Dynamic Systems and Applications 32 (2023) 165-188

APPROXIMATE MAXIMUM LIKELIHOOD ESTIMATION IN SEMILINEAR SPDE

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
E-mail: J.Bishwal@uncc.edu

ABSTRACT. We estimate the drift in the semilinear SPDE by approximation by space and time discretization. We study the asymptotic properties of the approximate maximum likelihood estimators and also the rates of convergence. We also study parameter estimation in controlled semilinear SPDE.

AMS (MOS) Subject Classification. 35R60, 37H15, 37L55, 60G44, 60H10, 60H15, 60H35, 60J60, 62M05, 62M99, 62F12, 65L09, 93E20.

Key Words and Phrases. Semilinear SPDE, time and space discretization, finite-element method, Galerkin approximation, Crank-Nicholson scheme, Euler scheme, rate of convergence, Markov control, admissibility, approximate maximum likelihood estimation, consistency, asymptotic normality

1. Introduction and Preliminaries

Parameter estimation in semilinear stochastic partial differential equations (SPDEs) has received some recent attention. Altermeyer, Bretschneider, Janak and Reiss (2022) studied parameter estimation in an SPDE model for cell repolarization. Altermeyer, Cialenko and Pasemann (2023) studied parameter estimation for semilinear SPDEs from local measurements. Cialenko, Kim and Pasemann (2023) studied statistical analysis of discretely sampled semilinear SPDEs through a power variation approach. Altermeyer, Bretschneider, Janak, and Reiss (2021) studied nonparametric estimation for linear SPDEs from local measurements. Pasemann and Stannnat (2020) studied Drift estimation for stochastic reaction-diffusion systems. Pasemann, Flemming, Alonso, Beta and Stannat (2021) studied diffusivity estimation for activator-inhibitor models with application to the intracellular dynamics of the actin cytoskeleton.

Received November 11, 2022 ISSN 1056-2176(Print); ISSN 2693-5295 (online)
www.dynamicpublishers.org https://doi.org/10.46719/dsa2023.32.10

Parameter estimation in infinite dimensional stochastic differential equations was first studied by Loges (1984). 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. Bayes and sequential estimation in Hilbert space valued stochastic differential equations were studied in Bishwal (1999). Bishwal (2008) studied asymptotics of maximum likelihood and Bayes estimators in finite and infinite dimensional SDEs.

In the spectral approach with large number of Fourier modes, the Bernstein-von Mises theorem and spectral asymptotics of Bayes estimators for parabolic SPDEs were studied in Bishwal (2001). Bernstein-von Mises theorem and Bayesian asymptotics for small noise intensity for parabolic stochastic partial differential equations were studied in Bishwal (2018). Hypothesis testing for fractional stochastic partial differential equations with applications to neurophysiology and finance was studied in Bishwal (2017). Bishwal (2021) studied parameter estimation and hypothesis testing in non-linear SPDE based on continuous and discrete observations. Bishwal (2022a) studied estimation in SPDE by mixingale estimation function method based on observations at the jump times of a Poisson process.

In this paper we consider discretization of the spectral approach and also temporal discretization. Within spectral approach, there has not been much attempt so far to quantify the amount of spatial information needed to recover its asymptotics for drift estimation. We determine how much spatial information is needed in order to reconstruct spectral asymptotics. Cialenko *et al.* (2020) consider the discretization in time of the maximum likelihood estimator from the spectral approach.

Pasemann and Stannnat (2020) studied asymptotic normality of the AMLE of the drift parameter in a semilinear SPDE, the reaction-diffusion equation, based on finite dimensional approximation to the solution trajectory obtained by truncation in the Fourier space. Burgers equation and Cahn-Hillard equations are treated as special cases. However, our estimators are based on spatial-discretization. We also obtain rates of convergence of the estimators. In the later part of the paper we consider controlled SPDE where the unknown parameter appears in the nonlinear part.

Let H be a separable Hilbert space. Let A be an infinitesimal generator of an analytical semigroup of negative type. Let W_t be Wiener process taking values in H with covariance operator Q. Consider the stochastic evolution equation

$$du_t = (\theta A u_t + f(t, u_t))dt + \sigma(u_t)dW_t, \quad u_0 = u^0 \in U \tag{1.1}$$

where U is a certain interpolation space of H and A.

The functions f and σ satisfy certain smoothness conditions for the existence and uniqueness of solution of the evolution equation.

Fix $\gamma, \alpha, \rho, \alpha_\sigma$ such that $\gamma > 0,\ 0 \le \tau \le 1/2,\ 0 \le \rho < \min(1/2, \gamma)$ and $\rho + \tau_\sigma < \rho$. We assume

(A1) (i) f is Lipschitz continuous:

$$\|f(t,x) - f(t,y)\|_\delta \le C\|x - y\|_\delta$$

for $\delta \in [\rho, \gamma]$ and $x \in D((-A)^\gamma))$.

(ii) f satisfies linear growth

$$\|f(t,x)\|_\delta^2 \le K(1 + \|x\|_\delta^2).$$

for $\delta \in [\rho, \gamma]$ and $x, y \in D((-A)^\gamma))$.

