Bernoulli **9**(5), 2003, 745–762

Likelihood inference for a discretely observed stochastic partial differential equation

BO MARKUSSEN

Department of Applied Mathematics and Statistics, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen Ø, Denmark. E-mail: markusb@math.ku.dk.

Parabolic and hyperbolic stochastic partial differential equations in one-dimensional space have been proposed as models for the term structure of interest rates. The solution to these equations is reviewed, and their sample path properties are studied. In the parabolic case the sample paths essentially are Hölder continuous of order $\frac{1}{2}$ in space and $\frac{1}{4}$ in time, and in the hyperbolic case the sample paths essentially are Hölder continuous of order $\frac{1}{2}$ simultaneously in time and space. Parametric likelihood inference given an observation at discrete lattice points in time and space is also considered. The associated infinite-dimensional state-space model is described, and a finite-dimensional approximation is proposed. Conditions are presented under which the resulting approximate maximum likelihood estimator is asymptotically efficient when the number of observations in time increases to infinity at a fixed time step. The asymptotic distribution of the approximate likelihood ratio test for a parabolic equation against the hyperbolic alternative is found to be a truncated chi-square. Moreover, explicit moment estimators are derived which can be used as a starting point for a numerical optimization of the likelihood function.

Keywords: discrete observations; likelihood ratio test for parabolic equation; moment estimators; sample path properties; stochastic partial differential equation

1. Introduction

The purpose of this paper is to propose an approximate likelihood and study the asymptotic properties of the associated maximum likelihood estimator for the parameter $\theta = (\eta_1, \eta_2, \xi_0, \xi_1, \xi_2)$ given an observation at discrete points in time and space of the stationary solution of the parabolic $(\eta_2 = 0, \xi_2 > 0)$ or hyperbolic $(\eta_2 > 0, \xi_2 > 0)$ stochastic partial differential equation

$$\eta_2 \frac{\partial^2}{\partial t^2} V(t, x) + \eta_1 \frac{\partial}{\partial t} V(t, x) = \xi_0 V(t, x) + \xi_1 \frac{\partial}{\partial x} V(t, x) + \xi_2 \frac{\partial}{\partial x^2} V(t, x) + W_\xi(t, x), \quad (1)$$

$t \in \mathbb{R}$, $0 < x < 1$, with Dirichlet boundary conditions $V(t, 0) = V(t, 1) = 0$. Here the parameters satisfy $\eta_1, \xi_2 > 0$, $\eta_2 \geqslant 0$ and the stochastic disturbance term $W_\xi(t, x)$ is related to Brownian white noise $W(t, x)$ (see Holden *et al.* 1996), via the equation

1350–7265

$$W_\xi(t, x) = \exp\left(-\frac{\xi_1}{2\xi_2}x\right)W(t, x). \tag{2}$$

Parameter estimation in stochastic partial difference equations has been discussed in several papers, but to our knowledge always assuming an observation of the Fourier coefficients in the spectral representation of the equation. In Huebner and Rozovskii (1995) a continuous observation over a fixed time interval of the M first Fourier coefficients is assumed to be available. The asymptotic behaviour of the maximum likelihood estimator when $M \to \infty$ is investigated and shown to depend on the structure of the partial differential operators involved. In Piterbarg and Rozovskii (1997) a similar estimation problem given the observation of the M first Fourier coefficients at n discrete points in time is considered for $M \to \infty$. The stochastic partial differential equation (1) has been used in mathematical finance in models of the term structure for bonds of different maturity times; see Cont (1998) and Santa-Clara and Sornette (2001). In these models the spatial component represents time to maturity. In Cont (1998) it is argued that the short rate ($x = 0$) and the long rate ($x = 1$) can be modelled independently of the profile from the short rate to the long rate, and that the deviation from the average profile can be modelled by (1). Realistic data thus consist of observations at discrete points in time and space organized in a lattice. The spatial resolution N is usually fairly low, consisting of between 10 and 20 maturity times, and calculating the discrete Fourier transforms and using the Galerkin approximation (see Huebner and Rozovskii 1995) would be inadequate and result in biased estimates. In this paper we thus assume that the stochastic partial differential equation has been observed at discrete points in time and space. Moreover, we assume that the number of observations in space is fixed and examine the asymptotic behaviour when the number of observations in time tends to infinity at a fixed time step. The parabolic and hyperbolic equations have different properties – see the discussion in Cont (1998) – whence it is of interest to test the hypothesis of a parabolic equation against a hyperbolic equation.

The paper is organized as follows. In Section 2 we describe the parameters for which there exists a stationary solution to (1), give a representation of the solution, and investigate the smoothness properties of the sample paths. In Section 3 we propose an approximation of the likelihood, prove a uniform version of the local asymptotic normality property and derive the likelihood ratio test for a parabolic equation against a hyperbolic equation. In Section 4 we derive easily calculated moment estimators, which can be used as a starting point for a numerical optimization of the likelihood function. Moreover, we describe how to perform the likelihood ratio test for a parabolic equation against the hyperbolic alternative. Readers mainly interested in implementing the estimation and test procedures can skip the more technical parts of Section 3.

2. The stationary solution and its properties

If the partial differential equation (1) is either parabolic or hyperbolic, i.e. $\eta_2 \geqslant 0$ and $\xi_2 > 0$, then the associated Green function $G(t, x, y)$ is given by

$$G(t, x, y) = \sum_{k=1}^{\infty} T_k(t) X_k(x) X_k(y),$$

where $X_k(x)$ and λ_k, $k \in \mathbb{N}$, are respectively the eigenfunctions and eigenvalues of the differential operator $\xi_0 + \xi_1 \partial/\partial_x + \xi_2 \partial^2/\partial x^2$ with boundary conditions $X(0) = X(1) = 0$. Moreover, in the parabolic case $T_k(t)$ is the solution of the differential equation

$$\eta_1 T'(t) = \lambda_k T(t), \qquad \eta_1 T(0) = 1,$$

and in the hyperbolic case $T_k(t)$ is the solution of the differential equation

$$\eta_2 T''(t) + \eta_1 T'(t) = \lambda_k T(t), \qquad T(0) = 0, \qquad \eta_2 T'(0) = 1. \tag{3}$$

Solving these differential equations, we find that $X_k(x)$ and λ_k are given by

$$X_k(x) = \sqrt{2} \sin(\pi k x) \exp\left(-\frac{\xi_1}{2\xi_2} x\right), \qquad \lambda_k = \xi_0 - \frac{\xi_1^2}{4\xi_2} - \pi^2 k^2 \xi_2, \tag{4}$$

and $T_k(t)$ is given by

$$T_k(t) = \begin{cases} \dfrac{1}{\eta_1} \exp\left(\dfrac{\lambda_k}{\eta_1} t\right) & \text{if } \eta_2 = 0, \\ \dfrac{1}{\sqrt{\mu_k}} \left(\exp\left(\dfrac{-\eta_1 + \sqrt{\mu_k}}{2\eta_2} t\right) - \exp\left(\dfrac{-\eta_1 - \sqrt{\mu_k}}{2\eta_2} t\right)\right) & \text{if } \eta_2 > 0,\ \mu_k > 0, \\ t \dfrac{1}{\eta_2} \exp\left(-\dfrac{\eta_1}{2\eta_2} t\right) & \text{if } \eta_2 > 0,\ \mu_k = 0, \\ \dfrac{2}{\sqrt{-\mu_k}} \sin\left(\dfrac{\sqrt{-\mu_k}}{2\eta_2} t\right) \exp\left(-\dfrac{\eta_1}{2\eta_2} t\right) & \text{if } \eta_2 > 0,\ \mu_k < 0, \end{cases} \tag{5}$$

where $\mu_k = \eta_1^2 + 4\eta_2\lambda_k$. Thus if $\lambda_1 < 0 < \eta_1$, then there exists a unique stationary solution $V(t, x)$ to (1) given by

$$\begin{aligned} V(t, x) &= \int_{-\infty}^{t} \int_0^1 G(t - s, x, y) W_\xi(s, y) \exp\left(\frac{\xi_1}{\xi_2} y\right) \mathrm{d}y\, \mathrm{d}s \\ &= \int_{-\infty}^{t} \int_0^1 \sum_{k=1}^{\infty} T_k(t - s) X_k(x) X_k(y) W(s, y) \exp\left(\frac{\xi_1}{2\xi_2} y\right) \mathrm{d}y\, \mathrm{d}s \\ &= \sum_{k=1}^{\infty} U_k(t) X_k(x). \end{aligned} \tag{6}$$

