©2022 Ada Academica https://adac.ee

Eur. J. Stat. 2 (2022) 3
doi: 10.28924/ada/stat.2.3

Mixingale Estimation Function for SPDEs with Random Sampling

Jaya P. N. Bishwal

Department of Mathematics and Statistics, University of North Carolina at Charlotte, 376 Fretwell Bldg, 9201 University City Blvd. Charlotte, NC 28223-0001, USA
**Correspondence: J.Bishwal@uncc.edu*

Abstract. We study the mixingale estimation function estimator of the drift parameter in the stochastic partial differential equation when the process is observed at the arrival times of a Poisson process. We use a two stage estimation procedure. We first estimate the intensity of the Poisson process. Then we substitute this estimate in the estimation function to estimate the drift parameter. We obtain the strong consistency and the asymptotic normality of the mixingale estimation function estimator.

1. Introduction

Parameter estimation in stochastic partial differential equations (SPDEs) is a very young area of research in view of its applications in finance, physics, biology and oceanography. Loges (1984) initiated the study of parameter estimation in infinite dimensional stochastic differential equations. When the length of the observation time becomes large, he obtained consistency and asymptotic normality of the maximum likelihood estimator (MLE) of a real valued drift parameter in a Hilbert space valued SDE. Koski and Loges (1986) extended the work of Loges (1984) to minimum contrast estimators. Koski and Loges (1985) applied the work to a stochastic heat flow problem. Mohapl (1997) studied maximum likelihood and least squares estimators for discrete observations of an elliptic SPDE where the dependent structure of the observations is completely different and simple from the parabolic case. Martingale estimation function for discretely observed diffusions was studied in Bibby and Srensen (1995). Bishwal (2007) studied a new estimating function for discretely sampled diffusions by removing the stochastic integral in Girsanov likelihood. Bishwal (2008) studied likelihood asymptotics and Bayesian asymptotics for drift estimation of finite and infinite dimensional stochastic differential equations. Benstein-von Mises theorem and small noise

Received: 10 Oct 2021.
Key words and phrases. stochastic partial differential equations; space-time white noise; cylindrical Brownian motion; diffusion field; Poisson sampling; mixingale estimating function; quasi likelihood estimator; consistency; asymptotic normality.

Bayesian asymptotics for parabolic stochastic partial differential equations was studied in Bishwal (2018). HJM type forward interest rate models viewed as an SPDE along corresponding estimation and testing problem is studied in Bishwal (2017), see also Bishwal (2021). Chong (2020) established a limit theorem for integrated volatility estimation in SPDE by conducting a martingale approximation by truncation and blocking techniques to apply results by Jacod (1997).

Based on continuous observations, usually there can be two asymptotic settings in SPDE: 1) $T \to \infty$ 2) $N \to \infty$ where T is the length of the observations and N is the number of Fourier coefficients in the expansion of the solution to the SPDE.

In a Bayesian approach, using the first setting, Bishwal (1999) proved the Bernstein-von Mises theorem and asymptotic properties of regular Bayes estimator of the drift parameter in a Hilbert space valued SDE when the corresponding ergodic diffusion process is observed continuously over a time interval $[0, T]$. The asymptotics are studied as $T \to \infty$ under the condition of absolute continuity of measures generated by the process. Results are illustrated for the example of an SPDE.

Bishwal (2002) proved the Bernstein-von Mises theorem and spectral asymptotics of Bayes estimators for parabolic SPDEs when the number of Fourier coefficients becomes large. In this case, the measures generated by the process for different parameters are singular.

In this paper we study the asymptotic properties of the quasi maximum likelihood estimator when we have observations of finite-dimensional projections at Poisson arrival time points. The asymptotic setting is the large number of observations at random time points which are the arrivals of a Poisson process and large number of Fourier coefficients.

The rest of the paper is organized as follows : Section 2 contains model, assumptions and preliminaries which include a review of the limit theorems for triangular array of dependent random variables. In Section 3 we prove the main results of the paper. Section 4 demonstrates the results through heat equation as an example of SPDE. Section 5 provides concluding remarks.

2. Model and Preliminaries

Let us fix θ_0, the unknown true value of the parameter θ. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space and $W(t, x)$ be a process on this space with values in the Schwarz space of distributions $D'(G)$ where $x \in G \subset \mathbb{R}^d$ such that for $\phi, \psi \in C_0^\infty(G)$, $\|\phi\|_{L^2(G)}^{-1} \langle W(t, \cdot), \phi(\cdot)\rangle$ is a one dimensional Wiener process and

$$E(\langle W(s, \cdot), \phi(\cdot)\rangle \langle W(t, \cdot), \psi(\cdot)\rangle) = (s \wedge t)(\phi, \psi)_{L^2(G)}. \tag{2.1}$$

This process is usually referred to as the cylindrical Brownian motion (CBM).

Consider the stochastic evolution equation

$$du(t,x) + (\mathcal{A}_0 + \theta\mathcal{A}_1)u(t,x)dt = dW(t,x), \quad t \in [0,T], x \in G, \quad u(0,x) = 0 \tag{2.2}$$

where G is a smooth bounded domain in $\mathbb{R}^d$, $\mathcal{A}_0$ and $\mathcal{A}_1$ are linear operators on a smooth bounded domain G in $\mathbb{R}^d$ with orders m_0 and m_1 respectively with $m_1 \geq m - d/2$ where $2m = \max(m_0, m_1)$, $\{W_t(x)\}$ is a cylindrical Brownian motion based on the observations of the solution $u(t,x)$, $t \in [0,T]$, $x \in G$. Let $A^\theta = \mathcal{A}_0 + \theta\mathcal{A}_1$.