(iii) f is Hölder continuous in time

$$\|f(t,x) - f(s,x)\|_\rho \le C|t-s|^{\min(1/2, \gamma-\rho-\alpha)}\|x\|_{\min(1+\rho+\alpha+\alpha_\sigma,\ \gamma+\alpha)}$$

for $x, y \in D((-A)^\gamma))$.

(A2) Assume that σ is an operator such that $(-A)^{-\alpha}\sigma : X \to L_2^0$ is bounded and satisfies the following conditions:

(i) $(-A)^{-\alpha_\sigma}\sigma$ is Lipschitz continuous in space, i.e.,

$$\|(-A)^{-\alpha_\sigma}[\sigma(x) - \sigma(y)]\|_{L_2^{0,\delta}} \le C\|x - y\|_\delta$$

for $\delta \in [\rho, \gamma]$ and $x, y \in D((-A)^\gamma))$.

(ii)

$$\|(-A)^{-\tau}\sigma(x)\|_{L_2^{0,\delta}} \le C\|x\|_\delta$$

for $\delta \in [\rho, \gamma]$.

(iii) $(-A)^{-\alpha_\sigma}$ is globally Lipschitz, i.e., satisfies

$$\|(-A)^{-\alpha}[\sigma(x) - \sigma(y)]\|_{L_2^{0,\delta}} \le \xi\|x - y\|_\delta$$

for $\delta \in [\rho, \gamma]$ and $x, y \in D((-A)^\gamma))$.

(A3) The semigroup $T_A(t)$ associated with the operator A satisfies

$$\int_0^t \|T_A(t)(x-y)\|^2 ds \le (\zeta + \phi(t))\|x - y\|_\delta^2$$

for all $x, y \in D((-A)^\gamma))$. for all $\delta \in [0, \gamma]$, such that $\phi(t) \to 0$ as $t \to 0$.

Space Discretization: Method of Moments

A function f in H can be written as $f = \sum_{1=1}^\infty f_i\phi_i$ where ϕ_i is a complete set of basis function in H. The approximation is done by taking a finite number of basis function. Define the d_n dimensional subspace $H_n = \text{span}\{\phi_i : 1 \le i \le d_n\}$ and the approximation

$$f \approx \sum_{i=1}^{d_n} \phi_i f_i. \tag{1.2}$$

Substituting the approximation f into the operator equation we get

$$\sum_{i=1}^{d_n} f_i A\phi_i = g.$$

Now taking the inner product of g with a set of test functions $\{\chi_j, 1 \le j \le d_n\}$, we can write

$$\sum_{i=1}^{d_n} f_i \langle A\phi_i, \chi_j\rangle = \langle g, \chi_j\rangle, \quad 1 \le j \le d_n$$

which can be written in the matrix form $[a_{ij}][f_i] = [g_j]$ where

$$a_{i,j} = \langle A\phi_i, \chi_j\rangle$$

and $g_j = \langle \chi_j, g\rangle, \quad 1 \le i \le d_n 1 \le j \le d_n$.

The approximating operator A_n is now defined by the matrix $A_n = (a_{i,j})_{i,j=1}^{d_n}$, i.e. $A_n := P_n A E_n$ where the projection operator is defined by

$$(P_n f)_i = \langle f, \chi_i\rangle, \ i = 1, \dots, d_n$$

where the 'embedding' or 'interpolation' operator E_n is given by

$$E_n c = \sum_{i=1}^{d_n} c_i \phi_i, \quad c \in \mathbb{R}^{d_n}.$$

If $\chi_j = \phi_j$, then the method is known as Galerkin method. P_n coincided with the orthogonal projection operator on the subspace $\{\phi_i : 1 \le i \le d_n\}$. In spectral methods, one takes usually orthogonal rectangular function, e.g., eigenfunction. In this case the projection operator P_n coincides with the orthogonal projection operator on the subspace $\{\phi_i : 1 \le i \le d_n\}$. In finite elements, one uses the variation form of $\sum_{i=1}^{d_n} f_i A\phi_i = g$ to reduce the regularity assumptions on the basis. Here it is sufficient as regularity condition that $\phi_i \in D((-A)^{1/2}), \ 1 \le i \le d_n$. A typical example is finite difference.

The approximation will satisfy

$$du_t^n + \theta A_n u_t^n dt = P_n f(t, E_n u_t^n)dt + \sigma_n(E_n u_t^n) dP_n W_t, \quad u_0^n = P_n u^0 \tag{1.3}$$

where σ_n is a bounded operator on H_n approximating σ such that trace $(\sigma_n Q_n \sigma_n^T)$ is exact on $E_n P_n H$, that is,

$$\text{trace}\langle(\sigma Q \sigma^T)\phi_j, \xi_i\rangle = \text{trace}\langle(\sigma_n Q_n \sigma_n^T)\phi_j, \xi_i\rangle, \quad i, j = 1, 2, \dots, d_n$$

where $Q_n = P_n Q E_n$. Notice that $P_n W_t$ is a d_n-dimensional Wiener process with nuclear covariance matrix Q_n.

