_i, \mathcal{B}_i$ are positive-definite, self-adjoint elliptic differential or pseudo-differential operators, either on a smooth bounded domain in $\mathbb{R}^d$ with suitable boundary conditions, or on a smooth compact d-dimensional manifold. Then*

(1) *the maximum likelihood estimator of θ_1 is consistent and asymptotically normal in the limit $N \to \infty$ if and only if*

$$\operatorname{order}(\mathcal{A}_1) \geq \frac{\operatorname{order}(\mathcal{A}_0 + \theta_1\mathcal{A}_1) + \operatorname{order}(\mathcal{B}_0 + \theta_2\mathcal{B}_1) - d}{2}; \tag{1.13}$$

(2) *the maximum likelihood estimator of* θ_2 *is consistent and asymptotically normal in the limit* $N \to \infty$ *if and only if*

$$\text{order}(\mathcal{B}_1) \geq \frac{\text{order}(\mathcal{B}_0 + \theta_2\mathcal{B}_1) - d}{2}. \tag{1.14}$$

Similar to the parabolic case (Huebner [5]), this result extends to a multichannel model with multiple parameters, which, in the equivalent operator formulation can be written as

$$\ddot{u} + \sum_{i=0}^{n} \theta_{1i}\mathcal{A}_i u = \sum_{j=0}^{m} \theta_{2j}\mathcal{B}_j\dot{u} + \dot{W}.$$

For example, the coefficient θ_{1p} can be consistently estimated if and only if

$$\text{order}(\mathcal{A}_p) \geq \frac{\text{order}\left(\sum_{i=0}^{n} \theta_{1i}\mathcal{A}_i\right) + \text{order}\left(\sum_{j=0}^{m} \theta_{2j}\mathcal{B}_j\right) - d}{2}.$$

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 ξ, $\mathbb{E}\xi$ and $\text{Var}\,\xi$ denote the expectation and variance respectively. The time derivative of a function is denote either by a dot on top (as in $\dot{u}$) or by a subscript t (as in u_t).

The following notations are used for two non-negative sequences a_n, b_n:

$$a_n \bowtie b_n \tag{1.15}$$

if there exist positive numbers c_1, c_2 such that $c_1 \leq a_n/b_n \leq c_2$ for all sufficiently large n;

$$a_n \asymp b_n \tag{1.16}$$

if

$$\lim_{n\to\infty} \frac{a_n}{b_n} = c \quad \text{for some} \quad c > 0; \tag{1.17}$$

$$a_n \sim b_n \tag{1.18}$$

if (1.17) holds with $c = 1$. Note that if $a_n \sim b_n$ and $\sum_n a_n$ diverges, then $\sum_{k=1}^{n} a_k \sim \sum_{k=1}^{n} b_k$.

2. Hyperbolic Multichannel Models

We start by introducing the following objects:

(1) H, a separable Hilbert space with an orthonormal basis $\{h_k,\ k \geq 1\}$;
(2) X, a separable Hilbert space such that H is densely and continuously embedded into X and

$$\sum_{k\geq 1} \|h_k\|_X^2 < \infty \tag{2.1}$$

(in other words, the embedding operator from H to X is Hilbert-Schmidt);
(3) $\Theta_1,\ \Theta_2$, two compact sets in $\mathbb{R}$;

(4) $\theta_1,\ \theta_2$, two real numbers, $\theta_1 \in \Theta_1$, $\theta_2 \in \Theta_2$;

(5) $(\Omega, \mathcal{F},\ (\mathcal{F}_t)_{t\geq 1}, \mathbb{P})$, a stochastic basis with the usual assumptions and a countable collection of independent standard Brownian motions $\{w_k = w_k(t),\ k \geq 1\}$.

In this setting, a cylindrical Brownian motion $W = W(t)$ on H is a continuous X-valued Gaussian process with representation

$$W(t) = \sum_{k\geq 1} h_k w_k(t). \tag{2.2}$$

The process W indeed has values in X rather than H because

$$\mathbb{E}\|W(t)\|_X^2 = t\sum_{k\geq 1} \|h_k\|_X^2 < \infty.$$

For fixed non-random $T > 0$, consider the multichannel model

$$\begin{aligned} &\ddot{u}_k(t) - \mu_k(\theta_2)\dot{u}(t) + \lambda_k(\theta_1)u_k(t) = \dot{w}_k(t),\ 0 < t \leq T,\ k \geq 1 \\ &u_k(0) = \dot{u}_k(0) = 0, \end{aligned} \tag{2.3}$$

with

$$\lambda_k(\theta_1) = \kappa_k + \theta_1\tau_k, \quad \mu_k(\theta_2) = \rho_k + \theta_2\nu_k. \tag{2.4}$$

We also consider the equivalent operator formulation

$$\ddot{u} + \mathcal{A}(\theta_1)u = \mathcal{B}(\theta_2)\dot{u} + \dot{W}, \qquad 0 < t \leq T, \tag{2.5}$$

where $\mathcal{A}(\theta_1) = \mathcal{A}_0 + \theta_1\mathcal{A}_1$, $\mathcal{B}(\theta_2) = \mathcal{B}_0 + \theta_2\mathcal{B}_1$,

$$\mathcal{A}_0 h_k = \kappa_k h_k, \quad \mathcal{B}_0 = \rho_k h_k, \quad \mathcal{A}_1 h_k = \tau_k h_k, \quad \mathcal{B}_1 = \nu_k h_k. \tag{2.6}$$

We will refer to $\mathcal{A}(\theta_1)$ and $\mathcal{B}(\theta_2)$ as the `evolution` and `dissipation` operators, respectively. Hyperbolicity of (2.3) means that the evolution operator is bounded from below and dominates, in some sense, the dissipation operator. More precisely, we have

Definition 2.1. *Multichannel model* (2.3) *or* (2.5) *is called* `hyperbolic` *on the time interval* $[0, T]$ *if*

(1) *there exist positive numbers* C^*, c_1, c_2 *such that* $\{\kappa_k + \theta\tau_k + C^*,\ k \geq 1\}$ *is a positive, non-decreasing, and unbounded sequence for all* $\theta \in \Theta_1$ *and*

$$c_1 \leq \frac{\kappa_k + \theta\tau_k + C^*}{\kappa_k + \theta'\tau_k + C^*} \leq c_2 \tag{2.7}$$

for all $\theta, \theta' \in \Theta_1$*;*

(2) *there exist positive numbers* C, J *such that, for all* $k \geq J$ *and all* $\theta_1 \in \Theta_1,\ \theta_2 \in \Theta_2$,

$$T(\rho_k + \theta_2\nu_k) \leq \ln(\kappa_k + \theta_1\tau_k) + C. \tag{2.8}$$

Condition (2.8) means that there is no restriction on the strength of dissipation, but amplification must be weak. For example, let $\boldsymbol{\Delta}$ be the Laplace operator in a smooth bounded domain $G \subset \mathbb{R}^d$ with zero boundary conditions, and $H = L_2(G)$. Then each of the following models is hyperbolic on $[0, T]$ for all $T > 0$:

$$\begin{aligned} &u_{tt} = \boldsymbol{\Delta}u + u_t + \dot{W}, \quad u_{tt} = \boldsymbol{\Delta}u - u_t + \dot{W}, \\ &u_{tt} = \boldsymbol{\Delta}(u + u_t) + \dot{W}, \quad u_{tt} = \boldsymbol{\Delta}u - \boldsymbol{\Delta}^2 u_t + \dot{W}, \end{aligned} \tag{2.9}$$

while

$$u_{tt} = \mathbf{\Delta}(u - u_t) + \dot{W} \quad \text{and} \quad u_{tt} = \mathbf{\Delta} u + \mathbf{\Delta}^2 u_t + \dot{W}$$

are not hyperbolic on any $[0,T]$. To construct a model that is hyperbolic on every time interval $[0,T]$ and has unbounded amplification, take $\theta_1 = \theta_2 = 1$, $\kappa_k = \rho_k = 0$, $\tau_k = e^k$, $\nu_k = \ln\ln k$.

The following result shows that, in a hyperbolic model, the evolution operator is uniformly bounded from below.

Proposition 2.2. *If* (2.3) *is hyperbolic, then*

$$\lim_{k\to\infty} \big(\kappa_k + \theta\tau_k\big) = +\infty \tag{2.10}$$

uniformly in $\theta \in \Theta_1$, *and there exists an index* $J \geq 1$ *and a number* c_0 *such that, for all* $k \geq J$ *and* $\theta \in \Theta_1$,

$$\kappa_k + \theta\tau_k > 1, \tag{2.11}$$

$$\frac{|\tau_k|}{\kappa_k + \theta\tau_k} \leq c_0. \tag{2.12}$$

Proof. Recall that $\lambda_k(\theta) = \kappa_k + \theta\tau_k$. Since $\{\lambda_k(\theta) + C^*,\ k \geq 1\}$ is a positive, non-decreasing, and unbounded sequence for all $\theta \in \Theta_1$ and (2.7) holds, we have (2.10), and then (2.11) follows.