We assume that there exists a complete orthonormal system $\{h_j\}_{j=1}^\infty$ in $L_2(G))$ such that for every $j = 1, 2, \ldots,$ the system $h_j \in W_0^{m,2}(G) \cap C^\infty(\overline{G})$ and

$$\Lambda_\theta h_j = \beta_j(\theta)h_j, \text{ and } \mathcal{L}_\theta h_j = \mu_j(\theta)h_j \text{ for all } \theta \in \Theta \tag{2.3}$$

where $\mathcal{L}_\theta$ is a closed self adjoint extension of A^θ, $\Lambda_\theta := (k(\theta)I - \mathcal{L}_\theta)^{1/2m}, k(\theta)$ is a constant and the spectrum of the operator Λ_θ consists of eigenvalues $\{\beta_j(\theta)\}_{j=1}^\infty$ of finite multiplicities and $\mu_j(\theta) = -\beta_j^{2m} + k(\theta)$.

The cylindrical Brownian motion $W(t,x)$ can be expanded in the series

$$W(t,x) = \sum_{j=1}^\infty W_j(t)h_j \tag{2.4}$$

where $\{W_j(t), t \geq 0\}_{j=1}^\infty$ are independent one dimensional Brownian motions. The latter series converges P-a.s. in $H^{-\alpha}$ for $\alpha > d/2$. Indeed

$$\|W(t,x)\|_{-\alpha}^2 = \sum_{j=1}^\infty W_j^2(t)\|h_j\|_{-\alpha}^2 = \sum_{j=1}^\infty W_j^2(t)\beta_j^{-2\alpha} \tag{2.5}$$

and the later series converges P-a.s.

Let

$$\psi_N := \sum_{j=1}^N \frac{\beta_j^2}{\mu_j} \tag{2.6}$$

Here $\theta \in \Theta \subseteq \mathbb{R}$ is the unknown parameter to be estimated on the basis of the observations of the random field $u^\theta(t,x), t \geq 0, \ x \in [0,1]$.

Consider the Fourier expansion of the process

$$u(t,x) = \sum_{j=1}^\infty u_j(t)\phi_j(x) \tag{2.7}$$

corresponding to some orthogonal basis $\{\phi_j(x)\}_{j=1}^\infty$. Bagchi and Kumar (2001) used this representation for infinite factor model. Note that $u_j^\theta(t), j \geq 1$ are independent one dimensional Ornstein-Uhlenbeck processes

$$du_j^\theta(t) = \mu_j(\theta)u_j^\theta(t)dt + \beta_j^{-\alpha}dW_j(t) \tag{2.8}$$

$$u_j^\theta(0) = u_{0j}^\theta,$$

Recall that $\mu_j(\theta) = k(\theta) - \beta_j^{2m}$. Thus the jth Fourier coefficient satisfies the linear SDE of the Ornstein-Uhlenbeck type

$$du_j^\theta(t) = (k(\theta) - \beta_j^{2m})u_j^\theta(t)dt + \beta_j^{-\alpha}dW_j(t) \tag{2.9}$$

Note that for a fixed t, the processes are $\{u_1(t), u_2(t), u_3(t), \ldots\}$ independent. This is like a continuous version of cross section time series, i.e, a joint regression auto-regression model of order 1. The random field $u(t, x)$ is observed at discrete times t and discrete positions x. Equivalently, the Fourier coefficients $u_j^\theta(t)$ are observed at discrete time points.

The discrete time points could be deterministic (equally spaced/homoscedastic or unequally spaced
/heteroscadastic) or random. We consider random time points. For $x \in [0, 1]$, for fixed j, we observe the process $\{u_j(t), t \geq 0\}$ at times $\{t_0, t_1, t_2, \ldots.\}$. We assume that the sampling instants $\{t_i, i = 0, 1, 2\ldots\}$ are generated by a Poisson process on $[0, \infty)$, i.e., $t_0 = 0$, $t_i = t_{i-1} + \alpha_i$, $i = 1, 2, \ldots$ where α_i are i.i.d. positive random variables with a common exponential distribution $F(x) = 1 - \exp(-\lambda x)$. Note that intensity parameter $\lambda > 0$ is the average sampling rate which is needs to be estimated. It is also assumed that the sampling process t_i, $i = 0, 1, 2, \ldots$ is independent of the observation process $\{u_j(t), t \geq 0, j \geq 1\}$. We note that the probability density function of $t_{k+i} - t_k$ is independent of k and is given by the gamma density

$$f_i(t) = \lambda(\lambda t)^{i-1} \exp(-\lambda t) I_t/(i-1)!, \;\; i = 0, 1, 2, \ldots. \tag{2.10}$$

where $I_t = 1$ if $t \geq 0$ and $I_t = 0$ if $t < 0$.