The space discretization satisfy the following assumptions:

(B1) (i) $H, H_1, H_2, \ldots$ are all real or complex Banach spaces. All norms will be denoted by $\|\cdot\|$.

(ii) P_n is a bounded linear operator satisfying $\|P_n x\| \leq p\|x\|$ for all $n \geq 1$, $x \in H$ and for some $p \geq 0$.

(iii) E_n is a bounded linear operator satisfying $\|E_n x\| \leq q\|x\|$ for all $n \geq 1$, $x \in H$ and for some $q \geq 0$.

(iv) $P_n E_n x = x$ for all $n \geq 1$ and $x \in H_n$.

(B2) Stability condition: A_n is a bounded operator and there exists some $M < \infty$ and $\omega \in \mathbb{R}$ such that

$$\|e^{A_n} t\| \leq M e^{\omega t} \text{ for } t \geq 0,\ n \geq 1.$$

Time Discretization

We discretize the time t at the same time as the space H. Let τ_n be the time step size corresponding to the space H_n. Following are the time discretization schemes.

The explicit Euler scheme is given by

$$\frac{u_n(t+\tau_n) - u_n(t)}{\tau_n} + \theta A_n u_n(t) = P_n f(t, E_n u_n(t)) + \sigma_n(E_n u_n(t))\Delta B_n(t)$$

where $\Delta B_n(t) := B_n(t+\tau_n) - B_n(t)$ and $B_n(t)$ is a d_n-dimensional Brownian motion with nuclear covariance $Q_n = P_n Q E_n$.

We discuss the simulation of d_n-dimensional Brownian motion B. To simulate $B_n \sim BM(\mu, \Sigma)$, we first find a matrix S such that $SS^T = \Sigma$. If S is $d \times k$, let $Z_1, Z_2, \ldots$ be independent standdard normal in $\mathbb{R}^k$. Set $B_0 = 0$. and $B_{t_{i+1}} = B_{t_i} + \mu(t_{i+1} - t_i) + \sqrt{t_{i+1} - t_i} S Z_i, i = 1, 2, \ldots, n-1$. Thus simulation of $BM(\mu, \Sigma)$ is straightforward once sigma has been factored.

Let $v_k^n := u_{k\tau_n}^n$. Let ξ_k^n are d_n-dimensional standard Gaussian random variables distributed according to $N(0, Q_n)$ where $Q_n = P_n Q E_n$ is the nuclear covariance operator.

Explicit Euler Scheme (Forward Rectangular Rule):

$$v_{k+1}^n = (1 + \tau_n \theta A_n) v_k^n + \tau_n P_n f(k\tau_n, E_n v_k^n) + \sqrt{\tau_n}\sigma_n(v_k^n)\xi_k^n, \quad v_0^n = P_n u^0.$$

Implicit Euler Scheme (Backward Rectangular Rule):

$$v_{k+1}^n = (1 - \tau_n \theta A_n)^{-1} v_k^n + \tau_n P_n f(k\tau_n, E_n v_k^n) + \sqrt{\tau_n}\sigma_n(v_k^n)\xi_k^n, \quad v_0^n = P_n u^0.$$

Crank-Nicholson Scheme (Trapezoidal Rule):

$$v_{k+1}^n = (1 - \frac{\tau_n}{2}\theta A_n)^{-1}(1 + \frac{\tau_n}{2}\theta A_n) v_k^n + \tau_n P_n f(k\tau_n, E_n v_k^n) + \sqrt{\tau_n}\sigma_n(v_k^n)\xi_k^n, \quad v_0^n = P_n u^0.$$

For $t = k\tau_n$, the solution is given by $v_n(t) = v_n^k$. Between the points $k\tau_n$ and $(k+1)\tau_n$, the solution can be linearly interpolated, that is by the polygonal approximation $v_n(t) = v_n^k + \tau_n^{-1}(t - k\tau_n)(v_n^{k+1} - v_n^k)$.

Homogeneous solution at a grid point $k\tau_n$ is approximated by $F_\tau(A_n)^k$ where $F_\tau(A)$ equals $(I + \tau A)$ in case of Explicit Euler scheme, equals $(I + \tau A)^{-1}$ incase of Implicit Euler scheme, equals $(I - \tau A)^{-1}(I + \frac{1}{2}\tau A)$ incase of Crank-Nicholson scheme. In the Crank-Nicholson scheme, the discretization is done symmetrically around the point $k\tau_n + \frac{1}{2}\tau_n$.

(B3) $\sigma : H \to H$ is unbounded and the stability condition

$$\|(1 + \frac{\tau_n}{2} A_n)^k\| \le M \exp(k\tau_n)$$

is satisfied in the case of Crank-Nicholson scheme. If σ is bounded (B3) is not necessary. However, we restrict ourselves to the unbounded case.