To prove (2.12), we argue by contradiction. Assume that the sequence $\{|\tau_k|\,\lambda_k^{-1}(\theta), k \geq 1\}$ is not uniformly bounded. Then there exists a sequence $\{|\tau_{k_j}|\,\lambda_{k_j}^{-1}(\theta_j),\ j \geq 1\}$ such that

$$\lim_{j\to\infty} \frac{|\tau_{k_j}|}{\theta_j\tau_{k_j} + \kappa_{k_j}} = +\infty. \tag{2.13}$$

With no loss of generality, assume that $\tau_{k_j} > 0$, and, since Θ_1 is compact, we also assume that $\lim_{j\to\infty} \theta_j = \theta^\circ \in \Theta_1$ (if not, extract a further sub-sequence).

Then (2.13) implies

$$\lim_{j\to\infty} \frac{\kappa_{k_j}}{\tau_{k_j}} = -\theta^\circ. \tag{2.14}$$

Note that $\lim_{j\to\infty} |\tau_{k_j}| = +\infty$, because $\lim_{j\to\infty}(\theta^\circ\tau_{k_j} + \kappa_{k_j}) = +\infty$. Consequently,

$$\lim_{j\to\infty} \frac{\lambda_{k_j}(\theta) + C^*}{\lambda_{k_j}(\theta^\circ) + C^*} = \frac{\theta - \theta^\circ}{\theta^\circ + \lim_{j\to\infty}(\kappa_{k_j}/\tau_{k_j})} = \infty, \quad \theta \neq \theta^\circ.$$

As a result, if (2.12) fails, then so does (2.7) for $\theta \neq \theta^\circ$, $\theta' = \theta^\circ$. □

Theorem 2.3. *The process* $u(t) = \sum_{k\geq 1} u_k(t)h_k$ *is an element of* X *for every* t *if and only if* (2.3) *is hyperbolic in the sense of Definition 2.1. If, in addition, there exists a real number* C_0 *such that* $\theta_2\nu_k + \rho_k \leq C_0$ *for all* $k \geq 1$ *and* $\theta_2 \in \Theta_2$, *then* $v(t) = \sum_{k\geq 1} \dot{u}_k(t)h_k$ *is also an* X*-valued process.*

Proof. For every $k \geq 1$, equation (2.3) has a unique solution, and direct computations show that

$$u_k(t) = \int_0^t \mathfrak{f}_k(t-s)\,dw_k(s), \tag{2.15}$$

where the `fundamental solution` $\mathfrak{f}_k$ satisfies

$$\ddot{\mathfrak{f}}_k(t) - \mu_k \dot{\mathfrak{f}}_k(t) + \lambda_k \mathfrak{f}_k(t) = 0, \quad \mathfrak{f}_k(0) = 0, \ \dot{\mathfrak{f}}_k(0) = 1; \tag{2.16}$$

see Appendix for details. Thus, $\mathbb{E}u_k(t) = 0$ and, since the processes u_k are independent for different k, the series $\sum_{k\geq 1} u_k(t)h_k$ defines an X-valued process if and only if

$$\sup_{k\geq 1} \sup_{t\in[0,T]} \mathbb{E}|u_k(t)|^2 < \infty. \tag{2.17}$$

By direct computation using (2.15) and the Itô isometry,

$$\mathbb{E}|u_k(t)|^2 = \int_0^t \mathfrak{f}_k^2(t-s)ds = \int_0^t \mathfrak{f}_k^2(s)ds. \tag{2.18}$$

The proof of the theorem is thus reduced to the study of the fundamental solution $\mathfrak{f}_k$ for sufficiently large k. More precisely, we will show that

$$\sup_{t\in[0,T]} \sup_{k} \mathfrak{f}_k^2(t) < \infty, \tag{2.19}$$

which, by (2.18), implies (2.17).

The solution of equation (2.16) is determined by the roots $r_{\pm}$ of the characteristic equation

$$r^2 - \mu_k r + \lambda_k = 0: \quad r_{\pm} = \frac{\mu_k \pm \sqrt{\mu_k^2 - 4\lambda_k}}{2}. \tag{2.20}$$

By Lemma 2.2, $\lim_{k\to\infty} \lambda_k = +\infty$, and, in particular, $\lambda_k > 0$ for all sufficiently large k. Also, condition (2.8) means that if $\mu_k > 0$, then $\mu_k \leq (\ln \lambda_k + C)/T$, and therefore $\mu_k < 2\sqrt{\lambda_k}$ for all sufficiently large k. Accordingly, we assume that $\lambda_k > 0$ and consider two cases: $|\mu_k| < 2\sqrt{\lambda_k}$ and $\mu_k \leq -2\sqrt{\lambda_k}$.

If $|\mu_k| < 2\sqrt{\lambda_k}$, then equation (2.20) has complex conjugate roots, and, with $\ell_k = \sqrt{\lambda_k - (\mu_k^2/4)}$,

$$\mathfrak{f}_k^2(t) = t^2 e^{\mu_k t} \left(\frac{\sin(\ell_k t)}{\ell_k t}\right)^2. \tag{2.21}$$

If $\mu_k \leq 0$, then $\mathfrak{f}_k^2 \leq T^2$ for all $t \in [0,T]$ and (2.19) follows. If $\mu_k > 0$, then, for sufficiently large k, condition (2.8) ensures that $e^{\mu_k t} \leq \lambda_k e^C$ and $\lambda_k/\ell_k^2 < 2$. Then $\mathfrak{f}_k^2 \leq 2T^2 e^C$ and (2.19) follows.

If $\mu_k \leq -2\sqrt{\lambda_k}$, then (2.20) has real roots (a double root if $\mu_k = -2\sqrt{\lambda_k}$), and, using the notations $\ell_k = \sqrt{\mu_k^2 - 4\lambda_k}$, $a = \mu_k + \ell_k$,

$$\mathfrak{f}_k^2(t) = t^2 e^{at} \left(\frac{1 - e^{-\ell_k t}}{\ell_k t}\right)^2; \tag{2.22}$$

the case of the double root corresponds to the limit $\ell_k \to 0$. By assumption, $a \leq 0$, so that $\mathfrak{f}_k^2(t) \leq T^2$ and (2.19) follows.

Similarly, $v(t) = \sum_k v_k(t)h_k$, $v_k(t) = \int_0^t \dot{\mathfrak{f}}_k(t-s)dw_k(s)$, and $\mathbb{E}v_k^2(t) = \int_0^t |\dot{\mathfrak{f}}_k(s)|^2 ds$. By direct computation, if $\mu_k \leq C_0$, then

$$\sup_{t\in[0,T]} \sup_{k} |\dot{\mathfrak{f}}_k(t)|^2 < \infty,$$

and therefore $v(t) \in X$.

This completes the proof of Theorem 2.3. □

3. Estimation of Parameters

Assume that the measurements of $u_k(t)$ from (2.3) and of $v_k(t) = \dot{u}_k$ are available for all $t \in [0, T]$ and $k = 1, \ldots, N$. The objective is to estimate the parameters θ_1, θ_2.

Since

$$dv_k(t) = \big(-\lambda_k(\theta_1)u_k + \mu_k(\theta_2)v_k\big)dt + dw_k(t) \tag{3.1}$$

(see (2.3)), and $u_k(t) = \int_0^t v_k(s)ds$, the vector process $\mathbf{v} = (v_1, \ldots, v_N)$ is a diffusion-type process in the sense of Liptser and Shiryaev [17, Definition 4.2.7]. Therefore, by Theorem 7.6 in [17] (see also Section 7.2.7 of the same reference), the measure $\mathbf{P}^{\mathbf{v}}$ generated by the process $\mathbf{v}$ in the space of $\mathbb{R}^N$-valued continuous functions on $[0, T]$ is absolutely continuous with respect to the measure $\mathbf{P}^{\mathbf{w}}$, generated in the same space by the N-dimensional standard Brownian motion $\mathbf{w} = (w_1, \ldots, w_N)$. Moreover, the density $Z = d\mathbf{P}^{\mathbf{v}}/d\mathbf{P}^{\mathbf{w}}$ satisfies

$$\begin{aligned} Z(\mathbf{v}) = \exp\Bigg(\sum_{k=1}^{N} \Big(\int_0^T & \big(-\lambda_k(\theta_1)u_k(t) + \mu_k(\theta_2)v_k(t)\big) dv_k(t) \\ & - \frac{1}{2}\int_0^T \big(-\lambda_k(\theta_1)u_k(t) + \mu_k(\theta_2)v_k(t)\big)^2 dt\Big)\Bigg). \end{aligned}$$