For a fixed $1 \leq j \leq N$, we denote $u_j(t_i)$ by $u_{j,t_i}, i = 1, 2, \ldots, n$. We observe the first N Fourier coefficient at the random time points $t_i, i = 1, 2, \ldots, n$. Thus the dataset is given by $u_{j,t_i}, j = 1, 2, \ldots, N, i = 1, 2, \ldots, n$. Thus the total number of observations of the random field u on the grid is nN. For instance, in the case of an application to term structure data, this is quite natural, since we can expect many observations over a time of a number of different maturities. Our asymptotic set up is $n \to \infty$ and $N \to \infty$.

Define

$$\rho := \rho(\lambda, \theta) = \frac{\lambda}{\lambda - \kappa(\theta) + \beta_j^{2m}}, j = 1, 2, \ldots, N. \tag{2.11}$$

Now recall some limit theorems for triangular array of dependent random variables, namely martingale difference sequences and mixing sequences.

Theorem 2.1 *Martingale CLT (Hall and Heyde (1980)) : Let $\{X_{n,i}, 1 \leq i \leq k_n\}$ be a triangular array of square integrable martingale difference sequence under the natural filtrations satisfying:*

a) $\max |X_{n,i}| \to^P 0$*;*

b) $\sum_{i=1}^{k_n} X_{n,i}^2 \to^P 1$*;*

c) $E \max X_{n,i}^2$ *is bounded in* n*;*

Then $S_n \to^{\mathcal{D}} \mathcal{N}(0,1)$ *as* $n \to \infty$ *where* $S_n = \sum_{i=1}^{k_n} X_{n,i}$.

Theorem 2.2 *Mixing CLT (Peligrad and Utev (1997)): Let* $\{X_{n,i}, 1 \le i \le k_n\}$ *be a tringular array of random variables satisfying:*

a) $var(\sum_{j=a}^{b} X_{n,j}) \le C \sum_{j=a}^{b} var(X_{n,j})$ *for every* $0 \le a \le b \le k_n$ *where* C *is a universal constant;*

b) $\liminf_{n\to\infty} \dfrac{var(\sum_{j=1}^{k_n} X_{n,j})}{\sum_{j=1}^{k_n} var(X_{n,j})} > 0$;

c) $$\left| cov\left(\exp\left(it\sum_{j=a}^{b} X_{n,j}\right), \quad \exp\left(it\sum_{j=b+u}^{c} X_{n,j}\right)\right)\right| \le h_t(u)\sum_{j=a}^{c} var(X_{n,j})$$

for every $0 \le a \le b \le c \le k_n$ *where* $h_t(u) \ge 0, \sum h_t(2^j) < \infty$ *and* u *is of the form* $u = [(c-a)^{1-\epsilon}]$ *for certain* $0 < \epsilon < 1$;

d) $\sigma_n^{-2} \sum_{j=1}^{k_n} EX_{n,j}^2 I(|X_{n,j}| > \epsilon\sigma_n)$ *as* $n \to \infty$ *for every* $\epsilon > 0$ *where* σ_n^2 *denotes* $var(\sum_{j=1}^{k_n} X_{n,j})$. *Then* $S_n/\sigma_n \to^{\mathcal{D}} \mathcal{N}(0,1)$ *as* $n \to \infty$ *where* $S_n = \sum_{j=1}^{k_n} X_{n,j}$.

Condition (C) in Theorem 2.2 is a ρ-mixing type condition generalizing the more restrictive condition from Utev (1991) that the triangular array is ρ-mixing with a certain decay of the mixing coefficients.

The ρ-mixing condition which gives the rate of decay of maximal correlation coefficient was introduced by Kolmogorov and Rozanov (1960). Note that ρ-mixing implies strong mixing.

Next we recall some facts from mixingale law of large numbers (Mixingale LLN) and mixingale central limit theorem (Mixingale CLT).

Mixingale was introduced by McLeish (1975a). Mixingale condition is a condition of weak temporal dependence that is weaker than most conditions considered in the literature. Examples include martingale difference sequences, integrable M-dependent sequences, stationary Gaussian processes whose correlations converge to zero as the time span increases to infinity, ϕ-mixing sequences, ρ-mixing sequences, α-mixing sequences, autoregressive moving average sequences, infinite order moving average sequences and near epoch dependent sequences.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $\{\mathcal{F}_{n,i}, i = \ldots, 0, 1, \ldots; n = 1, 2, \ldots\}$ be any nondecreasing sequence of sub σ-fields of $\mathcal{F}$ in i for each n. Often one will take $\mathcal{F}_{n,i} = \mathcal{F}$ for $i > k_n$ and $\mathcal{F}_{n,i} = \{\Phi, \Omega\}$ for $i \le 0$. Let $\{X_{n,i}, i = 1, 2, \ldots, k_n, n = 1, 2, \ldots\}$ where $k_n \uparrow \infty$ as $n \to \infty$.