Here the Fourier coefficient processes $U_k(t)$ are given by

$$U_k(t) = \int_{-\infty}^{t} T_k(t - s) W_k(s) \mathrm{d}s = \int_{-\infty}^{t} T_k(t - s) \mathrm{d}B_k(s), \tag{7}$$

and the pairwise independent white noise processes $W_k(t)$, $k \in \mathbb{N}$, and the pairwise independent two-sided normalized Brownian motions $B_k(t)$, $k \in \mathbb{N}$, are given by

$$W_k(t) = \int_0^1 X_k(y) W(t,\, y) \exp\left(\frac{\xi_1}{2\xi_2} y\right) \mathrm{d}y, \qquad B_k(t) = \int_0^t W_k(s) \mathrm{d}s.$$

Observe that the choice (2) of the noise process $W_\xi(t, x)$ is made in order for the Brownian motions $B_k(t)$ to become independent, thus facilitating the analysis. Whether this choice is adequate from a modelling point of view will of course depend on the particular application. We summarize the above considerations in the following theorem.

Theorem 1. *If the parameter* $\theta = (\eta_1, \eta_2, \xi_0, \xi_1, \xi_2)$ *belongs to the parameter space* $\Theta \subset \mathbb{R}^5$ *given by*

$$\eta_2 \geqslant 0, \qquad \eta_1, \xi_2 > 0, \qquad \xi_0, \xi_1 \in \mathbb{R}, \qquad \frac{\xi_1}{4\xi_2} + \pi^2 \xi_2 > \xi_0, \tag{8}$$

then $V(t, x) = \sum_{k=1}^{\infty} U_k(t) X_k(x)$ *is the unique stationary solution to* (1).

The pairwise independent coefficient processes $U_k(t)$, $k \in \mathbb{N}$, are characterized by the following proposition.

Proposition 1. *If* $\eta_2 = 0$, *then* $U_k(t)$ *is a stationary Ornstein–Uhlenbeck process and solves the stochastic differential equation*

$$\mathrm{d}U_k(t) = \frac{\lambda_k}{\eta_1} U_k(t) \mathrm{d}t + \frac{1}{\eta_1} \mathrm{d}B_k(t); \tag{9}$$

if $\eta_2 > 0$, *then* $U_k(t)$ *is the first component of the two-dimensional stationary Ornstein–Uhlenbeck process* $\overline{U}_k(t) = (U_k(t), \tilde{U}_k(t))$, *where*

$$\tilde{U}_k(t) = \int_{-\infty}^t T_k'(t-s) W_k(s) \mathrm{d}s = \int_{-\infty}^t T_k'(t-s) \mathrm{d}B_k(s).$$

Moreover, $\overline{U}_k(t)$ *solves the stochastic differential equation*

$$\mathrm{d}\overline{U}_k(t) = \begin{pmatrix} 0 & 1 \\ \dfrac{\lambda_k}{\eta_2} & -\dfrac{\eta_1}{\eta_2} \end{pmatrix} \overline{U}_k(t) \mathrm{d}t + \begin{pmatrix} 0 \\ \dfrac{1}{\eta_2} \end{pmatrix} \mathrm{d}B_k(t).$$

Proof. The statement is classical in the parabolic case (see Walsh 1986, p. 323). In the hyperbolic case we use the white noise calculus and the differential equation (3) to see that

$$\frac{\mathrm{d}}{\mathrm{d}t} U_k(t) = \int_{-\infty}^t T_k'(t-s) W_k(s) \mathrm{d}s + T_k(0) W_k(t) = \tilde{U}_k(t)$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{U}_k(t) = \int_{-\infty}^{t} T_k''(t-s)W_k(s)\mathrm{d}s + T_k'(0)W_k(t)$$
$$= \int_{-\infty}^{t} \left(\frac{\lambda_k}{\eta_2}T_k(t-s) - \frac{\eta_1}{\eta_2}T_k'(t-s)\right)W_k(s)\mathrm{d}s + \frac{1}{\eta_2}W_k(t)$$
$$= \frac{\lambda_k}{\eta_2}U_k(t) - \frac{\eta_1}{\eta_2}\tilde{U}_k(t) + \frac{1}{\eta_2}W_k(t).$$

If these equations are rewritten as Itô stochastic differential equations, then the stated equation for $\overline{U}_k(t)$ follows. □

Let $\Delta > 0$ be some fixed time step, and let $U_{k,\Delta}(t)$, $\tilde{U}_{k,\Delta}(t)$ be the time series associated with the kth coefficient processes at the discrete time points $t\Delta$, $t \in \mathbb{Z}$, i.e.

$$U_{k,\Delta}(t) = U_k(t\Delta), \qquad \tilde{U}_{k,\Delta}(t) = \tilde{U}_k(t\Delta), \qquad \tilde{U}_{k,\Delta}(t) = (U_{k,\Delta}(t), \tilde{U}_{k,\Delta}(t)).$$

Proposition 2. *If $\eta_2 = 0$, then $U_{k,\Delta}(t)$ is a first-order autoregressive process,*

$$U_{k,\Delta}(t) = \rho_{k,\Delta}U_{k,\Delta}(t+1) + \varepsilon_{k,\Delta}(t) \overset{d}{\sim} \mathcal{N}_1(0, \sigma_k^2),$$

where the innovations $\varepsilon_{k,\Delta}(t)$, $t \in \mathbb{Z}$, are independent and distributed as $\mathcal{N}_1(0, \sigma_k^2 - \rho_{k,\Delta}^2\sigma_k^2)$, and the autoregression coefficient and the stationary variance are given by

$$\rho_{k,\Delta} = \frac{T_k(\Delta)}{T_k(0)} = \exp\left(\frac{\lambda_k}{\eta_1}\Delta\right), \qquad \sigma_k^2 = \frac{1}{-2\eta_1\lambda_k}.$$

If $\eta_2 > 0$, then $\overline{U}_{k,\Delta}(t)$ is a first-order autoregressive process, i.e.

$$\overline{U}_{k,\Delta}(t) = \overline{\rho}_{k,\Delta}\overline{U}_{k,\Delta}(t-1) + \bar{\varepsilon}_{k,\Delta}(t) \overset{d}{\sim} \mathcal{N}_2(0, \overline{\sigma}_k^2),$$

where the innovations $\bar{\varepsilon}_{k,\Delta}(t) = (\varepsilon_{k,\Delta}(t), \tilde{\varepsilon}_{k,\Delta}(t))$, $t \in \mathbb{Z}$, are independent and distributed as $\mathcal{N}_2(0, \overline{\sigma}_k^2 - \overline{\rho}_{k,\Delta}\overline{\sigma}_k^2\overline{\rho}_{k,\Delta}^*)$, and the autoregression coefficient and the stationary variance are given by

$$\overline{\rho}_{k,\Delta} = \begin{pmatrix} \eta_2 T_k'(\Delta) + \eta_1 T_k(\Delta) & \eta_2 T_k(\Delta) \\ \lambda_k T_k(\Delta) & \eta_2 T_k'(\Delta) \end{pmatrix},$$

$$\overline{\sigma}_k^2 = \begin{pmatrix} \sigma_k^2 & 0 \\ 0 & \tilde{\sigma}_k^2 \end{pmatrix} = \begin{pmatrix} \dfrac{1}{-2\eta_1\lambda_k} & 0 \\ 0 & \dfrac{1}{2\eta_1\eta_2} \end{pmatrix}.$$

Moreover, the coefficient $\rho_{k,\Delta} = \eta_2 T_k'(\Delta) + \eta_1 T_k(\Delta)$ has the series expansion

$$\rho_{k,\Delta} = 1 + \frac{\lambda_k}{2\eta_2}\Delta^2 - \frac{\lambda_k \eta_1}{6\eta_2^2}\Delta^3 + o(\Delta^3).$$