In the case of implicit Euler scheme

$$\sqrt{E[\|v_k^n - u(k\tau_n)\|_\rho]^2} \le C\left(\tau_n^{\min(1/2,\gamma-\rho-\tau_\sigma)} + \tau_n \parallel (-A_n)^{\max(0,1+\rho+\tau_\sigma-\gamma)} \parallel \right.$$

$$\left. + \kappa(n)[\eta_{\gamma-\epsilon}(n) + (k\tau_n)^{-\rho} + 1]\eta_\gamma(n)\right).$$

In the case of explicit Euler scheme

$$\sqrt{E[\|v_k^n - u(k\tau_n)\|_\rho]^2} \le C\left(\tau_n^{\min(1/2,\gamma-\rho-\alpha_\sigma)} + \kappa(n)[\eta_{\gamma-\epsilon}(n) + (k\tau_n)^{-\rho} + 1]\eta_\gamma(n)\right).$$

In the case of Crank-Nicholson scheme

$$\sqrt{E[\|v_k^n - u(k\tau_n)\|_\rho]^2} \le C\left(\tau_n^{\min(1/2,\gamma-\rho-\tau_\sigma)} + \kappa(n)[\eta_{\gamma-\epsilon}(n) + (k\tau_n)^{-\rho} + 1]\eta_\gamma(n)\right).$$

The rate is same as the implicit Euler scheme. However under (B3), we have second order convergence of the Crank-Nicholson scheme. See Hausenblas (2003).

2. Space Discretized AMLE

Let P_θ^T the measure generated by the solution $\{u(t), t \in [0,T]\}$ of the SPDE on the space $C([0,T]; H)$ with the associated Borel σ-algebra B_T.

Note that condition (A1) is equivalent to

$$\int_0^T \|Au(s)\|^2 ds < \infty \text{ a.s. for fixed } \epsilon.$$

Thus under (A1), for different θ the measures P_θ are mutually absolutely continuous.

The Radon-Nikodym derivative (likelihood) of P_θ^T with respect to $P_{\theta_0}^T$ is given by

$$\begin{aligned} L_T^\theta(u) := \tfrac{dP_\theta^T}{dP_{\theta_0}^T}(u) &= \exp\left\{(\theta-\theta_0)\int_0^T (Au(s), du(s)) - \frac{1}{2}(\theta^2-\theta_0^2)\int_0^T \|Au(s)\|^2 ds \right. \\ &\left. -(\theta-\theta_0)\int_0^T (Au(s), f(s,u(s))ds\right\}. \end{aligned} \tag{2.1}$$

Maximizing $L_T^\theta(u)$ with respect to θ provides the maximum likelihood estimator (MLE) given by

$$\hat{\theta}_T = \frac{\int_0^T (Au(s), du(s) - f(s,u(s))ds)}{\int_0^T \|Au(s)\|^2 ds}. \tag{2.2}$$

Let $u^N(t)$ be the N-dimensional approximation to the solution trajectory obtained by truncation in Fourier space. The process $u^N(t)$ generates a probability measure on the space of continuous paths with values in $\mathbb{R}^N$, denoted by $P_\theta^{T,N}$. Of course, different values of θ lead to different measures on path space. Due to Girsanov theorem, the density of $P_\theta^{T,N}$ with respect to $P_{\theta_0}^{T,N}$ is given by

$$\begin{aligned} L_{T,N}^\theta(u^N) := \tfrac{dP_\theta^{T,N}}{dP_{\theta_0}^{T,N}}(u^N) &= \exp\left\{(\theta-\theta_0)\int_0^T (Au^N(s), du^N(s)) - \right. \\ \tfrac{1}{2}(\theta^2-\theta_0^2)\int_0^T \|Au^N(s)\|^2 ds & \\ &\left. -(\theta-\theta_0)\int_0^T (Au^N(s), f(s,u^N(s))ds\right\}. \end{aligned} \tag{2.3}$$

Maximizing $L_{T,N}^\theta(u^N)$ with respect to θ provides the approximate maximum likelihood estimator (AMLE) given by

$$\hat{\theta}_{N,T} = \frac{\int_0^T (Au^N(s), du^N(s) - f(s,u^N(s))ds)}{\int_0^T \|Au^N(s)\|^2 ds}. \tag{2.4}$$

Paseman and Stannat (2021) showed that the estimator is consistent and asymptotically nomal as the number of Fourier coefficients becomes large.

Now we proceed through the Galerkin method. We define the AMLE as the conditional least squares estimator (CLSE), due to Gaussian noise they are equal.

We need the following result on AML estimation in the sequel.

Lemma 2.1 *(Le Breton (1976, page 138)). Let $(\Omega, A, \{P_\theta; \theta \in \mathbb{R}\})$ be a statistical structure dominated by P, with a log-likelihood function $L(\theta, \cdot)$. Let $\{A_n, n \geq 1\}$ be a sequence of sub-σ-algebras of A and, for all $n \geq 1, L_n(\theta, \cdot)$ be the log-likelihood function on the statistical structure $(\Omega, A_n, \{P_{\theta|A_n}; \theta \in \mathbb{R}\})$ or any A_n - measurable function. Let us suppose that the following assumptions are satisfied.*

(C1) L and L_n are twice continuously differentiable with derivatives $L^{(i)}$ and $L_n^{(i)}$ respectively, $i = 1, 2$.