Definition 2.1 (*Mixingale*) A triangular array $\{X_{n,i}, \mathcal{F}_{n,i}\}$ is a L_1-mixingale if for all $i = 1, 2, \ldots, k_n, n = 1, 2, \ldots$ where $k_n \to \infty$ as $n \to \infty$, and $\psi_m : m = 0, 1, 2, \ldots$ such that $\psi_m \downarrow 0$ as

$m \to \infty$, we have

$$(a) \quad |E(X_{n,i}|\mathcal{F}_{n,i-m})| \leq c_{n,i}\psi_m,$$

$$(b) \quad |X_{n,i} - E(X_{n,i}|\mathcal{F}_{n,i+m})| \leq c_{n,i}\psi_{m+1}.$$

L_1-mixingales are necessarily sequences of mean zero random variables since $|EX_{n,i}| = E|E(X_{n,i}|\mathcal{F}_{n,i-m})|$
$\leq c_{n,i}\psi_m \to 0$ as $n \to \infty$. The constants $\{\psi_m\}$ are referred to as L_1-mixingale numbers. These numbers index the temporal dependence of the sequence $\{X_i\}$. If $\{X_i\}$ is independent, then $\psi_m = 0$ for all $m \geq 0$.

First we report weak law of large numbers for mixingales.

Theorem 2.3 *(Mixingale WLLN)(Andrews (1988)) : Suppose the triangular array $\{X_{n,i}, \mathcal{F}_{n,i}\}$ is a uniformly integrable L_1-mixingale. If $\limsup_{n\to\infty} \frac{1}{k_n}\sum_{i=1}^{k_n} c_{n,i} < \infty$, then*

$$E|\bar{X}_n| = E\left|\frac{1}{k_n}\sum_{i=1}^{k_n} X_{n,i}\right| \to 0$$

as $n \to \infty$ and in consequence $\bar{X}_n \to^P 0$ as $n \to \infty$.

Then we report the following strong law of large numbers which strengthens McLeish (1975a) by weakening the conditions.

Theorem 2.4 *(Mixingale SLLN) (Hansen (1991)) : Suppose the array $\{X_i, \mathcal{F}_i\}$ is a L_2-mixingale with $\sum_{m=1}^{\infty} \psi_m < \infty$ and $\sum_{i=1}^{\infty} i^{-2}c_i^2 < \infty$. Then $\bar{X}_n \to 0$ almost surely as $n \to \infty$.*

The following is the strong law of large numbers for triangular array of mixingales.

Theorem 2.5 *(Mixingale SLLN) (De Jong (1996)) : Suppose the triangular array $\{X_{n,i}, \mathcal{F}_{n,i}\}$ is a L_2-mixingale and for a positive integer-valued sequence m_n, we have $\sum_{n=1}^{\infty}(n^{-1}\sum_{i=1}^{n} c_{n,i}\psi_{m_n})^2 < \infty$. Then $\bar{X}_n \to 0$ almost surely as $n \to \infty$.*

McLeish (1975b) gave an invariance principle similar to the central limit theorem of Serfling (1968) under assumption of the conditional expectations of variables with respect to distant last. Tightness is proved under and "asymptotic martingale" type condition, and when conditional variances of the partial sums are asymptotically constant, the limit is shown to be Brownian motion.

Theorem 2.6 *(Mixingale CLT) (McLeish (1977)) : Suppose the triangular array $\{X_{n,i}, \mathcal{F}_{n,i}\}$ is a uniformly integrable L_1-mixingale. Then*

$$\sum_{i=1}^{nt} X_{n,i} \to^{\mathcal{D}} W_t$$

as $n \to \infty$ *where* $h^{-1}E\left[\left(\sum_{i=\lfloor nt \rfloor}^{\lfloor n(t+h) \rfloor} X_{n,i}\right)^2 |\mathcal{F}_{n,\lfloor ns \rfloor}\right] - 1 \to^P 0$ *as* $n^{-1} + h + (nh)^{-1} \to 0$ *for some* s, t *such that* $0 \leq s < t < t+h < 1$ *where* W_t *is standard Brownian motion.*

Theorem 2.6 is generalized in the following:

Theorem 2.7 *(Mixingale CLT) (Ikeda (2013)) : Suppose the triangular array* $\{X_{n,i}, \mathcal{F}_{n,i}\}$ *is a uniformly integrable* L_1*-mixingale. Then*

$$\sum_{i=1}^{nt} X_{n,i} \to^{\mathcal{D}} \int_0^t \delta_v^{1/2} dW_v$$

as $n \to \infty$ *where* $h^{-1}E\left[\left(\sum_{i=\lfloor nt \rfloor}^{\lfloor n(t+h) \rfloor} X_{n,i}\right)^2 |\mathcal{F}_{n,\lfloor ns \rfloor}\right] - \delta_t \to^P 0$ *as* $n^{-1} + h + (nh)^{-1} \to 0$ *for some* s, t *such that* $0 \leq s < t < t+h < 1$ *with* $(\delta_t)_{t\in[0,1]}$ *is non-negative,* t*-continuous, uniformly integrable and* $\mathcal{F}$*-measurable and* $\int_0^1 \delta_s ds$ *is uniformly bounded away from zero.*

Mixingale estimation function (MEF) estimator, which is also the quasi maximum likelihood estimator (QMLE) is the solution of the estimating equation:

$$G^*_{n,N}(\theta) = 0 \tag{2.12}$$

where

$$G^*_{n,N}(\theta) = \sum_{j=1}^{N}\sum_{i=1}^{n} \frac{\beta_j^{2\alpha}\lambda(\rho(\lambda,\theta))^2}{\rho(\lambda, 2\theta)} u_{j,t_{i-1}} \left[(u_{j,t_{i-1}}\theta\rho(\lambda,\theta))^2 + \lambda\right]^{-1} [u_{j,t_i} - \rho(\lambda,\theta)u_{j,t_{i-1}}]. \tag{2.13}$$

We call the solution of the estimating equation the quasi maximum likelihood estimator (QLE). There is no explicit solution for this equation.