Proof. We will first consider the hyperbolic case. Since $\overline{U}_k(t)$ is an Ornstein–Uhlenbeck process the time series $\overline{U}_{k,\Delta}(t)$ is a first-order autoregressive process. The stationary variance can be calculated by

$$\overline{\sigma}_k^2 = \mathrm{var}\left(\int_{-\infty}^{t} \begin{pmatrix} T_k(t-s) \\ T_k'(t-s) \end{pmatrix} \mathrm{d}B_k(s)\right) = \int_0^\infty \begin{pmatrix} T_k(s) \\ T_k'(s) \end{pmatrix} \begin{pmatrix} T_k(s) \\ T_k'(s) \end{pmatrix}^* \mathrm{d}s.$$

The components of $\overline{U}_k(t)$ are a priori seen to be independent, i.e.

$$\mathrm{cov}(U_k(t), \tilde{U}_k(t)) = \int_0^\infty T_k(s) T_k'(s) \mathrm{d}s = \frac{1}{2} T_k^2(s)|_{s=0}^{s=\infty} = 0,$$

and the marginal variances are found by direct computation of the integrals. In order to determine the autoregression coefficient we use the equation

$$\overline{U}_{k,\Delta}(t) = \underbrace{\int_{-\infty}^{(t-1)\Delta} \begin{pmatrix} T_k(t\Delta - s) \\ T_k'(t\Delta - s) \end{pmatrix} \mathrm{d}B_k(s)}_{=\bar{\rho}_{k,\Delta}\overline{U}_{k,\Delta}(t-1)} + \underbrace{\int_{(t-1)\Delta}^{t\Delta} \begin{pmatrix} T_k(t\Delta - s) \\ T_k'(t\Delta - s) \end{pmatrix} \mathrm{d}B_k(s)}_{\text{innovation } \bar{\varepsilon}_{k,\Delta}(t)}$$

to conclude that $\bar{\rho}_{k,\Delta}$ satisfies the equation

$$\bar{\rho}_{k,\Delta} \begin{pmatrix} T_k(u) \\ T_k'(u) \end{pmatrix} = \begin{pmatrix} T_k(u+\Delta) \\ T_k'(u+\Delta) \end{pmatrix}, \qquad u \geqslant 0. \tag{10}$$

Inserting $u = 0$, normalizing by $T_k'(0)$ and using the differential equation (3) gives the second column of $\bar{\rho}_{k,\Delta}$, i.e.

$$\bar{\rho}_{k,\Delta} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{T_k'(0)} \bar{\rho}_{k,\Delta} \begin{pmatrix} T_k(0) \\ T_k'(0) \end{pmatrix} = \begin{pmatrix} \eta_2 T_k(\Delta) \\ \eta_2 T_k'(\Delta) \end{pmatrix}.$$

Differentiating (10) with respect to u, inserting $u = 0$, normalizing by $T_k'(0)$ and using the differential equation (3) gives

$$\bar{\rho}_{k,\Delta} \begin{pmatrix} 1 \\ -\frac{\eta_1}{\eta_2} \end{pmatrix} = \frac{1}{T_k'(0)} \bar{\rho}_{k,\Delta} \begin{pmatrix} T_k'(0) \\ T_k''(0) \end{pmatrix} = \begin{pmatrix} \eta_2 T_k'(\Delta) \\ \lambda_k T_k(\Delta) - \eta_1 T_k'(\Delta) \end{pmatrix}$$

and hence the first column of $\bar{\rho}_{k,\Delta}$, i.e.

$$\bar{\rho}_{k,\Delta} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \bar{\rho}_{k,\Delta} \begin{pmatrix} 1 \\ -\frac{\eta_1}{\eta_2} \end{pmatrix} + \frac{\eta_1}{\eta_2} \bar{\rho}_{k,\Delta} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \eta_2 T_k'(\Delta) + \eta_1 T_k(\Delta) \\ \lambda_k T_k(\Delta) \end{pmatrix}.$$

The series expansion for the coefficient $\rho_{k,\Delta} = \eta_2 T_k'(\Delta) + \eta_1 T_k(\Delta)$ follows by the Taylor formula since

$$\frac{\mathrm{d}}{\mathrm{d}\Delta}\rho_{k,\Delta} = \eta_2 T_k''(\Delta) + \eta_1 T_k'(\Delta) = \lambda_k T_k(\Delta)$$

gives that $\frac{\mathrm{d}^j}{\mathrm{d}\Delta^j}\rho_{k,\Delta}$ for $j = 1, 2, 3$ and $\Delta = 0$ equals

$$\lambda_k T_k(0) = 0, \qquad \lambda_k T_k'(0) = \frac{\lambda_k}{\eta_2}, \qquad \lambda_k T_k''(0) = \frac{-\lambda_k \eta_1}{\eta_2^2}.$$

The easier parabolic case is analysed similarly. □

In the hyperbolic case the paths of the coefficient processes $U_k(t)$ are continuously differentiable; cf. Proposition 1. This fact suggests that the solution $V(t, x)$ is smoother in the hyperbolic case than in the parabolic case. We conclude this section by considering the sample path properties of $V(t, x)$.

Lemma 1. *There exists $\alpha < \infty$ such that, for $|t - s|$ sufficiently small,*

$$\mathrm{var}\,(V(t, x) - V(t, y)) \leqslant \alpha|x - y|,$$

$$\mathrm{var}(V(s, x) - V(t, x)) \leqslant \begin{cases} \alpha\sqrt{|t - s|}, & \text{if } \eta_2 = 0, \\ \alpha|t - s|\log(|t - s|^{-1}), & \text{if } \eta_2 > 0. \end{cases}$$

Proof. The first estimate and the second estimate in the parabolic case are given in Walsh (1986, Proposition 3.7). We thus consider the second estimate in the hyperbolic case. For fixed $\delta > 0$ the representation (6) gives

$$\mathrm{var}\,(V(t + \delta, x) - V(t, x)) = \sum_{k=1}^{\infty} X_k^2(x)\,\mathrm{var}\,(U_k(t + \delta) - U_k(t)).$$

The squared eigenfunctions $X_k^2(x)$ are uniformly bounded, and using Proposition 2 and the differential equation (3), we see that

$$\begin{aligned}
&\mathrm{var}\,(U_k(t + \delta) - U_k(t)) \\
&\quad = (\eta_1 T_k(\delta) + \eta_2 T_k'(\delta) - 1)^2 \sigma_k^2 + \eta_2^2 T_k(\delta)^2 \tilde{\sigma}_k^2 + \mathrm{var}\,(\varepsilon_{k,\delta}) \\
&\quad = \{(\eta_1 T_k(\delta) + \eta_2 T_k'(\delta) - 1)^2 + 1 - (\eta_1 T_k(\delta) + \eta_2 T_k'(\delta))^2\}\sigma_k^2 \\
&\quad = 2(1 - \eta_1 T_k(\delta) - \eta_2 T_k'(\delta))\sigma_k^2 \\
&\quad = \frac{1}{\eta_1 \lambda_k}(\eta_2 T_k'(\delta) + \eta_1 T_k(\delta) - \eta_2 T_k'(0) - \eta_1 T_k(0)) \\
&\quad = \frac{1}{\eta_1}\int_0^{\delta} T_k(u)\mathrm{d}u.
\end{aligned}$$