Define

$$\mathfrak{z} = \ln Z(\mathbf{v}).$$

Note that $\mathfrak{z}$ is a function of θ_1, θ_2, and the maximum likelihood estimator of the parameters θ_1, θ_2 is computed by solving the system of equations

$$\frac{\partial \mathfrak{z}}{\partial \theta_1} = 0, \quad \frac{\partial \mathfrak{z}}{\partial \theta_2} = 0, \tag{3.2}$$

with unknowns θ_1, θ_2. This system can be written as

$$\begin{aligned} F_{1,N} + L_{1,N} + K_{1,N}\theta_1 + K_{12,N}\theta_2 &= A_{1,N} \\ F_{2,N} + L_{2,N} + K_{12,N}\theta_1 + K_{2,N}\theta_2 &= A_{2,N}, \end{aligned} \tag{3.3}$$

where

$$
\begin{aligned}
&A_{1,N} = -\sum_{k=1}^{N}\int_0^T \tau_k u_k(t)dv_k(t), \quad A_{2,N} = \sum_{k=1}^{N}\int_0^T \nu_k v_k(t)dv_k(t),\\
&F_{1,N} = -\sum_{k=1}^{N}\int_0^T \kappa_k\tau_k u_k^2(t)dt, \quad F_{2,N} = \sum_{k=1}^{N}\int_0^T \rho_k\nu_k v_k^2(t)dt,\\
&K_{1,N} = \sum_{k=1}^{N}\int_0^T \tau_k^2 u_k^2(t)dt,\ K_{2,N} = \sum_{k=1}^{N}\int_0^T \nu_k^2 v_k^2(t)dt,\\
&K_{12,N} = -\sum_{k=1}^{N}\int_0^T \nu_k\tau_k u_k(t)v_k(t)dt,\\
&L_{1,N} = -\sum_{k=1}^{N}\int_0^T \rho_k\tau_k u_k(t)v_k(t)dt, \quad L_{2,N} = -\sum_{k=1}^{N}\int_0^T \kappa_k\nu_k u_k(t)v_k(t)dt.
\end{aligned}
\tag{3.4}
$$

All the numbers A, F, L and K are computable from the observations of $u_k(t)$ and $v_k(t)$, $k = 1, \ldots, N$, $t \in [0, T]$.

Note that

$$
K_{12,N} = -\frac{1}{2}\sum_{k=1}^{N}\tau_k\nu_k u_k^2(T), \quad L_{1,N} = -\frac{1}{2}\sum_{k=1}^{N}\rho_k\tau_k u_k^2(T),
$$

$$
L_{2,N} = -\frac{1}{2}\sum_{k=1}^{N}\kappa_k\nu_k u_k^2(T),
$$

because, by assumption, $u_k(0) = 0$ and thus

$$
\int_0^T u_k v_k(t)dt = \int_0^T u_k(t)du_k(t) = \frac{1}{2}u_k^2(T).
$$

By the Cauchy-Schwartz inequality, $K_{1,N}K_{2,N} - K_{12,N}^2 > 0$ with probability one, because the process u_k is not a scalar multiple of v_k. Therefore (3.3) has a unique solution

$$
\begin{aligned}
\hat{\theta}_{1,N} &= \frac{K_{2,N}\big(A_{1,N} - F_{1,N} - L_{1,N}\big) - K_{12,N}\big(A_{2,N} - L_{2,N} - F_{2,N}\big)}{K_{1,N}K_{2,N} - K_{12,N}^2},\\
\hat{\theta}_{2,N} &= \frac{K_{1,N}\big(A_{2,N} - F_{2,N} - L_{2,N}\big) - K_{12,N}\big(A_{1,N} - F_{1,N} - L_{1,N}\big)}{K_{1,N}K_{2,N} - K_{12,N}^2}.
\end{aligned}
\tag{3.5}
$$

With notations (3.4) in mind, formulas (3.5) provide explicit expressions for the maximum likelihood estimators of θ_1 and θ_2. To study asymptotic properties of these estimators, we need expressions for $\hat{\theta}_{i,N} - \theta_i$, $i = 1, 2$:

$$
\begin{aligned}
\hat{\theta}_{1,N} - \theta_1 &= \frac{1}{1 - D_N}\left(\frac{\iota_{1,N}}{K_{1,N}} - \frac{\iota_{2,N}K_{12,N}}{K_{1,N}K_{2,N}}\right),\\
\hat{\theta}_{2,N} - \theta_2 &= \frac{1}{1 - D_N}\left(\frac{\iota_{2,N}}{K_{2,N}} - \frac{\iota_{1,N}K_{12,N}}{K_{1,N}K_{2,N}}\right),
\end{aligned}
\tag{3.6}
$$

where

$$
\begin{aligned}
\iota_{1,N} &= -\sum_{k=1}^{N} \int_0^T \tau_k u_k(t)dw_k(t), \quad \iota_{2,N} = \sum_{k=1}^{N} \int_0^T \nu_k v_k(t)dw_k(t),\\
D_N &= \frac{K_{12,N}^2}{K_{1,N}K_{2,N}}.
\end{aligned} \tag{3.7}
$$

It follows that, as $N \to \infty$, asymptotic behavior of the estimators is determined by $\iota_{i,N}/K_{i,N}$, $i = 1, 2$, and $K_{12,N}/(K_{1,N}K_{2,N})$. Note that each of $\iota_{i,N}, K_{i,N}, K_{12,N}$ is a sum of independent random variables. Moreover,

$$
\mathbb{E}\iota_{i,N}^2 = \mathbb{E}K_{i,N}, \ i = 1, 2. \tag{3.8}
$$

If $\mathfrak{f}_k$ is the function satisfying

$$
\ddot{\mathfrak{f}}_k(t) - \mu_k(\theta_2)\dot{\mathfrak{f}}_k(t) + \lambda_k(\theta_1)\mathfrak{f}_k(t) = 0, \quad \mathfrak{f}_k(0) = 0, \ \dot{\mathfrak{f}}_k(0) = 1, \tag{3.9}
$$

then, by direct computation, $u_k(t) = \int_0^t \mathfrak{f}_k(t-s)dw_k(s)$ (see Appendix for more details), so that

$$
\mathbb{E}u_k^2(t) = \int_0^t |\mathfrak{f}_k(s)|^2 ds, \ \mathbb{E}v_k^2(t) = \int_0^t |\dot{\mathfrak{f}}_k(s)|^2 ds,
$$

and

$$
\begin{aligned}
\Psi_{1,N} &:= \mathbb{E}K_{1,N} = \sum_{k=1}^{N} \tau_k^2 \int_0^T \int_0^t |\mathfrak{f}_k(s)|^2 dsdt,\\
\Psi_{2,N} &:= \mathbb{E}K_{2,N} = \sum_{k=1}^{N} \nu_k^2 \int_0^T \int_0^t |\dot{\mathfrak{f}}_k(s)|^2 dsdt,\\
\Psi_{12,N} &:= \mathbb{E}K_{12,N} = -\frac{1}{2} \sum_{k=1}^{N} \tau_k \nu_k \int_0^T |\mathfrak{f}_k(s)|^2 ds.
\end{aligned} \tag{3.10}
$$

The following is a necessary conditions for the consistency of the estimators.

Proposition 3.1. *If* $\lim_{N\to\infty} \hat{\theta}_{1,N} = \theta_i$ *in probability, then*

$$
\lim_{N\to\infty} \Psi_{1,N} = +\infty.
$$

Similarly, if $\lim_{N\to\infty} \hat{\theta}_{2,N} = \theta_i$ *in probability, then*

$$
\lim_{N\to\infty} \Psi_{2,N} = +\infty.
$$

Proof. Each of the sequences $\{\Psi_{i,N}, \ N \geq 1\}$ is monotonically increasing and thus has a limit, finite or infinite. If $\lim_{N\to\infty} \Psi_{i,N} < \infty$, then $\lim_{N\to\infty} \iota_{i,N}/K_{i,N}$ exists with probability one and is a non-degenerate random variable; by (3.6), $\hat{\theta}_{i,N}$ cannot converge to θ_i. □

Under the assumptions of Theorem 2.3, we derived a bound $|\mathfrak{f}_k(t)|^2 \leq const. \cdot T^2$, which was enough to establish existence and uniqueness of solution of (2.5). To study estimators $\hat{\theta}_{i,N}$, and, in particular, convergence or divergence of the sequences