(C2) $L^{(2)}$ *does not depend on* θ *and* P*-almost surely strictly negative.*

(C3) $L^{(1)}(\theta, \cdot) = 0$ *admits* P *almost surely a unique solution* $\hat{\theta}$*.*

(C4) There exists a sequence $\{\gamma_n, n \geq 1\}$ *of positive numbers converging to zero and for* $i = 1, 2$ *and all* $K > 0$ *there exists a sequence* $\{\nabla_n^i(K), n \geq 1\}$ *of positive random variables such that* $\theta \in \mathbb{R}$,

(a) $\{\nabla_n^i(K), n \geq 1\}$ *is bounded in* P_θ *probability,*

(b) $\sup_{|\theta| \leq K} |L_n^{(i)}(\theta, \cdot) - L^{(i)}(\theta, \cdot)| \leq \gamma_n \nabla_n^i(K) P_\theta$ *almost surely.*

Then there exists a sequence $\{\theta_n, n \geq 1\}$ *of random variables satisfying*

(i) θ_n *is* A_n*-measurable,*

and for all $\theta \in \mathbb{R}$

(ii) $\lim_{n \to \infty} P_\theta[L_n^{(1)}(\theta_n) = 0] = 1$

(iii) $P_\theta - \lim_{n \to \infty} \theta_n = \hat{\theta}$*, where* $\hat{\theta}$ *is the MLE based on* L*.*

Furthermore, if $\{\theta_n', n \geq 1\}$ *is another sequence of random variables satisfying (i), (ii) and (iii), then for all* $\theta \in \mathbb{R}$

$$\lim_{n \to \infty} P_\theta[\theta_n = \theta_n'] = 1.$$

Lastly, if $\{\theta_n, n \geq 1\}$ *is a sequence satisfying (i), (ii) and (iii) then for all* $\theta \in \mathbb{R}$*, the sequence* $\{\gamma_n^{-1}(\theta_n - \hat{\theta}), n \geq 1\}$ *is bounded in* P_θ *probability.* □

Forward AMLE (Based on the Explicit Euler scheme):

Define

$$F_{n,T} = \sum_{k=1}^{n} \left[\frac{v_{k+1}^n - (1 + \tau_n \theta A_n) v_k^n - \tau_n P_n f(k\tau_n, v_k^n)}{\sqrt{\tau_n} \sigma_n(v_k^n)} \right]^2 \tag{2.5}$$

and

$$\widehat{\theta}_{n,T} = \text{argmin}_\theta F_{n,T}$$

which is given by

$$\widehat{\theta}_{n,T} = \frac{\sum_{k=1}^{n} A_n v_k^n (v_{k+1}^n - v_k^n) - \tau_n P_n f(k\tau_n, v_k^n)}{\sum_{k=1}^{n} (A_n v_k^n)^2 \tau_n}.$$

The following rate of convergence holds for the forward estimator:

Theorem 2.1

$$\widehat{\theta}_{n,T} - \theta_T = O_P\left(\tau_n^{\min(1/2, \gamma - \rho - \alpha_\sigma)} + \kappa(n)[\eta_{\gamma - \epsilon}(n) + (k\tau_n)^{-\rho} + 1]\eta_\gamma(n)\right).$$

Proof. We use Le Breton's general theorem on AML estimation (Lemma 2.1). As one can see, in the linearly parametrized case, the rate of convergence of approximate MLE to continuous MLE is given by the rate of convergence of the corresponding approximate log-likelihood to continuous log-likelihood. □

Backward AMLE (Based on the Implicit Euler Scheme)

Define

$$G_{n,T} = \sum_{k=1}^{n} \left[\frac{v_{k+1}^n - (1 - \tau_n \theta A_n)^{-1} v_k^n - \tau_n P_n f(k\tau_n, v_k^n)}{\sqrt{\tau_n}\sigma_n(v_k^n)} \right]^2 \tag{2.6}$$

and

$$\widetilde{\theta}_{n,T} = \text{argmin}_\theta G_{n,T}$$

which is given by

$$\widetilde{\theta}_{n,T} = \frac{\sum_{k=1}^n A_n v_{k+1}^n (v_{k+1}^n - v_k^n) - \tau_n P_n f(k\tau_n, v_k^n)}{\sum_{k=1}^n (A_n v_k^n)^2 \tau_n}.$$

The following rate of convergence holds for the backward estimator:

Theorem 2.2

$$\widetilde{\theta}_{n,T} - \theta_T = O_P\left(\tau_n^{\min(1/2, \gamma - \rho - \alpha_\sigma)} + \tau_n \|(-A_n)^{\max(0, 1+\rho+\tau_\sigma - \gamma)}\| \right.$$
$$\left. + \kappa(n)[\eta_{\gamma-\epsilon}(n) + (k\tau_n)^{-\rho} + 1]\eta_\gamma(n)\right).$$

Proof. It is a direct consequence of Lemma 2.1. □

Midpoint AMLE (Based on the Crank-Nicholson Scheme)