We wish to bound the sum of the variances by a constant times $\delta\log(\delta^{-1})$ for $\delta > 0$ sufficiently small. Thus, given $n \in \mathbb{N}$,

$$\sum_{k=1}^{n} \operatorname{var}(U_k(t+\delta) - U_k(t)) = \frac{1}{\eta_1} \sum_{k=1}^{n} \int_0^{\delta} T_k(u) \mathrm{d}u$$

$$\approx \frac{1}{\eta_1} \sum_{k=1}^{n} \int_0^{\delta} \frac{1}{k\pi} \frac{1}{\sqrt{\eta_2 \xi_2}} \sin\left(k\pi \sqrt{\frac{\xi_2}{\eta_2}} u \right) \exp\left(-\frac{\eta_1}{2\eta_2} u \right) \mathrm{d}u$$

$$= \frac{1}{\pi \eta_1} \frac{1}{\sqrt{\eta_2 \xi_2}} \int_0^{\delta} f_n\left(\pi \sqrt{\frac{\xi_2}{\eta_2}} u \right) \exp\left(-\frac{\eta_1}{2\eta_2} u \right) \mathrm{d}u,$$

where $f_n(x) = \sum_{k=1}^{n} \sin(kx)/k$. Some comments regarding the somewhat subtle approximation are needed. Firstly, since the kernals $T_k(u)$ are bounded we may replace finitely many of the integrands $T_k(u)$ by arbitrary bounded functions. Secondly, if k is large, then $\mu_k < 0$ and the replacement for $T_k(u)$ corresponds to the approximation $-\mu_k \approx 4\pi \eta_2 \xi_2 k^2$; see equation (5). Thirdly, the approximations of $T_k(u)$ for large k are small perturbations of lower order in k and may be disregarded. If $\delta^2 < \frac{1}{4} \eta_2 \xi_2^{-1}$, then $f_n(x)$ is only employed for $0 \leqslant x \leqslant \pi/2$, where

$$-f_n'(x) = -\sum_{k=1}^{n} \cos(kx) = \frac{1 + \cos((n+1)x)}{2} - \frac{1 + \cos(x)}{2\sin(x)} \sin((n+1)x)$$

$$\leqslant 1 + \frac{1 + \cos(x)}{2\sin(x)} \leqslant 1 + x^{-1}.$$

Since $f(\pi/2) \leqslant 1$ we thus have, for $0 \leqslant x \leqslant \pi/2$,

$$0 \leqslant f_n(x) = f_n\left(\frac{\pi}{2}\right) - \int_x^{\pi/2} f_n'(u) \mathrm{d}u \leqslant 1 + \int_x^{\pi/2} (1 + u^{-1}) \mathrm{d}u < 4 - \log(x).$$

Using this bound, we see that

$$\int_0^{\delta} f_n\left(\pi \sqrt{\frac{\xi_2}{\eta_2}} u \right) \exp\left(-\frac{\eta_1}{2\eta_2} u \right) \mathrm{d}u \leqslant \int_0^{\delta} \left(4 - \log\left(\pi \sqrt{\frac{\xi_2}{\eta_2}} u \right) \right) \mathrm{d}u$$

$$= 5\delta + \delta \log\left(\frac{1}{\pi} \sqrt{\frac{\eta_2}{\xi_2}} \delta^{-1} \right),$$

and so the desired bound follows. □

Theorem 2. *The solution $V(t, x)$ to the stochastic partial difference equation* (1) *has a version that is continuous in* (t, x). *If, for some fixed $t_0 < \infty$,*

$$H_t(\delta) = \sup_{x,y\in[0,1]:|x-y|\leqslant\delta} |V(t, x) - V(t, y)|, \qquad t \in [0, t_0],$$

$$H(\delta) = \sup_{s,t\in[0,t_0],x,y\in[0,1]:((s-t)^2+(x-y)^2)^{1/2}\leqslant\delta} |V(s, x) - V(t, y)|$$

are the moduli of continuity in space and time t and in time and space, respectively, then there exist $\alpha < \infty$ and random variables Y_t, $t \in [0, t_0]$, and Y with exponential moments such that, for $0 \leqslant \delta \leqslant 1$,

$$H_t(\delta) \leqslant Y_t\delta^{1/2} + \alpha\delta^{1/2}\sqrt{\log(\delta^{-1})},$$

$$H(\delta) \leqslant \begin{cases} Y\delta^{1/4} + \alpha\delta^{1/4}\sqrt{\log(\delta^{-1})}, & \text{if } \eta_2 = 0, \\ Y\delta^{1/2}\sqrt{\log(\delta^{-1})} + \alpha\delta^{1/2}\log(\delta^{-1}), & \text{if } \eta_2 > 0. \end{cases}$$

Proof. Using the variance estimates given in Lemma 1, the proof is similar to the proof of Walsh (1986, Theorem 3.8). □

Theorem 2 states that the solution $V(t, x)$ to (1) has paths that essentially are Hölder continuous of order $\frac{1}{2}$ in space and $\frac{1}{4}$ in time in the parabolic case, and of order $\frac{1}{2}$ in time and space in the hyperbolic case. The paths are thus substantially rougher in time in the parabolic case. The roughness present in the parabolic case is also reflected in the property that the process in that case has non-vanishing quartic variation (see Walsh 1986, Theorem 3.10).

3. Likelihood inference given discrete observations

In this section we give a time series representation of the statistical model given an observation of $V(t, x)$ at discrete points in time and space at the lattice points (t, x) given by

$$t = \Delta, 2\Delta, \ldots, n\Delta, \qquad x = \frac{a_1}{b}, \ldots, \frac{a_N}{b}, \tag{11}$$

where $\Delta > 0$ and $a_1, \ldots, a_N, b \in \mathbb{N}$, $a_1 < \ldots < a_N < b$, are fixed. Moreover, we describe an approximate maximum likelihood estimation procedure which is asymptotically efficient as $n \to \infty$, and we describe the associated likelihood ratio test for a parabolic equation against a hyperbolic equation.

Given an observation of $V(t, x)$ at the lattice points (11), let the N-dimensional time series $V_\Delta(t)$, $t \in \mathbb{Z}$, the $2b$-dimensional time series $U_\Delta(t)$, $t \in \mathbb{Z}$, and the matrices $\Xi \in \mathbb{R}^{N\times N}$ and $\Psi \in \mathbb{R}^{N\times 2b}$ be given by

$$V_\Delta(t) = V\left(t\Delta, \frac{a_j}{b}\right)_{j=1,\ldots,N}, \qquad U_\Delta(t) = \left(\sum_{j=0}^{\infty} U_{k+2bj,\Delta}(t)\right)_{k=1,\ldots,2b},$$

$$\Xi = \operatorname{diag}\left(\exp\left(-\frac{\xi_1}{2\xi_2}\frac{a_j}{b}\right)\right)_{j=1,\ldots,N}, \qquad \Psi = \left(\sqrt{2}\sin\left(\pi k \frac{a_j}{b}\right)\right)_{\substack{j=1,\ldots,N \\ k=1,\ldots,2b}}.$$

Then the observable time series $V_\Delta(t)$ has the state-space representation

$$V_\Delta(t) = \Xi\Psi U_\Delta(t). \tag{12}$$

The components of the time series $U_\Delta(t)$ are independent and given as infinite sums of independent time series. We propose to approximate the tail in the representation of $U_\Delta(t)$ by independent white noise. Thus, given a cut-off point $K \in \mathbb{N}$, variances $\tau_k^2 > 0$ approximating the variance of the tails, and random variables

$$\upsilon_k(t) \overset{d}{\sim} \mathcal{N}(0, \tau_k^2), \qquad k = 1, \ldots, , 2b, \; t = 1, \ldots, n, \tag{13a}$$

independent of everything else, we approximate the distribution of $V_\Delta(t)$ with the distribution of

$$\Xi\Psi\left(\textstyle\sum_{j\in\mathbb{N}_0: k+2bj<K} U_{k+2bj,\Delta}(t) + \upsilon_k(t)\right)_{k=1,\ldots,2b} = \Xi\overline{\Psi}\begin{pmatrix} (U_{k,\Delta}(t))_{k=1,\ldots,K} \\ (\upsilon_k(t))_{k=1,\ldots,2b} \end{pmatrix}, \tag{13b}$$

where the matrix $\overline{\Psi} \in \mathbb{R}^{N\times(K+2b)}$ is given by

$$\overline{\Psi} = \left(\sqrt{2}\sin\left(\pi k \frac{a_j}{b}\right)\right)_{\substack{j=1,\ldots,N \\ k=1,\ldots,K,1,\ldots,2b}}. \tag{13c}$$

Below we describe how to choose the cut-off point K and the variances τ_k^2 as functions of $n \in \mathbb{N}$ such that the resulting approximate likelihood is asymptotically efficient as $n \to \infty$. In order to measure the quality of the proposed approximation we introduce metrics on the space of N-dimensional matrices and on the space of N-dimensional spectral densities. The Schatten p-norm $\|A\|_p$ of a matrix $A \in \mathbb{C}^{N\times N}$ is defined as the l^p-norm of the eigenvalues of the positive semidefinite matrix $|A| = (A^*A)^{1/2}$, i.e. $\|A\|_\infty$ is the operator norm of A and

$$\|A\|_p = \left(\operatorname{tr}(A^*A)^{p/2}\right)^{p^{-1}}, \qquad p \in [1, \infty).$$

The L^p-norm $\|\psi\|_p$ of a matrix-valued function $\psi\colon (-\frac{1}{2}, \frac{1}{2}] \to \mathbb{C}^{N\times N}$ is defined as the usual L^p-norm of the real-valued function $\|\psi(\cdot)\|_p$, i.e. $\|\psi\|_\infty$ is the essential supremum of $\|\psi(\cdot)\|_\infty$ and

$$\|\psi\|_p = \left(\int_{-1/2}^{1/2} \|\psi(\omega)\|_p^p \,\mathrm{d}\omega\right)^{p^{-1}}, \qquad p \in [1, \infty).$$

These L^p-norms behave much like the usual L^p-norms and, in particular, satisfy the Hölder inequality. The remaining analysis relies on the following theorem proved in Markussen (2001).