$\{\Psi_{i,N},\ N \geq 1\}$, we need more delicate bounds on both $|\mathfrak{f}_k(t)|^2$ and $|\dot{\mathfrak{f}}_k(t)|^2$. The computations, while relatively straightforward, are rather long and lead to the following relations (see (1.16) for the definition of $\sim$):

$$\mathbb{E}u_k^2(T) \sim \frac{e^{\mu_k(\theta_2)T}-1}{2\mu_k(\theta_2)\lambda_k(\theta_1)}, \quad \operatorname{Var} u_k^2(T) \sim 3\left(\frac{e^{\mu_k(\theta_2)T}-1}{2\mu_k(\theta_2)\lambda_k(\theta_1)}\right)^2; \tag{3.11}$$

$$\mathbb{E}\int_0^T u_k^2(t)dt \sim \frac{T^2 M\big(T\mu_k(\theta_2)\big)}{\lambda_k(\theta_1)}, \quad \operatorname{Var}\int_0^T u_k^2(t)dt \sim \frac{T^4 V\big(T\mu_k(\theta_2)\big)}{\lambda_k^2(\theta_1)}, \tag{3.12}$$

$$\mathbb{E}\int_0^T v_k^2(t)dt \sim T^2 M\big(T\mu_k(\theta_2)\big), \quad \operatorname{Var}\int_0^T v_k^2(t)dt \sim T^4 V\big(T\mu_k(\theta_2)\big), \tag{3.13}$$

where

$$M(x) = \begin{cases} \dfrac{e^x - x - 1}{2x^2}, & \text{if } x \neq 0, \\ \dfrac{1}{4}, & \text{if } x = 0; \end{cases} \tag{3.14}$$

$$V(x) = \begin{cases} \dfrac{e^{2x} + 4e^x - 4xe^x - 2x - 5}{4x^4}, & \text{if } x \neq 0, \\ \dfrac{1}{24}, & \text{if } x = 0. \end{cases} \tag{3.15}$$

Note that the functions M and V are continuous and positive on $\mathbb{R}$, and

$$M(x) \sim \begin{cases} (2|x|)^{-1}, & x \to -\infty, \\ 2(2x)^{-2}\, e^x, & x \to +\infty; \end{cases} \quad V(x) \sim \begin{cases} 4(2|x|)^{-3}, & x \to -\infty, \\ 4(2x)^{-4}\, e^{2x}, & x \to +\infty. \end{cases} \tag{3.16}$$

The computations leading to (3.11)–(3.13) rely on the fact that u_k and v_k are Gaussian processes, so that, for example,

$$\operatorname{Var}\int_0^T u_k^2(t)dt = 4\int_0^T\int_0^t \Big(\mathbb{E}\big(u(t)u(s)\big)\Big)^2 ds dt.$$

It follows from (3.12) and (3.13) that if $\lim_{N\to\infty}\Psi_{i,N} = +\infty$, then

$$\Psi_{1,N} \sim T^2 \sum_{k=1}^N \frac{\tau_k^2\, M\big(T\mu_k(\theta_2)\big)}{\lambda_k(\theta_1)}, \ \ \Psi_{2,N} \sim T^2 \sum_{k=1}^N \nu_k^2\, M\big(T\mu_k(\theta_2)\big). \tag{3.17}$$

Relations (3.17) show that conditions for consistency and asymptotic normality of the estimators require additional assumptions about κ_k, τ_k, ρ_k, and ν_k. To understand the nature of the assumptions, let $\mathcal{D}$ be an operator defined on smooth functions by

$$\mathcal{D}f(x) = -\sum_{i,j=1}^d \frac{\partial}{\partial x_i}\left(a_{ij}(x)\frac{\partial f(x)}{\partial x_j}\right),$$

in a smooth bounded domain $G \subset \mathbb{R}^d$, with zero Dirichlet boundary conditions. Assume that the functions a_{ij} are all infinitely differentiable in G and are bounded with all the derivatives, and the matrix $(a_{ij}(x),\ i,j = 1,\dots,d)$ is symmetric and uniformly positive-definite in G. Then the eigenvalues d_k of $\mathcal{D}$ can be enumerated so that

$$d_k \asymp k^{2/d} \tag{3.18}$$

in the sense of notation (1.16). More generally, for a positive-definite elliptic self-adjoint differential or pseudo-differential operator $\mathcal{D}$ of order m on a smooth bounded domain in $\mathbb{R}^d$ with suitable boundary conditions or on a smooth compact d-dimensional manifold, the asymptotic of the eigenvalues $d_k,\ k \geq 1$, is

$$d_k \asymp k^{m/d}; \tag{3.19}$$

note that m can be an arbitrary positive number. This result is well-known; see, for example, Safarov and Vassiliev [23, Section 1.2]. An example of $\mathcal{D}$ is $(1-\boldsymbol{\Delta})^{m/2}$, $m > 0$, where $\boldsymbol{\Delta}$ is the Laplace operator; for this operator relation (3.19) holds even when $m \leq 0$.

In our setting, when the operators are *defined* by their eigenvalues and eigenfunctions, more exotic eigenvalues are possible, for example, $\tau_k = e^k$ or $\nu_k = (-1)^k/k$. On the other hand, it is clear that the analysis of the estimators should be easier when all the sequences are of the type (3.19). Accordingly, we make the following

Definition 3.2. *Model* (2.3) *is called* `algebraically hyperbolic` *if it is hyperbolic, and the sequences* $\lambda_k(\theta) = \kappa_k + \theta\tau_k$, $\mu_k(\theta) = \rho_k + \theta\nu_k$ *have the following properties:*

(1) *There exist real numbers* α, α_1 *such that, for all* $\theta \in \Theta_1$,

$$\lambda_k(\theta) \asymp k^{\alpha}, \quad |\tau_k| \asymp k^{\alpha_1}; \tag{3.20}$$

(2) *Either* $|\mu_k(\theta)| \leq C$ *for all* $\theta \in \Theta_2$ *or there exist numbers* $\beta > 0,\ \beta_1 \in \mathbb{R}$ *such that, for all* $\theta \in \Theta_2$,

$$-\mu_k(\theta) \asymp k^{\beta},\ |\nu_k| \asymp k^{\beta_1}. \tag{3.21}$$

To emphasize the importance of the numbers α and β, we will sometimes say that the model is (α, β)-algebraically hyperbolic; $\beta = 0$ includes the case of uniformly bounded $\mu_k(\theta)$.

It follows that

- under hyperbolicity assumption, $\alpha > 0$ and no unbounded amplification is possible;
- each of the models in (2.9) is algebraically hyperbolic.

4. Analysis of Estimators: Algebraic Case

Theorem 4.1. *Assume that model* (2.3) *is* (α, β)*-algebraically hyperbolic in the sense of Definition 3.2.*

(1) *If*

$$\alpha_1 \geq \frac{\alpha + \beta - 1}{2}, \tag{4.1}$$

then the estimator $\hat{\theta}_{1,N}$ *is strongly consistent and asymptotically normal with rate* $\sqrt{\Psi_{1,N}}$ *as* $N \to \infty$:

$$\lim_{N\to\infty} \hat{\theta}_{1,N} = \theta_1 \ \text{ with probability one;} \tag{4.2}$$

$$\lim_{N\to\infty} \sqrt{\Psi_{1,N}}\Big(\hat{\theta}_{1,N} - \theta_1\Big) = \xi_1 \text{ in distribution,} \tag{4.3}$$

where ξ_1 *is a standard Gaussian random variable.*

(2) *If*

$$\beta_1 \geq \frac{\beta - 1}{2}, \tag{4.4}$$

then the estimator $\hat{\theta}_{2,N}$ is strongly consistent and asymptotically normal with rate $\sqrt{\Psi_{2,N}}$ as $N \to \infty$:

$$\lim_{N\to\infty} \hat{\theta}_{2,N} = \theta_2 \ \text{ with probability one;} \tag{4.5}$$

$$\lim_{N\to\infty} \sqrt{\Psi_{2,N}}\Big(\hat{\theta}_{2,N} - \theta_2\Big) = \xi_2 \text{ in distribution,} \tag{4.6}$$

where ξ_2 is a standard Gaussian random variable.

(3) *If both* (4.1) *and* (4.4) *hold, then the random variables ξ_1, ξ_2 are independent.*

Remark 4.2. *(a) In terms of the orders of the operators in the equivalent operator formulation* (2.5) (*see* (3.19)), *condition* (4.1) *becomes*

$$\text{order}(\mathcal{A}_1) \geq \frac{\text{order}(\mathcal{A}_0 + \theta_1 \mathcal{A}_1) + \text{order}(\mathcal{B}_0 + \theta_2 \mathcal{B}_1) - d}{2}, \tag{4.7}$$

and condition (4.4) *becomes*

$$\text{order}(\mathcal{B}_1) \geq \frac{\text{order}(\mathcal{B}_0 + \theta_2 \mathcal{B}_1) - d}{2}. \tag{4.8}$$

(b) The condition for consistency of $\hat{\theta}_{2,N}$ does not depend on the evolution operator and is similar to the consistency condition in the parabolic case [7, Theorem 2.1].

The operator formulation (2.5) reveals intuition behind conditions (4.7) and (4.8). The information about the numbers θ_1, θ_2 is carried by the terms $\mathcal{A}_1 u$ and $\mathcal{B}_1 \dot{u}$, respectively, and these terms must be irregular enough to be distinguishable in the noise $\dot{W}$ *during a finite observation window* $[0, T]$. The higher the orders of the operators, the more irregular the terms, the easier the estimation.