If in addition, $\sigma : H \to H$ is unbounded and satisfies the stability condition

$$\|(1 + \frac{\tau_n}{2} A_n)^k\| \le M e^{k\tau_n}$$

or $\sigma : H \to H$ is bounded, then define

$$J_{n,T} = \sum_{k=1}^{n} \left[\frac{v_{k+1}^n - (1 + \tau_n \theta A_n)(\frac{v_k^n + v_{k+1}^n}{2}) - \tau_n P_n f(k\tau_n, v_k^n)}{\sqrt{\tau_n}\sigma_n(v_k^n)} \right]^2 \tag{2.7}$$

and

$$\check{\theta}_{n,T} = \text{argmin}_\theta J_{n,T}$$

which is given by

$$\check{\theta}_{n,T} = \frac{\sum_{k=1}^n A_n (\frac{v_k^n + v_{k+1}^n}{2})(v_{k+1}^n - v_k^n) - \tau_n P_n f(k\tau_n, v_k^n)}{\sum_{k=1}^n (A_n v_k^n)^2 \tau_n}.$$

The following rate of convergence holds for the trapezoidal estimator:

Theorem 2.3

$$\check{\theta}_{n,T} - \theta_T = O_P\left(\tau_n^{\min(1/2,\gamma-\rho-\tau_\sigma)} + \kappa(n)[\eta_{\gamma-\epsilon}(n) + (k\tau_n)^{-\rho} + 1]\eta_\gamma(n)\right).$$

The rate is same as the backward estimator.

Proof. It is a direct consequence of Lemma 2.1. □

3. Approximate Maximum Likelihood Estimation in Controlled Semilinear SPDE

We estimate the drift in the controlled semilinear SPDE by approximation by space and time discretization. We study parameter dependent adaptive, ergodic cost stochastic control problem in a semilinear SPDE. Adaptive control problem is to find a family of consistent estimates of the unknown parameter and to determine an adaptive control from a family of admissible controls such that the optimal ergodic cost is achieved. Oksendal and Sulem (2007) applied stochastic control problem to finance. Stannat and Wessels (2022) provided necessary and sufficient conditions for optimal control of semilinear stochastic partial differential equations. Clairon and Samson (2022) studied optimal control for parameter estimation in partially observed hypoelliptic stochastic differential equations. Bishwal (2022b) studied parameter estimation in stochastic volatility models.

Let H be a separable Hilbert space. Let A be an infinitesimal generator of an analytical semigroup of negative type. Let W_t be Wiener process taking values in H with covariance operator Q. Consider the stochastic evolution equation

$$dX_t = (AX_t + f(\theta, X_t) - Y_t)dt + \sqrt{Q}dW_t, \quad X_0 = x \in H. \tag{3.1}$$

Controls Y are taken from a set of admissible controls Y which consist of progressively measurable processes $Y : \mathbb{R}_+ \times \Omega \to H$ such that $P(Y_t \in B(r_0)) = 1$ for all $t \geq 0$ where $B(r_0) \subset H$ stands for a centered ball with radius r_0.

Let the family of admissible controls be given by $Y = \{Y : \mathbb{R}_+ \times \Omega \to B_R \mid Y$ is measurable and (F_t) adapted$\}$ where $B_R = \{y \in H : |y| > R\}$ and $R > 0$ is fixed. A family of Markov control $Y(t) = \widetilde{Y}(X_t)$ is also considered where $\widetilde{Y} \in \widetilde{Y}$ where $\widetilde{Y} = \{\widetilde{Y} : H \to B_R |$ is measurable$\}$.

The cost functionals $J(x, \lambda, Y)$ and $\widetilde{J}(x, Y)$ are given by

$$J(x, \lambda, Y) = E_{x,Y} \int_0^\infty e^{-\lambda t}(\psi(X_t + h(Y_t))dt$$

and

$$\widetilde{J}(x,Y) = \liminf_{T\to\infty} E_{x,Y}\frac{1}{T}\int_0^T e^{-\lambda t}(\psi(X_t + h(Y_t))dt$$

where $\lambda > 0$, $h : B_R \to \mathbb{R}_+$, and $\psi : H \to \mathbb{R}$, that describe a discounted and ergodic cost control problem, respectively.

The adaptive control problem is to find a family of strongly consistent estimates of the unknown parameter θ and to determine an adaptive control from the family of admissible controls such that the optimal ergodic cost is achieved. Thus one wants to achieve $\inf_{Y\in Y} \widetilde{J}(x,Y)$.

Duncan *et al.* (2000) studied the linearly parametrized case. The following are the standing assumptions of this section:

(D1) f is Lipschitz and Gateaux differentiable and A generates a strongly continuous semigroup on H.

(D2) f is Hölder continuous in θ:

$$|f(\theta_1,x) - f(\theta_2,x)| \le k|\theta_1-\theta_2|^\alpha(1+|x|^p),\quad \theta_1,\theta_2\in\theta,\ x\in H.$$

(D3) A generates a C_0-semigroup (S_t) on H. Moreover, there exists $\omega\in\mathbb{R}$ such that

$$\|S_t\|_{HS} \le e^{-\omega t},\ \forall\, t\ge 0$$

for all $t\ge 0$ and some $\omega > 0$.