Theorem 3. *Let $V(t)$, $t \in \mathbb{Z}$, be an N-dimensional Gaussian time series, i.e.*

$$\mathbb{V}_n = (V(1), \ldots, V(n))^* \overset{d}{\sim} \mathcal{N}_{n\times N}(0, \Sigma_n(\psi)),$$

where $\Sigma_n(\psi)$ is the Toeplitz matrix associated to the spectral density ψ. For each $n \in \mathbb{N}$, let $\psi_{n,\theta}$, $\theta \in \Theta \subseteq \mathbb{R}^d$, be a family of spectral densities and let $l_n(\theta)$ be the corresponding log-likelihood function given by

$$l_n(\theta) = -\tfrac{1}{2}\log\det\Sigma_n(\psi_{n,\theta}) - \tfrac{1}{2}\operatorname{tr}(\Sigma_n(\psi_{n,\theta})^{-1}\mathbb{V}_n\mathbb{V}_n^*). \tag{14}$$

Let $\theta_0 \in \Theta$ be some fixed and unknown parameter, and suppose that the Fisher information matrix $\mathcal{J}$ given by

$$\mathcal{J} = \lim_{n\to\infty}\left(\frac{1}{2}\int_{-1/2}^{1/2}\operatorname{tr}\left(\psi_{n,\theta_0}^{-1}(\omega)\partial_i\psi_{n,\theta_0}(\omega)\psi_{n,\theta_0}^{-1}(\omega)\partial_j\psi_{n,\theta_0}(\omega)\right)\mathrm{d}\omega\right)_{i,j=1,\ldots,d}$$

exists and is positive definite. If $n^{1/2}\|\psi - \psi_{n,\theta_0}\|_2$ vanishes as $n \to \infty$ and some additional regularity conditions are satisfied, then the maximum likelihood estimator $\hat{\theta}_n =$ $\operatorname{argmax}_{\theta\in\Theta} l_n(\theta)$ is a $\sqrt{n}$-consistent estimator for θ_0 and there exists a sequence G_n, $n \in \mathbb{N}$, of d-dimensional random variables converging in distribution to $\mathcal{N}_d(0, \mathcal{J})$ such that

$$\mathrm{E}\left(\sup_{u\in\mathbb{R}^d:|u|\leqslant r,\theta_0+n^{-1/2}u\in\Theta}|l_n(\theta_0+n^{-1/2}u) - l_n(\theta_0) - \left(u^*G_n - \tfrac{1}{2}u^*\mathcal{J}u\right)|\right) \tag{15}$$

vanishes for every $r > 0$ as $n \to \infty$. The local asymptotical normality property (15) *implies that $\hat{\theta}_n$ is asymptotically efficient in the sense of Hájek and Le Cam.*

Let $\Theta \subset \mathbb{R}^5$ be the parameter space described in Theorem 1, and suppose that the stochastic partial differential equation described by the parameter $\theta_0 \in \Theta$ has been observed at the lattice points (11). Then the additional regularity conditions are satisfied for the associated time series model (see Markussen 2001), and Theorem 3 applies if

$$n^{1/2}\|\psi_{\theta_0} - \psi_{n,\theta_0}\|_2 \underset{n\to\infty}{\to} 0,$$

where ψ_{θ_0} is the exact spectral density and ψ_{n,θ_0} is the nth approximate spectral density. Using the state-space representation (12) and the approximate state-space representation (13), we see that

$$\begin{aligned}\psi_{\theta_0} &= \Xi\Psi\operatorname{diag}(\textstyle\sum_{j=0}^{\infty}\varphi_{k+2bj,\Delta})_{k=1,\ldots,2b}\Psi^*\Xi^*,\\ \psi_{n,\theta_0} &= \Xi\Psi\operatorname{diag}(\textstyle\sum_{j\in\mathbb{N}_0:k+2bj<K}\varphi_{k+2bj,\Delta} + \tau_k^2)_{k=1,\ldots,2b}\Psi^*\Xi^*,\end{aligned} \tag{16}$$

where $\varphi_{k,\Delta}$ is the spectral density for the coefficient process $U_{k,\Delta}(t)$.

Lemma 2. *In the parabolic case the spectral density $\varphi_{k,\Delta}$ satisfies the bound*

$$\frac{1}{-2\eta_1\lambda_k}\frac{1-\mathrm{e}^{\lambda_k\Delta/\eta_1}}{1+\mathrm{e}^{\lambda_k\Delta/\eta_1}} < \varphi_{k,\Delta}(\omega) < \frac{1}{-2\eta_1\lambda_k}\frac{1+\mathrm{e}^{\lambda_k\Delta/\eta_1}}{1-\mathrm{e}^{\lambda_k\Delta/\eta_1}}.$$

In the hyperbolic case and for $\mu_k < 0$, with $k > (1/2\pi)\sqrt{\eta_1^2/\eta_2\xi_2 + 4\xi_0/\xi_2 + \xi_1^2/\xi_2^2}$, the spectral density $\varphi_{k,\Delta}$ satisfies

$$\varphi_{k,\Delta}(\omega) \geqslant \frac{1}{-2\eta_1\lambda_k}\frac{e^{\eta_1\Delta/2\eta_2}-1}{e^{\eta_1\Delta/2\eta_2}+1} - \frac{1}{-\lambda_k\sqrt{-\mu_k}}\left(\exp\left(\frac{\eta_1}{2\eta_2}\Delta\right) - \exp\left(\frac{-\eta_1}{2\eta_2}\Delta\right)\right)^{-1},$$

$$\varphi_{k,\Delta}(\omega) \leqslant \frac{1}{-2\eta_1\lambda_k}\frac{e^{\eta_1\Delta/2\eta_2}+1}{e^{\eta_1\Delta/2\eta_2}-1} + \frac{1}{-\lambda_k\sqrt{-\mu_k}}\left(\exp\left(\frac{\eta_1}{2\eta_2}\Delta\right) - \exp\left(\frac{-\eta_1}{2\eta_2}\right)\right)^{-1}.$$

Proof. In the parabolic case the first-order autoregressive process $U_{k,\Delta}(t)$ has spectral density

$$\begin{aligned}\varphi_{k,\Delta}(\omega) &= \sigma_k^2(1-\rho_{k,\Delta}^2)(1 - 2\rho_{k,\Delta}\cos(2\pi\omega) + \rho_{k,\Delta}^2) \\ &= \frac{1}{-2\eta_1\lambda_k}\left(1 - \exp\left(\frac{2\lambda_k}{\eta_1}\Delta\right)\right)\left(1 - 2\exp\left(\frac{\lambda_k}{\eta_1}\Delta\right)\cos(2\pi\omega) + \exp\left(\frac{2\lambda_k}{\eta_1}\Delta\right)\right)^{-1},\end{aligned} \tag{17}$$

and so the stated bound immediately follows. In the hyperbolic case we first do a linear transformation of the autoregressive process $\overline{U}_{k,\Delta}(t)$. Let the matrix A and the inverse A^{-1} be given by

$$A = \begin{pmatrix} 1 & 0 \\ \dfrac{\eta_1}{\sqrt{-\mu_k}} & \dfrac{2\eta_2}{\sqrt{-\mu_k}} \end{pmatrix}, \qquad A^{-1} = \begin{pmatrix} 1 & 0 \\ \dfrac{-\eta_1}{2\eta_2} & \dfrac{\sqrt{-\mu_k}}{2\eta_2} \end{pmatrix},$$

and consider the autoregression

$$A\overline{U}_{k,\Delta}(t+1) = (A\overline{\rho}_{k,\Delta}A^{-1})A\overline{U}_{k,\Delta}(t) + A\overline{\varepsilon}_{k,\Delta}(t+1).$$