The optimal estimating function for estimation of the unknown parameter θ is given by

$$G_{n,N}(\theta) = \sum_{j=1}^{N}\sum_{i=1}^{n} \beta_j^{2\alpha} u_{j,t_{i-1}}[u_{j,t_i} - \rho(\lambda,\theta)u_{j,t_{i-1}}]. \tag{2.14}$$

The mixingale estimation function (MEF) estimator of ρ is the solution of

$$G_{n,N}(\theta) = 0 \tag{2.15}$$

and is given by

$$\hat{\rho}_{n,N} := \frac{\sum_{j=1}^{N}\sum_{i=1}^{n} u_{j,t_{i-1}} u_{j,t_i}}{\sum_{j=1}^{N}\sum_{i=1}^{n} u^2_{j,t_{i-1}}}. \tag{2.16}$$

3. Main Results

We do the parameter estimation in two steps: The rate λ of the Poisson process can be estimated given the arrival times t_i, therefore it is done at a first step. Since we observe total number of

arrivals n of the Poisson process over the T intervals of length one, the MLE of λ is given by

$$\hat{\lambda}_n := \frac{n}{T}. \tag{3.1}$$

Theorem 3.1 *We have*

$$\hat{\lambda}_n \to \lambda \text{ a.s. as } n \to \infty.$$

$$\sqrt{n}(\hat{\lambda}_n - \lambda) \to^{\mathcal{D}} \mathcal{N}(0,\ e^{\lambda}(1 - e^{-\lambda})) \text{ as } n \to \infty.$$

Proof. Let V_i be the number of arrivals in the interval $(i-1, i]$. Then $V_i,\ i = 1, 2, \ldots, n$ are i.i.d. Poisson distributed with parameter λ. Since Φ is continuous, we have for every fixed $j \geq 1$, $I_{\{0\}}(V_i) = I_{\{0\}}(u_{j,t_i})$ a.s. $i = 1, 2, \ldots, n$. Note that for every fixed $j \geq 1$,

$$\frac{1}{n}\sum_{i=1}^{n} I_{\{0\}}(u_{j,t_i}) \to^{a.s.} E(I_{\{0\}}V_1) = P(V_1 = 0) = e^{-\lambda} \text{ as } n \to \infty. \tag{3.2}$$

LLN and CLT and delta method applied to the sequence $I_{\{0\}}(u_{t_i}),\ i = 1, 2, \ldots, n$ give the results. □

The CLT result above allows us to construct confidence interval for the jump rate λ.

Corollary 3.1
A $100(1-\alpha)\%$ confidence interval for λ is given by
$\left[\frac{n}{T} - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{1}{n} - \frac{1}{T}},\ \ \frac{n}{T} + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{1}{n} - \frac{1}{T}}\right]$ *where $Z_{1-\frac{\alpha}{2}}$ is the $(1-\frac{\alpha}{2})$-quantile of the standard normal distribution.*

We obtain the strong consistency and asymptotic normality of the MEF estimator in the following theorem.

Theorem 3.2 *We have*

$$a)\ \ \hat{\rho}_n \to^{a.s.} \rho \text{ as } n \to \infty \text{ and } N \to \infty.$$

$$b)\ \ \sqrt{n\psi_N}(\hat{\rho}_{n,N} - \rho) \to^{\mathcal{D}} \mathcal{N}(0,\ \lambda^{-i}(1 - e^{-\rho})) \text{ as } n \to \infty \text{ and } N \to \infty.$$

Proof. By using the fact that every stationary mixing process is ergodic, it is easy to show that if $u_j(t)$ is a stationary ergodic O-U Markov process and t_i is a process with nonnegative i.i.d. increments which is independent of $u_j(t)$, then $\{u_{j,t_i}, i \geq 1, j \geq 1\}$ is a stationary ergodic Markov process. Hence $\{u_{j,t_i}, i \geq 1\}$ is a stationary ergodic Markov process. Thus the extra randomness of the sampling instants preserves the stationarity and ergodicity of the Markov process in order for the law of large numbers to be applicable.

Observe that $u_j^{\theta}(t) := v_j$ is a stationary ergodic Markov chain and $v_j \sim \mathcal{N}(0, \sigma^2)$ where σ^2 is the variance of u_{1,t_0}. Thus by SLLN for zero mean square integrable mixingales (Theorem 2.5),

Peligrad and Utev (1997, Theorem B) and arguments in Bibinger and Trabs (2017, Proposition 7.6), we have

$$\frac{1}{n\psi_N}\sum_{j=1}^{N}\sum_{i=1}^{n} u_{j,t_{i-1}}u_{j,t_i} \to^{a.s.} E(u_{1,t_0}u_{1,t_1}) = \rho E(u_{1,t_0}^2) \tag{3.3}$$

$$\frac{1}{n\psi_N}\sum_{j=1}^{N}\sum_{i=1}^{n} u_{j,t_{i-1}}^2 \to^{a.s.} E(u_{1,t_0}^2) \tag{3.4}$$