Proof of Theorem 4.1. Note that if $\beta > 0$, then $\lim_{k\to\infty} \mu_k(\theta) = -\infty$, and therefore, by (3.16),

$$M_k\big(T\mu_k(\theta)\big) \sim \frac{1}{2T|\mu_k(\theta)|} \asymp k^{-\beta}, \ V_k\big(T\mu_k(\theta)\big) \sim \frac{1}{2|T\mu_k(\theta)|^3} \asymp k^{-3\beta}. \tag{4.9}$$

Let

$$\gamma_1 = 2\alpha_1 - \alpha - \beta, \quad \gamma_2 = 2\beta_1 - \beta, \ \gamma_{12} = \alpha_1 - \alpha + \beta_1 - \beta.$$

We have (see (1.15) for the definition of $\bowtie$)

$$\tau_k^2 \mathbb{E} \int_0^T u_k^2(t)dt \bowtie k^{\gamma_1}, \ \tau_k^4 \operatorname{Var} \int_0^T u_k^2(t)dt \bowtie k^{2\gamma_1 - \beta}, \tag{4.10}$$

$$\nu_k^2 \mathbb{E} \int_0^T v_k^2(t)dt \bowtie k^{\gamma_2}, \ \nu_k^4 \operatorname{Var} \int_0^T v_k^2(t)dt \bowtie k^{2\gamma_2 - \beta}, \tag{4.11}$$

$$|\nu_k \tau_k| \, \mathbb{E} u_k^2(T) \bowtie k^{\gamma_{12}}, \ \nu_k^2 \tau_k^2 \operatorname{Var} u_k^2(T) \bowtie k^{2\gamma_{12}}, \tag{4.12}$$

and therefore

$$\Psi_{1,N} \bowtie \begin{cases} \text{const.}, & \text{if } \gamma_1 < -1, \\ \ln N, & \text{if } \gamma_1 = -1, \\ N^{\gamma_1+1}, & \text{if } \gamma_1 > -1, \end{cases} \qquad \Psi_{2,N} \bowtie \begin{cases} \text{const.}, & \text{if } \gamma_2 < -1, \\ \ln N, & \text{if } \gamma_2 = -1, \\ N^{\gamma_2+1}, & \text{if } \gamma_2 > -1, \end{cases} \tag{4.13}$$

$$|\Psi_{12,N}| \bowtie \begin{cases} \text{const.}, & \text{if } \gamma_{12} < -1, \\ \ln N, & \text{if } \gamma_{12} = -1, \\ N^{\gamma_{12}+1}, & \text{if } \gamma_{12} > -1. \end{cases} \tag{4.14}$$

Next, we show that condition (4.1) implies

$$\lim_{N\to\infty} \frac{K_{1,N}}{\Psi_{1,N}} = 1 \text{ with probability one,} \tag{4.15}$$

condition (4.4) implies

$$\lim_{N\to\infty} \frac{K_{2,N}}{\Psi_{2,N}} = 1 \text{ with probability one,} \tag{4.16}$$

and either (4.1) or (4.4),

$$\lim_{N\to\infty} D_N = 0 \text{ with probability one.} \tag{4.17}$$

Indeed, (4.15) follows from (3.12) and (B.8), because (4.1) implies

$$\sum_n \frac{n^{2\gamma_1-\beta}}{\Psi^2_{n,1}} < \infty.$$

Similarly, (4.16) follows from (3.13) and (B.8), because (4.4) implies

$$\sum_n \frac{n^{2\gamma_2}}{\Psi^2_{n,2}} < \infty.$$

For (4.17), we first observe that $\lim_{N\to\infty} K_{12,N}/\Psi_{12,N}$ exists with probability one. If $\gamma_{12} < -1$, then the limit is a $\mathbb{P}$-a.s finite random variable. If $\gamma_{12} \geq -1$, then (3.11) and (B.8) imply that the limit is 1. Then direct analysis shows that

$$\lim_{N\to\infty} \frac{\Psi^2_{12,N}}{\Psi_{1,N}\Psi_{2,N}} = 0$$

if at least one of $\Psi_{1,N}$, $\Psi_{2,N}$ is unbounded.

Next, we show that (4.1) implies

$$\lim_{N\to\infty} \frac{\iota_{1,N}}{\Psi_{1,N}} = 0 \text{ with probability one,} \tag{4.18}$$

and

$$\lim_{N\to\infty} \frac{\iota_{1,N}}{\sqrt{\Psi_{1,N}}} = \xi_1 \text{ in distribution,} \tag{4.19}$$

whereas (4.4) implies

$$\lim_{N\to\infty} \frac{\iota_{2,N}}{\Psi_{2,N}} = 0 \text{ with probability one,} \tag{4.20}$$

and

$$\lim_{N\to\infty} \frac{\iota_{2,N}}{\sqrt{\Psi_{2,N}}} = \xi_2 \text{ in distribution.} \tag{4.21}$$

Indeed, (4.18) follows from (3.12) and (B.3), because (4.1) implies that $\sum_k k^{\gamma_1} = +\infty$. Similarly, (4.15) follows from (3.13) and (B.3).

Both (4.19) and (4.21) follow from Corollary B.4. Together with (4.17), the same Corollary also implies independence of ξ_1 and ξ_2 if both (4.1) and (4.4) hold.

To complete the proof of the theorem it remains to verify that

$$\lim_{N\to\infty} \frac{\iota_{i,N} K_{12,N}}{K_{1,N}K_{2,N}} = 0, \ i = 1,2, \text{ with probability one,} \tag{4.22}$$

and

$$\lim_{N\to\infty} \frac{\sqrt{\Psi_{i,N}}\, \iota_{i,N} K_{12,N}}{K_{1,N}K_{2,N}} = 0, \ i = 1,2, \text{ in probability.} \tag{4.23}$$

Direct computations show that (B.3) implies (4.22), and (4.17), (4.19), (4.21) imply (4.23).

This completes the proof of Theorem 4.1. □

Remark 4.3. *From Proposition 3.1, we see that condition* (4.1) *is both necessary and sufficient for consistency and asymptotic normality of* $\hat{\theta}_{1,N}$. *Similarly, condition* (4.4) *is necessary and sufficient for consistency and asymptotic normality of* $\hat{\theta}_{2,N}$.

Since in the algebraic case the sum $\sum_{k=1}^N k^{\gamma}$ appears frequently, we introduce a special notation to describe the asymptotic of this sum as $N \to \infty$ for $\gamma \geq -1$:

$$\Upsilon_N(\gamma) = \begin{cases} N^{\gamma+1}, & \text{if } \gamma > -1, \\ \ln N, & \text{if } \gamma = -1. \end{cases} \tag{4.24}$$

With this notation, $\sum_{k=1}^N k^{\gamma} \asymp \Upsilon_N(\gamma)$, $\gamma \geq -1$.

Let us consider several examples, in which $\boldsymbol{\Delta}$ is the Laplace operator in a smooth bounded domain G in $\mathbb{R}^d$ with zero boundary conditions; $H = L_2(G)$. We start with these three equations:

$$u_{tt} = \theta_1 \boldsymbol{\Delta} u + \theta_2 u_t + \dot{W}, \ \theta_1 > 0, \ \theta_2 \in \mathbb{R}; \tag{4.25}$$

$$u_{tt} = \boldsymbol{\Delta}(\theta_1 u + \theta_2 u_t) + \dot{W}, \ \theta_1 > 0, \ \theta_2 > 0; \tag{4.26}$$

$$u_{tt} = \theta_1 \boldsymbol{\Delta} u - \theta_2 \boldsymbol{\Delta}^2 u_t + \dot{W}, \ \theta_1 > 0, \ \theta_2 > 0. \tag{4.27}$$

The following table summarizes the results:

Asymptotic	Eq. (4.25)	Eq. (4.26)	Eq. (4.27)
$\Psi_{1,N}$	$N^{\frac{2}{d}+1}$	N	$\Upsilon_N(-2/d),\ d \geq 2$
$\Psi_{2,N}$	N	$N^{\frac{2}{d}+1}$	$N^{\frac{4}{d}+1}$

In equations (4.25)–(4.27), $\mathcal{A}_1$ and $\mathcal{B}_1$ are the *leading* operators, that is, $\alpha = \alpha_1$ and $\beta = \beta_1$. This, in particular, ensures that $\hat{\theta}_{2,N}$ is always consistent.