Using Proposition 2, we see that the autoregression coefficient $A\overline{\rho}_{k,\Delta}A^{-1}$ is given by

$$A\overline{\rho}_{k,\Delta}A^{-1} = \begin{pmatrix} \dfrac{\eta_1}{2}T_k(\Delta) + \eta_2 T'_k(\Delta) & \dfrac{\sqrt{-\mu_k}}{2}T_k(\Delta) \\ \dfrac{-\sqrt{-\mu_k}}{2}T_k(\Delta) & \dfrac{\eta_1}{2}T_k(\Delta) + \eta_2 T'_k(\Delta) \end{pmatrix}.$$

The powers of $A\overline{\rho}_{k,\Delta}A^{-1}$ can be calculated easily since $A\overline{\rho}_{k,\Delta}A^{-1}$ is a scaled rotation matrix. The differential equation (3) gives

$$\frac{d}{d\Delta}\left\{\exp\left(\frac{\eta_1}{2\eta_2}\Delta\right)\left(\frac{\eta_1}{2}T_k(\Delta) + \eta_2 T'_k(\Delta)\right)\right\} = \frac{-\sqrt{-\mu_k}}{2\eta_2}\left(\exp\left(\frac{\eta_1}{2\eta_2}\Delta\right)\frac{-\sqrt{-\mu_k}}{2}T_k(\Delta)\right),$$

$$\frac{d}{d\Delta}\left\{\exp\left(\frac{\eta_1}{2\eta_2}\Delta\right)\frac{\sqrt{-\mu_k}}{2}T_k(\Delta)\right\} = \frac{\sqrt{-\mu_k}}{2\eta_2}\left\{\exp\left(\frac{\eta_1}{2\eta_2}\Delta\right)\left(\frac{\eta_1}{2}T_k(\Delta) + \eta_2 T'_k(\Delta)\right)\right\},$$

whence

$$A\overline{\rho}_{k,\Delta}A^{-1} = \exp\left(\frac{-\eta_1}{2\eta_2}\Delta\right)\begin{pmatrix} \cos\left(\frac{\sqrt{-\mu_k}}{2\eta_2}\Delta\right) & \sin\left(\frac{\sqrt{-\mu_k}}{2\eta_2}\Delta\right) \\ -\sin\left(\frac{\sqrt{-\mu_k}}{2\eta_2}\Delta\right) & \cos\left(\frac{\sqrt{-\mu_k}}{2\eta_2}\Delta\right) \end{pmatrix}.$$

Moreover, the variance of the innovations $A\overline{\varepsilon}_{k,\Delta}(t)$ is given by

$$A\overline{\sigma}_k^2 A^* = \sigma_k^2 \begin{pmatrix} 1 & \frac{\eta_1}{\sqrt{-\mu_k}} \\ \frac{\eta_1}{\sqrt{-\mu_k}} & \frac{\eta_1^2 - 2\eta_2\lambda_k}{-\mu_k} \end{pmatrix}.$$

The first component of the time series $A\overline{U}_{k,\Delta}(t)$, i.e. the coefficient process $U_{k,\Delta}(t)$, thus has spectral density $\varphi_{k,\Delta}$ given by

$$\begin{aligned} \varphi_{k,\Delta}(\omega) = \sigma_k^2 + 2\sigma_k^2 \sum_{t=1}^{\infty} \exp\left(\frac{-\eta_1}{2\eta_2}\Delta t\right)\cos\left(\frac{\sqrt{-\mu_k}}{2\eta_2}\Delta t\right)\cos(2\pi\omega t) \\ + 2\frac{\eta_1}{\sqrt{-\mu_k}}\sigma_k^2 \sum_{t=1}^{\infty} \exp\left(\frac{-\eta_1}{2\eta_2}\Delta t\right)\sin\left(\frac{\sqrt{-\mu_k}}{2\eta_2}\Delta t\right)\cos(2\pi\omega t), \end{aligned} \tag{18}$$

and thus the stated bounds follow by inserting the trigonometric relations

$$2\cos(x)\cos(y) = \cos(x+y) + \cos(x-y),$$
$$2\sin(x)\cos(y) = \sin(x+y) + \sin(x-y)$$

and using the bounds

$$-\frac{1}{e^a+1} \leqslant \sum_{n=1}^{\infty} e^{-an}\cos(bn) = \frac{e^a\cos(b) - 1}{1 - 2e^a\cos(b) + e^{2a}} \leqslant \frac{1}{e^a - 1},$$

$$-\frac{1}{e^a - e^{-a}} \leqslant \sum_{n=1}^{\infty} e^{-an}\sin(bn) = \frac{e^a\sin(b)}{1 - 2e^a\cos(b) + e^{2a}} \leqslant \frac{1}{e^a - e^{-a}}.$$

$\square$

The required cut-off point and white noise variances are given in the following lemma.

Lemma 3. *The quantity* $n^{1/2}\|\psi_{\theta_0} - \psi_{n,\theta_0}\|_2$ *vanishes as* $n \to \infty$ *if the cut-off point K and the white noise variances* τ_k^2*, as functions of* $n \in \mathbb{N}$*, in the parabolic case are given by*

$$K(n) = \left\lceil \frac{1}{\pi}\sqrt{\frac{\eta_1}{2\xi_2\Delta}}\sqrt{\log n}\right\rceil,$$

$$\tau_k^2(n) = \frac{1}{2\eta_1}\left(\frac{1}{2b\pi^2\xi_2 n^{1/4}} + \sum_{j\in\mathbb{N}_0:K(n)\leqslant k+2bj<n^{1/4}} \frac{1}{-\lambda_{k+2bj}}\right), \tag{19a}$$

and in the hyperbolic case for some arbitrary $\alpha > \frac{1}{4}$ *are given by*

$$K(n) = \lceil n^{\alpha}\rceil, \qquad \tau_k^2(n) = \frac{1}{4b\pi^2\eta_1\xi_2 n^{\alpha}}. \tag{19b}$$

Proof. Given fixed $\alpha > 0$, let the factors $\gamma_{\alpha,n}$ be defined by

$$\gamma_{\alpha,n} = \left(\pi^2\xi_2 n^{\alpha}\sum_{k=\lceil n^{\alpha}\rceil}^{\infty}\frac{1}{-\lambda_k}\right)^{-1} = 1 + o(n^{-\alpha}),$$

where the estimate follows from the approximation $-\lambda_k \approx \pi^2 k^2\xi_2$. We first consider the parabolic case. Using equation (16), the Hölder and the triangular inequality, we see that $n^{1/2}\|\psi_{\theta_0} - \psi_{n,\theta_0}\|_2$ is bounded by

$$n^{1/2}\|\Xi\|_{\infty}^2\|\Psi\|_{\infty}^2\left(\sum_{k=K(n)}^{\lfloor n^{1/4}\rfloor}\left\|\varphi_{k,\Delta} - \frac{1}{-2\eta_1\lambda_k}\right\|_2 + \sum_{k=\lceil n^{1/4}\rceil}^{\infty}\left\|\varphi_{k,\Delta} - \frac{1}{-2\eta_1\lambda_k}\gamma_{1/4,n}\right\|_2\right),$$

which by Lemma 2 is bounded by

$$n^{1/2}\|\Xi\|_{\infty}^2\|\Psi\|_{\infty}^2\left(\sum_{k=K(n)}^{\infty}\frac{1}{-\eta_1\lambda_k}\frac{e^{\lambda_k\Delta/\eta_1}}{1-e^{\lambda_k\Delta/\eta_1}} + \sum_{k=\lceil n^{1/4}\rceil}^{\infty}\frac{1}{-2\eta_1\lambda_k}|\gamma_{1/4,n} - 1|\right). \tag{20}$$

The term $\sum_{k=\lceil n^{1/4}\rceil}^{\infty}(1/-2\eta_1\lambda_k)|\gamma_{1/4,n} - 1|$ is of order $o(n^{-1/2})$, and using the choice of $K(n)$ given by (19a), we see that

$$\exp\left(\frac{\lambda_{K(n)}}{\eta_1}\Delta\right) \approx \exp\left(-\pi^2\frac{\xi_2}{\eta_1}\Delta K(n)^2\right) \leqslant n^{-1/2}.$$

The bound (20) thus vanishes as $n\to\infty$. In the hyperbolic case similar considerations lead us to consider the estimate

$$n^{1/2}\sum_{k=\lceil n^{\alpha}\rceil}^{\infty}\left\|\varphi_{k,\Delta} - \frac{1}{-2\eta_1\lambda_k}\gamma_{\alpha,n}\right\|_2 \leqslant n^{1/2}\sum_{k=\lceil n^{\alpha}\rceil}\frac{C_1}{-\lambda_k\sqrt{-\mu_k}} \leqslant C_2 n^{1/2-2\alpha}$$

for some constants C_1, C_2. If $\alpha > \frac{1}{4}$, then the right-hand side of the last inequality vanishes as $n\to\infty$ and the statement of the lemma follows. □

The following theorem states the asymptotic properties of the proposed approximate likelihood function for $n\to\infty$ and fixed $N\in\mathbb{N}$.