Thus

$$\frac{\sum_{j=1}^{N}\sum_{i=1}^{n} u_{j,t_{i-1}}u_{j,t_i}}{\sum_{j=1}^{N}\sum_{i=1}^{n} u_{j,t_{i-1}}^2} \to^{a.s.} \rho. \tag{3.5}$$

Further,

$$\sqrt{n\psi_N}(\hat{\rho}_{n,N}-\rho) = \frac{(n\psi_N)^{-1/2}\sum_{j=1}^{N}\sum_{i=1}^{n} u_{j,t_{i-1}}(u_{j,t_i}-\theta u_{j,t_{i-1}})}{(n\psi_N)^{-1}\sum_{j=1}^{N}\sum_{i=1}^{n} u_{j,t_{i-1}}^2}. \tag{3.6}$$

Since

$$E(u_{j,t_1}u_{j,t_2}|u_{j,t_1}) = \theta u_{j,t_1}^2, \quad j=1,2,\ldots,N, \tag{3.7}$$

it follows by Theorem 2.7 and Theorem 2.2 which an generalization of Theorem 2.1 obtained Peligrad and Utev (1997, Theorem B), along with the arguments in Bibinger and Trabs (2019), that

$$(n\psi_N)^{-1/2}\sum_{j=1}^{N}\sum_{i=1}^{n} u_{j,t_{i-1}}(u_{j,t_i}-\theta u_{j,t_{i-1}}) \tag{3.8}$$

converges in distribution to normal distribution with mean zero and variance equal to

$$E[(u_{1,t_1}u_{1,t_2}) - E(u_{1,t_1}u_{1,t_2}|u_{j,t_1})]^2 = 1 - e^{2(\theta-\beta_1\delta)}\{2(\beta_1-\theta)(\beta_1+1)\}^{-1}. \tag{3.9}$$

Applying delta method the result follows. □

In the next step, we use the estimator of λ to estimate θ.

Note that

$$\frac{1}{\hat{\rho}_{n,N}} = \frac{\sum_{j=1}^{N}\sum_{i=1}^{n} u_{j,t_{i-1}}^2}{\sum_{j=1}^{N}\sum_{i=1}^{n} u_{j,t_{i-1}}u_{j,t_i}}. \tag{3.10}$$

Thus

$$1 + \frac{\beta_1^{2m}-\kappa(\theta)}{\lambda} = \frac{\sum_{j=1}^{N}\sum_{i=1}^{n} u_{j,t_{i-1}}^2}{\sum_{j=1}^{N}\sum_{i=1}^{n} u_{j,t_{i-1}}u_{j,t_i}}. \tag{3.11}$$

Thus

$$\begin{aligned}\frac{\beta_1^{2m}-\kappa(\theta)}{\lambda} &= \frac{\sum_{j=1}^{N}\sum_{i=1}^{n} u_{j,t_{i-1}}^2}{\sum_{j=1}^{N}\sum_{i=1}^{n} u_{j,t_{i-1}}u_{j,t_i}} - 1 \\ &= -\frac{\sum_{j=1}^{N}\sum_{i=1}^{n} u_{j,t_{i-1}}[u_{j,t_i}-u_{j,t_{i-1}}]}{\sum_{j=1}^{N}\sum_{i=1}^{n} u_{j,t_{i-1}}u_{j,t_i}}\end{aligned} \tag{3.12}$$

Now replace λ by its estimator MLE $\hat{\lambda}_n = \frac{n}{T}$.

$$\beta_1^{2m} - \kappa(\theta) = -\frac{\sum_{j=1}^{N}\sum_{i=1}^{n} u_{j,t_{i-1}}[u_{j,t_i} - u_{j,t_{i-1}}]}{\frac{T}{n}\sum_{j=1}^{N}\sum_{i=1}^{n} u_{j,t_{i-1}} u_{j,t_i}} \tag{3.13}$$

Thus

$$\hat{\theta}_{n,N} = \kappa^{-1}\left(\beta_1^{2m} + \frac{\sum_{j=1}^{N}\sum_{i=1}^{n} u_{j,t_{i-1}}[u_{j,t_i} - u_{j,t_{i-1}}]}{\frac{T}{n}\sum_{j=1}^{N}\sum_{i=1}^{n} u_{j,t_{i-1}} u_{j,t_i}}\right). \tag{3.14}$$

Since the function $\kappa^{-1}(\cdot)$ is a continuous function, by application of delta method, the following result is a consequence of Theorem 3.2.

Theorem 3.3

$$a) \quad \hat{\theta}_{n,N} \to^{a.s.} \theta \text{ as } n \to \infty \text{ and } N \to \infty$$

$$b) \ \sqrt{n\psi_N}(\hat{\theta}_{n,N} - \theta) \to^{\mathcal{D}} \mathcal{N}(0,\ (\kappa'(\theta))^{-2}\lambda^2(1 - e^{-2\lambda^{-1}(\kappa(\theta)-\beta_1^{2m})}))$$

$$\text{as } n \to \infty \text{ and } N \to \infty \text{ such that } \frac{N}{\sqrt{n}} \to 0.$$

In the second stage, we substitute λ by its estimator $\hat{\lambda}_n = \frac{n}{T}$.