Let us now consider examples when $\mathcal{A}_1$ and $\mathcal{B}_1$ are not the leading operators:

$$u_{tt} = \big(\mathbf{\Delta} u + \theta_1 u\big) + \big(\mathbf{\Delta} u_t + \theta_2 u_t\big) + \dot{W},\ \theta_1 \in \mathbb{R},\ \theta_2 \in \mathbb{R}; \tag{4.28}$$

$$u_{tt} + \big(\mathbf{\Delta}^2 u + \theta_1 u\big) = \big(\theta_2 \mathbf{\Delta} u_t - \mathbf{\Delta}^2 u_t\big) + \dot{W}, \theta_1 \in \mathbb{R},\ \theta_2 \in R; \tag{4.29}$$

$$u_{tt} + \big(\mathbf{\Delta}^2 u + \theta_1 \mathbf{\Delta} u\big) = \big(\theta_2 u_t - \mathbf{\Delta}^2 u_t\big) + \dot{W},\ \theta_1 \in \mathbb{R},\ \theta_2 \in \mathbb{R}. \tag{4.30}$$

The following table summarizes the results:

Asymptotic	Eq. (4.28)	Eq. (4.29)	Eq. (4.30)
$\Psi_{1,N}$	$\Upsilon_N(-4/d),\ d \geq 4$	$\Upsilon_N(-8/d),\ d \geq 8$	$\Upsilon_N(-4/d),\ d \geq 4$
$\Psi_{2,N}$	$\Upsilon_N(-2/d),\ d \geq 2$	N	$\Upsilon_N(-4/d),\ d \geq 4$

As was mentioned in the Introduction, a multi-parameter estimation problem, such as

$$u_{tt} + \big(\theta_{11}\mathbf{\Delta}^2 u + \theta_{12}\mathbf{\Delta} u\big) = \big(\theta_{21} u_t - \theta_{22}\mathbf{\Delta}^2 u_t\big) + \dot{W},$$

can be studies in the same way.

5. Analysis of Estimators: General Case

As the proof of Theorem 4.1 shows, the key arguments involve a suitable law of large numbers. Verification of the corresponding conditions is straightforward in the algebraic case, but is impossible in the general case without additional assumptions. Indeed, as we work with weighted sums of independent random variables, we need some conditions on the weights for a law of large numbers to hold. In particular, the weights should not grow too fast: if $\xi_k,\ k \geq 1$, are iid standard Gaussian random variables, then the sequence $\{n^{-2}\sum_{k=1}^n n\xi_k^2,\ n \geq 1\}$ converges with probability one to 1/2, but $\{e^{-n}\sum_{k=1}^n e^k\xi_k^2,\ n \geq 1\}$ does not have a limit, even in probability.

Theorem B.2 in Appendix summarizes some of the laws of large numbers, and leads to the following

Definition 5.1. *The sequence $\{a_n,\ n \geq 1\}$ of positive numbers is called* `slowly increasing` *if*

$$\lim_{n\to\infty} \frac{\sum_{k=1}^n a_n^2}{\Big(\sum_{k=1}^n a_k\Big)^2} = 0. \tag{5.1}$$

The purpose of this definition is to simplify the statement of the main theorem (Theorem 5.2 below). It was not necessary in the algebraic case because the sequence $\{n^\gamma,\ n \geq 1\}$ is slowly increasing if and only if $\gamma \geq -1$. The reason for the terminology is that the sequence $\{e^{n^r},\ n \geq 1\}$ has property (5.1) if and only if $r < 1$. Further discussion of (5.1), including the connections with the weak law of large numbers, is after the proof of Theorem B.2 in Appendix.

In general, we have to replace (4.1) with

Condition 1. The sequence $\{\tau_k^2 M\big(T\mu_k(\theta_2)\big)/\lambda_k(\theta_1),\ k \geq 1\}$ is slowly increasing, and (4.4), with

Condition 2. The sequence $\{\nu_k^2 M\big(T\mu_k(\theta_2)\big),\ k\geq 1\}$ is slowly increasing.

Theorem 5.2. *Assume that* (2.3) *is hyperbolic.*

(1) *If Condition 1 holds, then*

$$\lim_{N\to\infty}\hat{\theta}_{1,N}=\theta_1 \ \text{ in probability;} \tag{5.2}$$

$$\lim_{N\to\infty}\sqrt{\Psi_{1,N}}\Big(\hat{\theta}_{1,N}-\theta_1\Big)=\xi_1 \text{ in distribution,} \tag{5.3}$$

where ξ_1 *is a standard Gaussian random variable.*

(2) *If Condition 2 holds then*

$$\lim_{N\to\infty}\hat{\theta}_{2,N}=\theta_2 \ \text{ in probability;} \tag{5.4}$$

$$\lim_{N\to\infty}\sqrt{\Psi_{2,N}}\Big(\hat{\theta}_{2,N}-\theta_2\Big)=\xi_2 \text{ in distribution,} \tag{5.5}$$

where ξ_2 *is a standard Gaussian random variable.*

(3) *If both Conditions 1 and 2 hold, then the random variables* ξ_1,ξ_2 *are independent.*

Proof. The main steps are the same as in the algebraic case (Theorem 4.1). In particular, (4.18) and (4.20) continue to hold as long as $\Psi_{1,N}\to\infty$ and $\Psi_{2,N}\to\infty$, respectively. The only difference is that Conditions 1 and 2 do not provide enough information about the almost sure behavior of $K_{12,N}/\mathbb{E}K_{12,N}$, and, in this general setting, there is no natural condition that would do that. As a result, in (4.17), the convergence is in probability rather than with probability one, and then, in both (4.15) and (4.16), convergence in probability will suffice. Conditions 1 and 2 ensure (4.15) and (4.16), respectively, but with convergence in probability rather than almost sure. This is a direct consequence of the weak law of large numbers.

In the case of (4.17), we have

$$\mathbb{E}|K_{12,N}|\leq\sum_{k=1}^{N}|\tau_k\nu_k|\,\mathbb{E}u_k^2(T)$$

and, for all sufficiently large k,

$$\mathbb{E}u_k^2(T)\leq\frac{4T}{\lambda_k(\theta_1)}M\big(T\mu_k(\theta_2)\big)\Big(1+\max\big(0,T\mu_k(\theta_2)\big)\Big),$$

because $xe^x-x\leq 4(e^x-x-1)(1+\max(0,x))$ for all $x\in\mathbb{R}$. Then

$$\lim_{N\to\infty}\frac{\mathbb{E}|K_{12,N}|}{\sqrt{\Psi_{1,N}\Psi_{2,N}}}=0. \tag{5.6}$$

Indeed, under Condition 1, (5.6) follows from

$$\frac{\big(\mathbb{E}|K_{12,N}|\big)^2}{\Psi_{1,N}\Psi_{2,N}}\leq\frac{16\sum_{k=1}^{N}\frac{\tau_k^2M\big(T\mu_k(\theta_2)\big)}{\lambda_k(\theta_1)}\frac{\Big(1+\max\big(0,T\mu_k(\theta_2)\big)\Big)^2}{\lambda_k(\theta_1)}}{T^2\sum_{k=1}^{N}\frac{\tau_k^2M\big(T\mu_k(\theta_2)\big)}{\lambda_k(\theta_1)}}$$

(Cauchy-Schwartz inequality) and

$$\lim_{k\to\infty} \frac{\big(1+\max\big(0, T\mu_k(\theta_2)\big)\big)^2}{\lambda_k(\theta_1)} = 0 \tag{5.7}$$

(hyperbolicity condition), while, under Condition 2, (5.6) follows from

$$\frac{\big(\mathbb{E}|K_{12,N}|\big)^2}{\Psi_{1,N}\Psi_{2,N}} \leq \frac{16\sum_{k=1}^N \nu_k^2 M\big(T\mu_k(\theta_2)\big)\ \frac{\big(1+\max\big(0,T\mu_k(\theta_2)\big)\big)^2}{\lambda_k(\theta_1)}}{T^2\sum_{k=1}^N \nu_k^2 M\big(T\mu_k(\theta_2)\big)}$$

(Cauchy-Schwartz inequality with a different arrangement of terms) and (5.7).

The interested reader can fill in the details in the rest of the proof. □

As an example, consider the model with $\kappa_k = e^{2k}$, $\tau_k = e^k$, $\rho_k = 0$, $\nu_k = \ln\ln(k+3)$ and assume that $\theta_1 > 0$, $\theta_2 > 0$. Then

$$\lambda_k = e^{2k} + \theta_1 e^k,\ \ \mu_k = \theta_2 \ln\ln(k+3),$$

so that $\tau_k^2/\lambda_k \sim 1$. Next, for all sufficiently large k,

$$\big(\ln(k+3)\big)^{T\theta_2/2} < M(T\mu_k) < \big(\ln(k+3)\big)^{T\theta_2},$$

and also $\nu_k^2 M(T\mu_k) \asymp \big(\ln(k+3)\big)^{T\theta_2}$. Using integral comparison, we conclude that, for all $r > 0$,

$$\sum_{k=1}^N \big(\ln k\big)^r \sim N\big(\ln N\big)^r.$$

Thus, both Condition 1 and Condition 2 hold. By Theorem 5.2, both $\hat{\theta}_{1,N}$ and $\hat{\theta}_{2,N}$ are consistent and asymptotically normal. Further computations show that

$$\Psi_{1,N} \asymp \frac{N\big(\ln N\big)^{T\theta_2}}{\big(\ln\ln N\big)^2},\quad \Psi_{2,N} \asymp N\big(\ln N\big)^{T\theta_2}.$$

6. Acknowledgement

The work of both authors was partially supported by the NSF Grant DMS-0803378.