Theorem 4. *Let* $\theta_0 = (\eta_1, \eta_2, \xi_0, \xi_1, \xi_2)$ *be the true parameter, and let the cut-off points* $K(n)$ *and the white noise variances* $\tau_k^2(n)$ *be given by* (19). *Then* $\hat{\theta}_n = \mathrm{argmax}_{\theta\in\Theta} l_n(\theta)$ *is an asymptotically efficient estimator for* θ_0 *and* $\sqrt{n}(\hat{\theta}_n - \theta_0)$ *converges in distribution to* $\mathcal{N}_5(0, \mathcal{J}^{-1})$, *where the Fisher information* $\mathcal{J}$ *is given by*

$$\mathcal{J} = \left(\frac{1}{2}\int_{-1/2}^{1/2} \mathrm{tr}(\psi_{\theta_0}^{-1}(\omega)\partial_i\psi_{\theta_0}(\omega)\psi_{\theta_0}^{-1}(\omega)\partial_j\psi_{\theta_0}(\omega))\mathrm{d}\omega\right)_{i,j=1,\ldots,5} \tag{21}$$

and the spectral density ψ_{θ_0} *is given by* (16), (17) *and* (18). *If* $\eta_2 = 0$, *then the likelihood ratio test statistic* χ_n *for a parabolic against a hyperbolic equation satisfies*

$$\chi_n = 2\sup_{\theta\in\Theta} l_n(\theta) - 2\sup_{\theta\in\Theta:\eta_2=0} l_n(\theta) \xrightarrow[n\to\infty]{\mathcal{D}} \frac{1}{2}\epsilon_0 + \frac{1}{2}\chi_1^2, \tag{22}$$

where ϵ_0 *is the point measure in* 0 *and* χ_1^2 *is the chi-square distribution with one degree of freedom.*

Proof. Using Lemma 3, we see that the conditions of Theorem 3 are satisfied. The properties of the estimator $\hat{\theta}_n$ follow from the local asymptotic normality property (15), and the asymptotic distribution of the likelihood ratio test statistic follows since the hypothesis $\eta_2 = 0$ is a hyperplane on the border of the parameter space; see Self and Liang (1987). □

4. Moment estimators and implementation

In this section we describe how the analysis given in Section 3 can be used to perform maximum likelihood estimation given an observation of the stochastic partial differential equation at the lattice points (11). Moreover, we describe how to perform the likelihood ratio test for a parabolic equation against the alternative of a hyperbolic equation.

We first give an informal derivation of preliminary moment estimators which can be used as a starting point for a numerical optimization of the likelihood function. Using equation (6), we have

$$V(t, x) = \sum_{k=1}^{\infty} U_k(t)X_k(x) = \sqrt{2}\exp\left(\frac{-\xi_1}{2\xi_2}x\right)\sum_{k=1}^{\infty} U_k(t)\sin(\pi kx),$$

whence the ergodic theorem for the processes $U_k(t)$ and Proposition 2 give

$$\frac{1}{n}\sum_{t=1}^{n} V(t\Delta, x)^2 = 2\exp\left(-\frac{\xi_1}{\xi_2}x\right)\frac{1}{n}\sum_{t=1}^{n}\left(\sum_{k=1}^{\infty} U_k(t\Delta)\sin(\pi kx)\right)^2$$

$$\overset{\text{prob}}{\approx} 2\exp\left(-\frac{\xi_1}{\xi_2}x\right)\sum_{k=1}^{\infty}\sigma_k^2\sin^2(\pi kx),$$

where the approximation in probability is good for large n. Inserting

$$\sigma_k^2 = \frac{1}{-2\eta_1\lambda_k} = \frac{\lambda_1}{\lambda_k}\sigma_1^2$$

and

$$\lambda_k = \xi_0 - \frac{\xi_1^2}{4\xi_2} - \pi^2 k^2 \xi_2 = \frac{k^2-1}{3}\lambda_2 - \frac{k^2-4}{3}\lambda_1$$

gives

$$\exp\left(-\frac{\xi_1}{2\xi_2}x\right) \overset{\text{prob}}{\approx} \frac{1}{\sqrt{6\sigma_1^2}}\sqrt{\frac{n^{-1}\sum_{t=1}^n V(t\Delta, x)^2}{\sum_{k=1}^\infty \sin^2(\pi k x)/\{(k^2-1)\lambda_2/\lambda_1 - k^2 + 4\}}}, \tag{23}$$

whence the unobserved coefficient processes $U_{k,\Delta}(t)$ can be estimated via

$$\begin{aligned}\sqrt{3\sigma_1^2}U_{k,\Delta}(t) &= \sqrt{3\sigma_1^2}\int_0^1 V(t\Delta, x)X_k(x)\exp\left(\frac{\xi_1}{\xi_2}x\right)\mathrm{d}x \\ &\approx \sqrt{3\sigma_1^2}\frac{1}{N}\sum_{j=1}^N V\left(t\Delta, \frac{a_j}{b}\right)\sqrt{2}\sin\left(\pi k\frac{a_j}{b}\right)\exp\left(\frac{\xi_1}{2\xi_2}\frac{a_j}{b}\right) \\ &\overset{\text{prob}}{\approx} \frac{1}{N}\sum_{j=1}^N V\left(t\Delta, \frac{a_j}{b}\right)\sin\left(\pi k\frac{a_j}{b}\right)\sqrt{\frac{n^{-1}\sum_{s=1}^n V\left(s\Delta, \frac{a_j}{b}\right)^2}{\sum_{i=1}^\infty \sin^2(\pi i a_j/b)/\{(i^2-1)\lambda_2/\lambda_1 - i^2+4\}}}.\end{aligned} \tag{24}$$

The first approximation in (24) is good for small k, large N and observation points a_j/b equally spaced in the interval [0, 1]. In the applications we have in mind N will usually be small, and this approximation will hence be bad. We will, however, use (24) for $k = 1, 2$ in order to derive preliminary estimates. For numerical implementations of (24) it is, moreover, necessary to replace the infinite sum over $i \in \mathbb{N}$ by some finite sum. Below we list estimators for the parameters λ_2/λ_1, $\eta_1\lambda_1$, η_1/λ_1 in the parabolic case, and λ_2/λ_1, $\eta_1\lambda_1$, η_2/λ_1, η_1/λ_1 in the hyperbolic case, which can be computed successively using the approximation (24) for the quantities $\sqrt{3\sigma_1^2}U_{k,\Delta}(t)$. The quotient λ_2/λ_1 can be estimated solving the approximative identity

$$\frac{\lambda_2}{\lambda_1} = \frac{\sigma_1^2}{\sigma_2^2} \overset{\text{prob}}{\approx} \frac{n^{-1}\sum_{t=1}^n 3\sigma_1^2 U_{1,\Delta}(t)^2}{n^{-1}\sum_{t=1}^n 3\sigma_1^2 U_{2,\Delta}(t)^2} \tag{25}$$

numerically in $\lambda_2/\lambda_1 > 1$. Then the product $\eta_1\lambda_1$ can be estimated by

$$-\eta_1\lambda_1 = \frac{1}{\sqrt{4\sigma_1^4}} \overset{\text{prob}}{\approx} \frac{1}{\sqrt{\frac{4}{3}n^{-1}\sum_{t=1}^n 3\sigma_1^2 U_{1,\Delta}(t)^2}}, \tag{26}$$

where the estimate for λ_2/λ_1 is inserted in (24). In the parabolic case the quotient λ_1/η_1 can be estimated by

$$\frac{\lambda_1}{\eta_1} = \frac{\log(\rho_{1,\Delta})}{\Delta} \overset{\text{prob}}{\approx} \frac{1}{\Delta} \log\left(\frac{(n-1)^{-1}\sum_{t=2}^{n} 3\sigma_1^2 U_{1,\Delta}(t-1)U_{1,\Delta}(t)}{n^{-1}\sum_{t=1}^{n} 3\sigma_1^2 U_{1,\Delta}(t)^2} \right); \tag{27}$$

cf. Proposition 2. In the hyperbolic case the estimation of the parameters η_2/λ_1 and η_1/λ_1 is more difficult. These parameters do not appear in the formulae for the stationary variances σ_k^2, whence the estimators need to be based on the autoregression coefficients

$$\rho_{k,\Delta} = \eta_2 T_k'(\Delta) + \eta_1 T_k(\Delta) = 1 + \frac{\lambda_k}{2\eta_2}\Delta^2 - \frac{\lambda_k \eta_1}{6\eta_2^2}\Delta^3 + o(\Delta^3).$$