(D4) There exists $\gamma > 0$ such that $\int_0^T t^{-\gamma}\|S_t\|_{HS}dt < \infty$ for some $T > 0$.

(D5)

$$\langle Ax + f(\theta, x+y),\ x\rangle \le -K|x|^2 + g(|y|)|x|,\ x\in D(A),\ y\in H$$

where $g : \mathbb{R}_+ \to \mathbb{R}_+$ is a continuous increasing function.

(D6) (Identifiability condition) For every $\theta\in\Theta, \theta\ne\theta_0$, there exists $x\in H$ such that $f(\theta,x)\ne f(\theta_0,x)$.

Let $P : H \to P(H)$ be a fixed finite dimensional projection on H with range in $Dom(A^*)$ that is chosen to satisfy (D7) given below:

(D7) For each admissible control law

$$\liminf_{t\to\infty}\int_0^t |Pf(\theta,X_s)|^2 ds > 0 \ \text{ a.s.}$$

If (D2) is satisfied, then the nonlinear equation (3.1) has a unique mild solution

$$X_t = S_t x + \int_0^t S_{t-r}(f(\theta,X_r) - Y_r)dr + \int_0^t S_{t-r}Q^{1/2}dW_r \tag{3.2}$$

for each $d\in D$ and $\theta\in\Theta$. If the control in (3.1) has the feedback form $Y_t = \widetilde{Y}(X_t)$ where $\widetilde{Y}\in\widetilde{Y}$, then the solution of (3.1) is obtained by an absolute continuity of measures as a weak solution in the probabilistic sense.

For a certainty equivalence adaptive control and a consistent family of estimators of the unknown parameter, it is shown that the adaptive control is self-optimizing. Consider (3.1) with the true parameter value $\theta_0 \in \Theta$,

$$dX_t = (AX_t + f(\theta_0, X_t) - \widetilde{Y}(X_t))dt + Q^{1/2}dW_t, \quad X_0 = x \tag{3.3}$$

where the adaptive control can be defined in the feedback form

$$\widetilde{Y}(X_t) = D\widetilde{H}(DZ_{\theta_t}(X_t))$$

and $\widetilde{H}$ is the Hamiltonian of the problem, $(\theta_t, t \geq 0)$ is an adapted, measurable process satisfying $\lim_{t\to\infty} \theta_t = \theta_0$ in probability and Z_θ is the solution of the infinite dimensional HJB equation.

Let $\bar{Y}_n = D\widetilde{H}(DZ_n)$ and $\bar{Y} = D\widetilde{H}(DZ_\theta)$ be controls where D is Gateux derivative. For an arbitrary $\Phi \in C([0,T], H)$, it follows that

$$\begin{aligned} &|E_{\tau,x,\bar{Y}_n}\Phi(X(\cdot)) - E_{\tau,x,\bar{Y}}\Phi(X(\cdot))| \\ &\leq E_{\tau,x}\Phi(X(\cdot))\left|\exp\left(\int_0^T \langle Q^{-1/2}\bar{Y}_n(X_s), dW_s\rangle - \frac{1}{2}\int_0^T |Q^{-1/2}\bar{Y}_n(X_s)|^2 ds\right)\right. \\ &\quad \left. - \exp\left(\int_0^T \langle Q^{-1/2}\bar{d}(X_s), dW_s\rangle - \frac{1}{2}\int_0^T |Q^{-1/2}\bar{Y}(X_s)|^2 ds\right)\right| \to 0 \end{aligned}$$

as $n \to \infty$ for $\tau \in [0,T)$. Finally, by the Skorohod's theorem

$$\lim_{n\to\infty} \sup_{s\in[\tau,T]} |X_n(s) - X_0(s)| = 0 \quad \text{a.s.}$$

for each $T > \tau$.

We turn to the estimation problem. The MLE $\hat{\theta}_T$ is the maximizer of the log-likelihood:

$$L(t,\theta) = \int_0^T \langle Pf(\theta, X_s), dPX(s)\rangle - \frac{1}{2}\int_0^T |Pf(\theta, X_s)|^2 ds.$$

In the linearly parametrized case $f(\theta, x) = f_0(x) + \sum_{r=1}^q \theta f_r(x)$, for the MLE of θ, using time change in the components of the stochastic integral and the law of large numbers for Brownian motion, we have the MLE error

$$e_j(T) = \left(\int_0^T \langle Pf(\theta, X_s), dPX_s\rangle\right)^{-1} \int_0^T \langle Pf(\theta, X_s), dPQ^{1/2}W_s\rangle \to 0 \quad \text{a.s.} \quad \text{as } T \to \infty$$

giving the strong consistency of the MLE, see Duncan *et al.* (2000).

We study the nonlinear case. Let

$$M_T(\theta) := \int_0^T \langle Q^{-1/2}(f(\theta, X_s) - f(\theta_0, X_s), dW_s\rangle,$$

$$S_T(\theta) := \int_0^T \langle Q^{-1/2}(f(\theta, X_s) - f(\theta_0, X_s), dX_s\rangle,$$

$$R_T(\theta) := \int_0^T |Q^{-1/2}(f(\theta, X_s) - f(\theta_0, X_s)|^2 ds.$$

It can be shown that the map $(t, \theta) \to M_t(\theta)$ has a continuous modification on $(0, \infty) \times \Theta$.