Appendix A. Second-order equations with constant coefficients

Consider the initial value problem

$$\ddot{y}(t) - 2b\dot{y}(t) + a^2 y(t) = 0,\ \ y(0)=0,\ \ \dot{y}(0)=1. \tag{A.1}$$

With $2b = \mu_k(\theta_2)$ and $a^2 = \lambda_k(\theta_1)$, we recover (2.16); recall that $\lambda_k(\theta_1) > 0$ for all sufficiently large k. If $\ell = \sqrt{|b^2 - a^2|}$, then

$$y(t) = \begin{cases} \dfrac{\sin(\ell t)}{\ell}\, e^{bt}, & a^2 > b^2;\\ te^{bt}, & a^2 = b^2;\\ \dfrac{\sinh(\ell t)}{\ell}\, e^{bt}, & a^2 < b^2; \end{cases} \tag{A.2}$$

as usual, $\sinh x = (e^x - e^{-x})/2$. Note that if $b < 0$ and $b^2 > a^2$, then $b + \ell < 0$. The solution of the inhomogeneous equation

$$\ddot{x}(t) - 2b\dot{x}(t) + a^2 x(t) = f(t), \quad x(0) = \dot{x}(0) = 0$$

is then $x(t) = \int_0^t y(t-s) f(s) ds$.

Appendix B. Some limit theorems

To begin, let us recall Kolmogorov's strong law of large numbers.

Theorem B.1. *Let $\{\xi_k, k \geq 1\}$ be a sequence of independent random variables with $\mathbb{E}\xi_n^2 < \infty$. If $\{b_n \geq 1\}$ is an unbounded increasing sequence of real numbers ($b_n \nearrow +\infty$) and*
$\sum_{n \geq 1} b_n^{-2} \operatorname{Var}(\xi_n) < \infty$, then

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

with probability one.

Proof. See, for example, Shiryaev [24, Theorem IV.3.2]. □

The following laws of large numbers, both strong and weak, are often used in the current paper.

Theorem B.2 (Several Laws of Large Numbers)**.** *Let χ_k, $k \geq 1$, be independent random variables, each with zero mean and positive finite variance. If*

$$\sum_{k \geq 1} \mathbb{E}\chi_k^2 = +\infty, \tag{B.2}$$

then

$$\lim_{N \to \infty} \frac{\sum_{k=1}^N \chi_k}{\sum_{k=1}^N \mathbb{E}\chi_k^2} = 0 \text{ with probability one.} \tag{B.3}$$

Next, assume in addition that

$$\mathbb{E}\chi_k^4 \leq c_1 \Big(\mathbb{E}\chi_k^2 \Big)^2 \tag{B.4}$$

for all $k \geq 1$, with $c_1 > 0$ independent of k. Then

$$\lim_{N \to \infty} \frac{\sum_{k=1}^N \Big(\mathbb{E}\chi_k^2 \Big)^2}{\Big(\sum_{k=1}^N \mathbb{E}\chi_k^2 \Big)^2} = 0 \tag{B.5}$$

implies

$$\lim_{N \to \infty} \frac{\sum_{k=1}^N \chi_k^2}{\sum_{k=1}^N \mathbb{E}\chi_k^2} = 1 \text{ in probability,} \tag{B.6}$$

while

$$\sum_{n\geq 1}\frac{\Big(\mathbb{E}\chi_n^2\Big)^2}{\Big(\sum_{k=1}^{n}\mathbb{E}\chi_k^2\Big)^2}<\infty, \tag{B.7}$$

implies

$$\lim_{N\to\infty}\frac{\sum_{k=1}^{N}\chi_k^2}{\sum_{k=1}^{N}\mathbb{E}\chi_k^2}=1 \text{ with probability one.} \tag{B.8}$$

Proof. To prove (B.3), we take $\xi_n=\chi_n$ and $b_n=\sum_{k=1}^{n}\mathbb{E}\chi_k^2$ and apply Theorem B.1; note that convergence of $\sum_n b_n^{-2}\mathbb{E}\chi_n^2$ follows from divergence of $\sum_{k\geq 1}\mathbb{E}\chi_k^2$:

$$\sum_n \frac{\mathbb{E}\chi_n^2}{b_n^2}\leq \sum_n\left(\frac{1}{b_{n-1}}-\frac{1}{b_n}\right).$$

To prove (B.8), we take $\xi_n=\chi_n^2$ and $b_n=\sum_{k=1}^{n}\mathbb{E}\chi_k^2$, and again apply Theorem B.1; this time, we have to *assume* convergence of the series $\sum_n \operatorname{Var}\xi_n b_n^{-2}$. Finally, (B.6) follows from (B.5) and Chebyshev's inequality. □

Theorem B.2 shows that normalizing a sum of zero-mean random variables by the total *variance* will give in the limit zero with probability one as ling as the total variance is unbounded, while normalizing a sum of positive random variables by the total *mean* will give in the limit one only under some additional assumptions. For example, given a collection of iid standard normal random variables $\{\xi_k,\ k\geq 1\}$, the sequence $\big(\sum_{k=1}^{n}e^k\xi_k^2\big)/\big(\sum_{k=1}^{n}e^k\big)$ does not converge in probability as $n\to\infty$.

To understand the meaning of conditions (B.5) and (B.7), note that if $\xi_k,\ k\geq 1$, are iid non-negative random variables with $\mathbb{E}\xi_1=A>0$, then, taking in Theorem B.1 $b_n=\sum_{k=1}^{n}\mathbb{E}\xi_k=An$, we recover the classical strong law of large numbers:

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\xi_k=A$$

with probability one. In the second part of Theorem B.2, we want to establish a similar result when the random variables ξ_k are positive and independent, but not identically distributed. Condition (B.4) (which holds, for example, for Gaussian random variables) allows us to apply Theorem B.1 with $b_n=\sum_{k=1}^{N}\mathbb{E}\xi_k$. If $a_k:=\mathbb{E}\xi_k>0$ for all k, then conditions (B.4) and (B.7) become, respectively,

$$\sum_{n\geq 1}a_n=+\infty, \tag{B.9}$$

$$\sum_{n\geq 1}\frac{a_n^2}{(\sum_{k=1}^{n}a_k)^2}<\infty. \tag{B.10}$$

On the other hand, if

$$\lim_{n\to\infty}\frac{\sum_{k=1}^{n}a_k^2}{(\sum_{k=1}^{n}a_k)^2}=0, \tag{B.11}$$

then Chebyshev's inequality leads to a weak law of large numbers.

In general, (B.9) does not imply (B.10) or (B.11) (take $a_n = e^n$), nor does (B.10) imply (B.9) (take $a_n = 1/n^2$), but obviously (B.11) implies (B.9). An interested reader can also verify that the sequence $\{e^{\sqrt{n}},\ n \geq 1\}$ satisfies (B.11) but not (B.10). On the other hand, since we use (B.9) and (B.10) to prove a strong law of large numbers, and use (B.11) to prove a weak law of large numbers, it will be natural to expect that conditions (B.9) and (B.10) together are stronger than (B.11). Kronecker's Lemma (see [24, Lemma IV.3.2] with $b_n = \left(\sum_{k=1}^n a_n\right)^2,\ x_n = a_n^2/b_n$) shows that this is indeed the case: (B.9) and (B.10) imply (B.11).

We say that a sequence of positive numbers $\{a_n,\ n \geq 1\}$ is `slowly increasing` if condition (B.11) holds. The notion of a slowly increasing sequence simplifies the conditions for consistency and asymptotic normality of the estimators in the general (non-algebraic) setting. Related conditions in the context of the law of large numbers are in [1]. If $a_n = n^\gamma$, $\gamma \in \mathbb{R}$ (algebraic case), then (B.9) (that is, $\gamma \geq -1$) implies (B.10), which is the reason for the strong consistency in Theorem 4.1.

The following theorem is used to prove asymptotic normality of the estimators.

Theorem B.3 (A Martingale Central Limit Theorem). *Let $M_{i,n} = M_{i,n}(t)$, $t \geq 0$, $n \geq 1$, $i = 1, 2$, be two sequences of continuous square–integrable martingales. If, for some $T > 0$,*

$$\lim_{n\to\infty} \frac{\langle M_{i,n}\rangle(T)}{\mathbb{E}\langle M_n\rangle(T)} = 1,\ i = 1, 2, \quad \text{in probability,}$$

and

$$\lim_{n\to\infty} \frac{\langle M_{1,n}, M_{2,n}\rangle(T)}{\Big(\mathbb{E}\langle M_{1,n}\rangle(T)\Big)^{1/2}\Big(\mathbb{E}\langle M_{2,n}\rangle(T)\Big)^{1/2}} = 0 \quad \text{in probability,}$$

then

$$\lim_{n\to\infty} \left(\begin{array}{c} M_{1,n}(T)\big(\mathbb{E}\langle M_{1,n}\rangle(T)\big)^{-1/2} \\ M_{2,n}(T)\big(\mathbb{E}\langle M_{2,n}\rangle(T)\big)^{-1/2} \end{array} \right) = \mathcal{N}(0, I) \quad \text{in distribution,}$$

where $\mathcal{N}(0, I)$ is a two-dimensional vector whose components are independent standard Gaussian random variables.