4. Examples

1) Consider the stochastic heat equation

$$du^\theta(t,x) = \theta\frac{\partial^2}{\partial x^2}u^\theta(t,x)dt + dW(t,x) \tag{4.1}$$

for $0 \le t \le T$ and $x \in (0,1)$ and $\theta > 0$ with periodic boundary conditions.

Here $2m = m_1 = 2$ and $\mu_j = -\theta\pi^2 j^2,\ \gamma > 1/2$ and $\psi_N = N^3$. Hence

$$\sqrt{nN^3}(\hat{\theta}_{n,N} - \theta) \to^{\mathcal{D}} \mathcal{N}(0,\ (\kappa'(\theta))^{-2}\lambda^2(1 - e^{-2\lambda^{-1}(\kappa(\theta)-\beta_i^{2m})})) \text{ as } n \to \infty \text{ and } N \to \infty.$$

2) As another example of the evolution equation consider the linear parabolic equation

$$du^\theta(t,x) = \theta u^\theta(t,x) + \frac{\partial^2}{\partial x^2}u^\theta(t,x)dt + dW(t,x),\ t \ge 0,\ x \in [0,1] \tag{4.2}$$

$$u(0,x) = u_0(x) \in L_2([0,1]) \tag{4.3}$$

$$u^\theta(t,0) = u^\theta(t,1),\ t \in [0,T], \tag{4.4}$$

If $d = 2$, then we have

$$\sqrt{n\log N}(\hat{\theta}_{n,N} - \theta) \to^{\mathcal{D}} \mathcal{N}(0,\ (\kappa'(\theta))^{-2}\lambda^2(1 - e^{-2\lambda^{-1}(\kappa(\theta)-\beta_i^{2m})})) \text{ as } n \to \infty \text{ and } N \to \infty.$$

If $d > 2$, then we have

$$\sqrt{nN^{(d-2)/d}}(\hat{\theta}_{n,N} - \theta) \to^{\mathcal{D}} \mathcal{N}(0,\ (\kappa'(\theta))^{-2}\lambda^2(1 - e^{-2\lambda^{-1}(\kappa(\theta)-\beta_i^{2m})})) \text{ as } n \to \infty \text{ and } N \to \infty.$$

5. Concluding Remarks

1. Observe that the parametric rate is $\sqrt{n\psi_N}$. This is expected in an ergodic situation.

2. For fixed time horizon T, with deterministic time grid $t_i = i\Delta_n, i = 1, 2, \ldots, n$, one needs the double asymptotics as follows: Either one can consider N to be fixed, e.g, $N = 1$ or N is a sequence satisfying $N = O(n^\gamma)$ for some $\gamma \in (0, 1/2)$. See Bibinger and Trabs (2017) for estimation of volatility in SPDE. However, for random time grid, with the arrivals being that of a Poisson process, one needs $N \to \infty$ to ensure that in finite time $T < \infty$, there may not have enough arrivals.

3. In deterministic time grid, Bibinger and Trabs (2017) provide asymptotic results with grid finer in time than in space. Thus the observation scheme is *high frequency* in time and *low resolution* in space. This is satisfied by the following scheme, so called rapidly increasing experimental design (REID): $\frac{N}{\sqrt{n}} \to 0$ as $n \to \infty$ and $N \to \infty$.

4. If the parameter is multidimensional, it would be interesting to see how the dimension of the parameter space, e.g., r relates to the space-time data sizes N and n. Also if the dimension of the parameter space increases with sample sizes, it would be interesting to see the triple asymptotics framework. Another problem would be sieve estimation of the nonparametric drift when the dimension of the sieve, e.g., q increases along with space-time data sizes N and n.

5. In standard semimartingale models leading to discrete increments, which are (almost) uncorrelated (martingale increments are uncorrelated), in our case the increments are negatively correlated. Hence, martingale central limit theorem can not be applied due to negative auto-correlation of the increments. Jacod's stable central limit theorem, see Jacod (1997, Theorem 3-1), which is typically exploited in the literature of high frequency statistics can not be directly applied. This fact highlights crucial difference between the discretized SPDE model and classical theory from the statistics of semimartinagle literature. Utev (1990) has proved a central limit theorems for ρ-mixing triangular arrays can be applied. However the random field here is not ρ-mixing. The abstract ρ-mixing assumption can be replaced by two explicit conditions on the variance of partial sums and on covariances of characteristic functions of partial sums. The resulting generalized central limit theorem, which satisfies a covariance inequality related to ρ-mixing process, has been reported by Peligrad and Utev (1997, Theorem B), which is presented as Theorem 2.2 and we verify the generalized mixing type conditions in our setup.

6. One advantage of random sampling is to remove the alising problem. Another advantage of random sampling is to gain optimality of discretization, see e.g, see Hofmann *et al.* (2001).

7. It remains to investigate estimation from random space sampling, e.g, Beta sampling of SPDE which have not been considered in this paper. Simultaneous random sampling of both time

and space would be more interesting.

References

[1] D.W.K. Andrews, Laws of large numbers for dependent non-identically distributed random variables, Econometric Theory, 4 (3) (1988), 458-467.

[2] A. Bagchi, K.S. Kumar, An infinite factor model for the interest rate derivatives, in: M. Kohlmann, S. Tang (Eds.), Mathematical Finance, Birkhäuser Basel, Basel, 2001: pp. 59–68.