Since these estimators only are reliable for a small time step $\Delta > 0$, we will use the series expansion in order to derive explicit estimators. The quotient η_2/λ_1 thus can be estimated by

$$\frac{-\lambda_1}{\eta_2} \approx \frac{2(1-\rho_{1,\Delta})}{\Delta^2} \overset{\text{prob}}{\approx} \frac{2}{\Delta^2}\left(1 - \frac{(n-1)^{-1}\sum_{t=2}^{n} 3\sigma_1^2 U_{1,\Delta}(t-1)U_{1,\Delta}(t)}{n^{-1}\sum_{t=1}^{n} 3\sigma_1^2 U_{1,\Delta}(t)^2} \right) \tag{28}$$

and the quotient η_1/λ_1 can be estimated by

$$\begin{aligned} \frac{\eta_1}{\lambda_1} &\approx \frac{1 + (\lambda_2/\lambda_1)(\lambda_1/\eta_2)\Delta^2/2 - \rho_{2,\Delta}}{(\lambda_2/\lambda_1)(\lambda_1/\eta_2)^2\Delta^3/6} \\ &\overset{\text{prob}}{\approx} \frac{3}{\Delta}\frac{\eta_2}{\lambda_1} + \frac{6}{\Delta^3}\frac{\lambda_1}{\lambda_2}\left(\frac{\eta_2}{\lambda_1}\right)^2 \left(1 - \frac{(n-1)^{-1}\sum_{t=2}^{n} 3\sigma_1^2 U_{2,\Delta}(t-1)U_{2,\Delta}(t)}{n^{-1}\sum_{t=1}^{n} 3\sigma_1^2 U_{2,\Delta}(t)^2} \right). \end{aligned} \tag{29}$$

Computing these estimators successively, we find estimates for the parameters η_2, η_1, λ_1 and λ_2. If observations at the points $x = \frac{1}{4}$ and $x = \frac{3}{4}$ are available, then the quotient ξ_1/ξ_2 can be estimated by

$$\begin{aligned} \frac{\xi_1}{\xi_2} &= 2\log\left(\frac{e^{-\xi_1/4\xi_2}}{e^{-3\xi_1/4\xi_2}} \right) \\ &\overset{\text{prob}}{\approx} 2\log\left(\frac{\sum_{k=1}^{\infty}\sigma_k^2 \sin^2(\pi k \frac{1}{4})}{\sum_{k=1}^{\infty}\sigma_k^2 \sin^2(\pi k \frac{3}{4})} \, \frac{n^{-1}\sum_{t=1}^{n} V(t\Delta, \frac{3}{4})^2}{n^{-1}\sum_{t=1}^{n} V(t\Delta, \frac{1}{4})^2} \right) \\ &= 2\log\left(\frac{\sum_{t=1}^{n} V(t\Delta, \frac{3}{4})^2}{\sum_{t=1}^{n} V(t\Delta, \frac{1}{4})^2} \right); \end{aligned} \tag{30}$$

cf. equation (23). Otherwise ξ_1/ξ_2 can be estimated by

$$\frac{\xi_1}{\xi_2} \overset{\text{prob}}{\approx} \frac{2b}{a_j} \log\left(\frac{\sum_{k=1}^{\infty} \sin^2(\pi k a_j/b)/\{(k^2-4)\lambda_1 - (k^2-1)\lambda_2\}}{3\eta_1 n^{-1}\sum_{t=1}^{n} V(t\Delta, a_j/b)^2} \right). \tag{31}$$

The estimator (31) probably has best properties for an observation point a_j/b close to $\frac{1}{2}$. Finally, estimates for the parameters ξ_0, ξ_1, ξ_2 can be found solving the equations

$$\xi_0 - \frac{\xi_1^2}{4\xi_2} = \frac{4\lambda_1 - \lambda_2}{3}, \qquad \xi_2 = \frac{\lambda_1 - \lambda_2}{3\pi^2}.$$

These considerations thus give an easily calculated preliminary estimate $\tilde{\theta}$ for the parameter $\theta = (\eta_2, \eta_1, \xi_0, \xi_1, \xi_2)$. The approximation (24) used is, however, poor for small N. In order to adjust the estimate it is asymptotically efficient by Theorem 4 to perform an optimization of the likelihood function given by the state-space model described by (13), (19) and Proposition 2. The likelihood for this state-space model can be calculated using the Kalman filter (see Brockwell and Davis 1991). Since closed-form expressions for the maximum likelihood estimator and the asymptotic Fisher information (21) are unavailable, it is necessary to do a numerical optimization of the likelihood function, for example starting from $\tilde{\theta}$. Estimation under a statistical hypothesis $\Theta_0 \subset \Theta$ is performed similarly, only optimizing over $\theta \in \Theta_0$.

The likelihood ratio test statistic χ_n for a parabolic equation against a hyperbolic equation is given by (22). Theorem 4 gives the asymptotic distribution

$$\Pr_{\eta_2=0}(\chi_n \geqslant x) \approx \tfrac{1}{2}\Pr(\chi_1^2 \geqslant x), \qquad x > 0.$$

The null hypothesis that the equation is parabolic against the alternative that the equation is hyperbolic is thus rejected at significance level α when $\chi_n > q_{1-2\alpha}$, where $q_{1-2\alpha}$ is the $1 - 2\alpha$ quantile of the chi-square distribution with one degree of freedom.

References

Brockwell, P.J. and Davis, R.A. (1991) *Time Series: Theory and Methods*. New York: Springer-Verlag.

Cont, R. (1998) Modeling term structure dynamics: an infinite dimensional approach. RI 402, Centre de Mathématiques Appliquées, École Polytechnique, Palaiseau, France.

Holden, H., Øksendal, B., Ubøe, J. and Zhang, T. (1996) *Stochastic Partial Differential Equations*. Boston: Birkhäuser.

Huebner, M. and Rozovskii, B.L. (1995) On asymptotic properties of maximum likelihood estimators for parabolic stochastic PDEs. *Probab. Theory Related Fields*, **103**, 143–163.

Markussen, B. (2001) Uniform convergence of approximate likelihoods for multivariate Gaussian time series. Preprint No. 8, University of Copenhagen. http://www.stat.ku.dk/research/preprint/.

Piterbarg, L. and Rozovskii, B.L. (1997) On asymptotical problems of parameter estimation in stochastic PDEs: Discrete time sampling. *Math. Methods Statist.*, **6**, 200–223.

Santa-Clara, P. and Sornette, D. (2001) The dynamics of the forward interest rate curve with stochastic string shocks. *Rev. Financial Stud.*, **14**, 149–185.

Self, S.G. and Liang, K.Y. (1987) Asymptotic properties of maximum likelihood estimators and likelihood ratio test under nonstandard conditions. *J. Amer. Statist. Assoc.*, **82**, 605–610.

Walsh, J.B. (1986) An introduction to stochastic partial differential equations. In R. Carmona, H. Kesten and J.B. Walsh (eds), *École d'Été de Probabilités de Saint-Flour XIV–1984*, Lecture Notes in Math. 1180, pp. 265–437. Berlin: Springer-Verlag.

Received December 2001 and revised February 2003