Proof. If X_n and X are continuous square-integrable martingales with values in $\mathbb{R}^d$ such that X is a Gaussian process and $\lim_{n\to\infty}\langle X_n\rangle(T) = \langle X\rangle(T)$ in probability, then $\lim_{n\to\infty} X_n(T) = X(T)$ in distribution; recall that, for a vector-valued martingale $X = (X^{(1)}, \dots, X^{(d)})$, $\langle X\rangle(t)$ is the symmetric matrix with entries $\langle X^{(i)}, X^{(j)}\rangle(t)$. This is one of the central limit theorems for martingales; see, for example, Lipster and Shiryaev [16, Theorem 5.5.11]. The result now follows if we take

$$X_n(t) = \left(\begin{array}{c} X_{1,n} \\ X_{2,n} \end{array} \right),\quad X(t) = \left(\begin{array}{c} w_1(t)/\sqrt{T} \\ w_2(t)/\sqrt{T} \end{array} \right),$$

where

$$X_{i,n} = \frac{M_{i,n}(t)}{\big(\mathbb{E}\langle M_{i,n}\rangle(T)\big)^{1/2}},\ i = 1, 2,$$

and w_1, w_2 are independent standard Brownian motions. □

Corollary B.4. *Let $f_{i,k} = f_{i,k}(t)$, $t \geq 0$, $i = 1, 2$, $k \geq 1$ be continuous, square-integrable processes and $w_k = w_k(t)$ be independent standard Brownian motions. Define*

$$\eta_{i,N} = \frac{\sum_{k=1}^{N} \int_0^T f_{i,k}(t) dw_k(t)}{\left(\sum_{k=1}^{N} \mathbb{E} \int_0^T f_{i,k}^2(t) dt\right)^{1/2}}, \quad i = 1, 2.$$

If

$$\lim_{N\to\infty} \frac{\sum_{k=1}^{N} \int_0^T f_{i,k}^2(t) dt}{\sum_{k=1}^{N} \mathbb{E} \int_0^T f_{i,k}^2(t) dt} = 1 \quad \text{in probability, and} \tag{B.12}$$

$$\lim_{N\to\infty} \frac{\sum_{k=1}^{N} \mathbb{E} \left| \int_0^T f_{1,k}(t) f_{2,k}(t) dt \right|}{\left(\sum_{k=1}^{N} \mathbb{E} \int_0^T f_{1,k}^2(t) dt\right)^{1/2} \left(\sum_{k=1}^{N} \mathbb{E} \int_0^T f_{2,k}^2(t) dt\right)^{1/2}} = 0,$$

then

$$\lim_{N\to\infty} \begin{pmatrix} \eta_{1,N} \\ \eta_{2,N} \end{pmatrix} = \mathcal{N}(0, I) \quad \text{in distribution,}$$

where $\mathcal{N}(0, I)$ is a two-dimensional vector whose components are independent standard Gaussian random variables.

Proof. This follows from Theorem B.3 by taking

$$M_{i,n}(t) = \frac{\sum_{k=1}^{n} \int_0^t f_{i,k}(s) dw_k(s)}{\left(\mathbb{E} \int_0^T f_{i,k}^2(t) dt\right)^{1/2}},$$

because $\mathbb{E}\langle M_{i,n}\rangle(T) = 1$ and

$$\langle M_{1,n}, M_{2,n}\rangle(T) = \frac{\sum_{k=1}^{N} \int_0^T f_{1,k}(t) f_{2,k}(t) dt}{\left(\sum_{k=1}^{N} \mathbb{E} \int_0^T f_{1,k}^2(t) dt\right)^{1/2} \left(\sum_{k=1}^{N} \mathbb{E} \int_0^T f_{2,k}^2(t) dt\right)^{1/2}}.$$

□

References

1. A. Adler and A. Rosalsky, *On the weak law of large numbers for normed weighted sums of i.i.d. random variables*, Internat. J. Math. & Math. Sci. **14** (1991), no. 1, 191–202.
2. P.-L. Chow, *Stochastic partial differential equations*, Chapman & Hall/CRC Applied Mathematics and Nonlinear Science Series, Chapman & Hall/CRC, Boca Raton, FL, 2007.
3. Yu. Golubev, *The principle of penalized empirical risk in severely ill-posed problems*, Probab. Theory Related Fields **130** (2004), no. 1, 18–38.
4. V. Holdai and A. Korostelev, *Image reconstruction in multi-channel model under Gaussian noise*, Math. Methods Statist. **17** (2008), no. 3, 198–208.
5. M. Huebner, *A characterization of asymptotic behaviour of maximum likelihood estimators for stochastic PDE's*, Math. Methods Statist. **6** (1997), no. 4, 395–415.
6. M. Huebner, R. Z. Khas′minskiĭ, and B. L. Rozovskii, *Two examples of parameter estimation*, Stochastic Processes: A volume in honor of G. Kallianpur (S. Cambanis, J. K. Ghosh, R. L. Karandikar, and P. K. Sen, eds.), Springer, New York, 1992, pp. 149–160.
7. M. Huebner and B. Rozovskii, *On asymptotic properties of maximum likelihood estimators for parabolic stochastic PDE's*, Probab. Theory Related Fields **103** (1995), 143–163.

8. I. A. Ibragimov and R. Z. Khas′minskiĭ, *Statistical estimation: Asymptotic theory*, Applications of Mathematics, vol. 16, Springer, 1981.
9. ______, *Problems of estimating the coefficients of stochastic partial differential equations. I*, Teor. Veroyatnost. i Primenen. **43** (1998), no. 3, 417–438, translation in Theory Probab. Appl. 43 (1999), no. 3, 370–387.
10. ______, *Problems of estimating the coefficients of stochastic partial differential equations. II*, Teor. Veroyatnost. i Primenen. **44** (1999), no. 3, 526–554, translation in Theory Probab. Appl. 44 (2000), no. 3, 469–494.
11. ______, *Problems of estimating the coefficients of stochastic partial differential equations. III*, Teor. Veroyatnost. i Primenen. **45** (2000), no. 2, 209–235, translation in Theory Probab. Appl. 45 (2001), no. 2, 210–232.
12. A. Korostelev and O. Lepski, *On a multi-channel change-point problem*, Math. Methods Statist. **17** (2008), no. 3, 187–197.
13. A. Korostelev and G. Yin, *Estimation of jump points in high-dimensional diffusion modulated by a hidden Markov chain*, Math. Methods Statist. **15** (2006), no. 1, 88–102.
14. Yu. A. Kutoyants, *Identification of dynamical systems with small noise*, Mathematics and its Applications, vol. 300, Kluwer Academic Publishers, 1994.
15. ______, *Statistical inference for ergodic diffusion processes*, Springer Series in Statistics, Springer, 2004.
16. R. Sh. Liptser and A. N. Shiryaev, *Theory of martingales*, Mathematics and its Applications (Soviet Series), vol. 49, Kluwer Academic Publishers, Dordrecht, 1989.
17. ______, *Statistics of random processes, I: General theory, 2nd ed.*, Applications of Mathematics, vol. 5, Springer, 2001.
18. W. Liu and S. V. Lototsly, *Estimating speed and damping in the stochastic wave equation*, http://arxiv.org/abs/0810.0046.
19. S. V. Lototsky, *Parameter estimation for stochastic parabolic equations: asymptotic properties of a two-dimensional projection-based estimator*, Stat. Inference Stoch. Process. **6** (2003), no. 1, 65–87.
20. S. V. Lototsky and B. L. Rozovskii, *Spectral asymptotics of some functionals arising in statistical inference for SPDEs*, Stochastic Process. Appl. **79** (1999), no. 1, 69–94.
21. R. Mikulevicius and B. L. Rozovskii, *Uniqueness and absolute continuity of weak solutions for parabolic SPDE's*, Acta Appl. Math. **35** (1994), 179–192.
22. M. Pensky and T. Sapatinas, *Functional deconvolution in a periodic setting: uniform case*, Ann. Statist. **37** (2009), no. 1, 73–104.
23. Yu. Safarov and D. Vassiliev, *The asymptotic distribution of eigenvalues of partial differential operators*, Translations of Mathematical Monographs, vol. 155, American Mathematical Society, Providence, RI, 1997.
24. A. N. Shiryaev, *Probability, 2nd ed.*, Graduate Texts in Mathematics, vol. 95, Springer, 1996.

Current address, W. Liu: Department of Mathematics, USC, Los Angeles, CA 90089 USA, tel. (+1) 213 821 1480; fax: (+1) 213 740 2424
E-mail address, W. Liu: `liu5@usc.edu`

Current address, S. V. Lototsky: Department of Mathematics, USC, Los Angeles, CA 90089 USA, tel. (+1) 213 740 2389; fax: (+1) 213 740 2424
E-mail address, S. V. Lototsky: `lototsky@usc.edu`
URL: `http://www-rcf.usc.edu/~lototsky`