[3] B.M. Bibby, M. Srensen, Martingale estimation functions for discretely observed diffusion processes, Bernoulli 1 (1995), 17-39.

[4] M. Bibinger, M. Trabs, Volatility estimation for stochastic PDEs using high-frequency observations, ArXiv. (2017).

[5] M. Bibinger, M. Trabs, On central limit theorems for power variations of the solution to the stochastic heat equation, in: A. Steland, E. Rafajłowicz, O. Okhrin (Eds.), Stochastic Models, Statistics and Their Applications, Springer International Publishing, Cham, 2019: pp. 69–84.

[6] J.P.N. Bishwal, Bayes and sequential estimation in Hilbert space valued stochastic differential equations, J. Korean Stat. Soc. 28 (1999), 93-106.

[7] J.P.N. Bishwal, Rates of convergence of the posterior distributions and the Bayes estimators in the Ornstein-Uhlenbeck process, Random Oper. Stoch. Equations 8 (2000), 51-70.

[8] J.P.N. Bishwal, The Bernstein-von Mises theorem and spectral asymptotics of Bayes estimators for parabolic SPDEs, J. Australian Math. Soc. 72 (2001), 289-300.

[9] J.P.N. Bishwal, A new estimating function for discretely sampled diffusions, Random Oper. Stoch. Equations 15 (2007), 65-88.

[10] J.P.N. Bishwal, Parameter Estimation in Stochastic Differential Equations, Lecture Notes in Mathematics, 1923, Springer-Verlag. (2008).

[11] J.P.N. Bishwal, Hypothesis testing for fractional stochastic partial differential equations with applications to neurophysiology and finance, Asian Res. J. Math. 4 (2017), 1-24.

[12] J.P.N. Bishwal, Benstein-von Mises theorem and small noise Bayesian asymptotics for parabolic stochastic partial differential equations, Theory Stoch. Processes 23 (2018), 6-17.

[13] J.P.N. Bishwal, Parameter Estimation in Stochastic Volatility Models, Springer Finance Series, Springer Nature Switzerland AG, (2021), (Forthcoming).

[14] Z. Cheng, I. Cialenko, R. Gong, Bayesian estimations for diagonalizable bilinear SPDEs, Stoch. Processes Appl. 130 (2020), 845-877.

[15] I. Cialenko, F. Delgado-Vences, H-Y. Kim, Drift estimation for discretely sampled SPDEs. ArXiv. (2019).

[16] I. Cialenko, Y. Huang, A note on parameter estimation for discretely sampled SPDEs, ArXiv. (2019).

[17] C. Chong, High-frequency analysis of parabolic stochastic PDEs, Ann. Stat. 48 (2020), 1143-1167.

[18] R.M. De Jong, A strong law of large numbers for triangular mixingale arrays, Stat. Probab. Lett. 27 (1996), 1-9.

[19] P. Hall, C. Heyde, Martingale Limit Theory and its Applications, Academic Press, New York. (1980).

[20] B.E. Hansen, Strong laws for dependent heterogeneous processes, Econometric Theory, 7 (1991), 213-221.

[21] N. Hofmann, T. Muller-Gronbach, K. Ritter, The optimal discretization of stochastic differential equations, J. Complex. 17 (2001), 117-153.

[22] S.S. Ikeda, A note on the mixingale limit theorem by McLeish (1977), GRIPS Discussion Paper (2013), 13-11.

[23] J. Jacod, On continuous conditional Gaussian martingales and stable convergence in law, Seminaire de Probabilites, Lecture Notes in Mathematics, Springer, (1997), 232-246.

[24] A.N. Kolmogorov, Yu. A. Rozanov, On strong mixing conditions for stationary Gaussian processes, Theory Probab. Appl. 5 (1960), 204-207.

[25] T. Koski, W. Loges, Asymptotic statistical inference for a stochastic heat flow problem, Stat. Probab. Lett. 3 (1985), 185-189.

[26] T. Koski, W. Loges, On minimum contrast estimation for Hilbert space valued stochastic differential equations, Stochastics 16 (1986), 217-225.

[27] D.L. McLeish, A maximal inequality and dependent strong law, Ann. Probab. 3 (1975a), 829-839.

[28] D.L. McLeish, Invariance principles for dependent sequences, Z. Wahrscheinlichkeitheorie verw. Gebiete 32 (1975b), 165-178.

[29] D.L. McLeish, On the invariance principle for nonstationary mixingales, Ann. Probab. 5 (1977) , 616-621.

[30] J. Mohapl, On estimation in the planar Ornstein-Uhlenbeck process, Commun. Stst.-Stoch. Models 13 (1997), 435-455.

[31] M. Peligrad, and Utev, Central limit theorem for linear processes, Ann. Probab. 25 (1997), 443-456.

[32] R.J. Serfling, Contributions to central limit theory for dependent variables, Ann. Math. Stat., 39 (1968), 1158-1175.

[33] S.A. Utev, Central limit theorem for dependent random variables, In Probability Theory and Mathematical Statistics, Vol. II (Vilinius, 1989), pages 519-528, "Mokslas", Vilnius. (1990).

[34] S.A. Utev, Sums of random variables with ϕ-mixing, Siberian Adv. Math. 1 (1990), 124-155.