Let P_θ be the measure generated by the process X_t and let θ_0 be the true value of the parameter θ. The likelihood ratio (Radon-Nikodym derivative) of is given by

$$\begin{aligned} L_T(\theta) &= \frac{dP_\theta}{dP_{\theta_0}} \\ &= \exp\left\{\int_0^T \langle Q^{-1/2}(f(\theta, X_s) - f(\theta_0, X_s), dX_s\rangle \right. \\ &\quad \left. -\frac{1}{2}\int_0^T |Q^{-1/2}(f(\theta, X_s) - f(\theta_0, X_s)|^2 ds. \right\} \\ &= \exp\left\{S_T(\theta) - \frac{1}{2}R_T(\theta)\right\}. \end{aligned}$$

In practice, the process $\{X_t, t \geq 0\}$ can not be observed continuously in time. We assume that the process $\{X_t, t \geq 0\}$ is observed at $0 = t_0 < t_1 < \ldots < t_n = T$ with $\Delta t_i := t_i - t_{i-1} = \frac{T}{n} = h$ and $T = dn^{1/2}$ for some fixed real number $d > 0$. We estimate θ from the observations $\{X_{t_0}, X_{t_1}, \ldots, X_{t_n}\}$. The Euler approximate log-likelihood is given by

$$l_{n,T}(\theta) := S_{n,T}(\theta) - \frac{1}{2}R_{n,T}(\theta)$$

where

$$S_{n,T}(\theta) := \sum_{i=1}^n \langle Q^{-1/2}(f(\theta, X_{t_{i-1}}) - f(\theta_0, X_{t_{i-1}}), (X_{t_i} - X_{t_{i-1}})\rangle,$$

$$R_{n,T}(\theta) := \sum_{i=1}^n |Q^{-1/2}(f(\theta, X_{t_{i-1}}) - f(\theta_0, X_{t_{i-1}})|^2 (t_i - t_{i-1}).$$

Approximate maximum likelihood estimator (AMLE) is defined as

$$\theta_{n,T} = \text{argmax}_{\theta\in\Theta} l_{n,T}.$$

We obtain the strong consistency of the AMLE in the next theorem.

Theorem 3.1 (Strong Consistency)

$\hat{\theta}_{n,T} \to \theta_0$ almost surely as $T \to \infty$ and $T/n \to 0$.

Proof. Observe that

$$\hat{\theta}_{n,T} - \theta_0 = \frac{M_{n,T}(\theta) + N_{n,T}(\theta)}{R_{n,T}(\theta)} = \frac{\frac{1}{T}M_{n,T}(\theta) + \frac{1}{T}N_{n,T}(\theta)}{\frac{1}{T}R_{n,T}(\theta)}$$

where

$$N_{n,T}(\theta) := \sum_{i=1}^{n} \left| Q^{-1/2} \int_{t_{i-1}}^{t_i} (f(\theta, X_{t_{i-1}}) - f(\theta_0, X_t))dt \right|^2$$

and

$$M_{n,T}(\theta) := \sum_{i=1}^{n} \langle Q^{-1/2}(f(\theta, X_{t_{i-1}}) - f(\theta_0, X_{t_{i-1}}),\ W_{t_i} - W_{t_{i-1}} \rangle.$$

Since M_T is a continuous martingale, by Dambis-Dubins-Schwarz theorem on clock change,

$$\lim_{T\to\infty} \frac{1}{T} M_T(\theta) = \lim_{T\to\infty} \frac{W^*(R_T(\theta))}{R_T(\theta)} \frac{R_T(\theta)}{T} = 0 \ \ \mathrm{P-a.s.}$$

where W^* is another Brownian motion independent of W. Due to ergodic theorem,

$$\lim_{T\to\infty} \frac{1}{T} R_T(\theta) = E|Q^{-1/2}(f(\theta, X_0) - f(\theta_0, X_0))|^2 \ \ \mathrm{P-a.s.}$$

It is easily seen that $M_{n,T}(\theta) - M_T(\theta) \to^P 0$ as $T/n \to 0$. We prove stronger result. We show that

a) $\frac{1}{T} M_{n,T}(\theta) \to 0$ almost surely as $T \to \infty$ and $T/n \to 0$,

b) $\frac{1}{T} N_{n,T}(\theta) \to 0$ almost surely as $T \to \infty$ and $T/n \to 0$.

a) Let $v(\theta, x) := f(\theta, x) - f(\theta_0, x)$. The Fourier expansion of $v(\theta, x)$ in $L(\Theta)$ be given by

$$v(\theta, x) = \sum_{m=1}^{\infty} a_m(x) e^{\pi j m \theta}, \quad j = \sqrt{-1},\ x \in \mathbb{R}$$

where $a_k(x)$ are the Fourier coefficients. Thus

$$\frac{1}{T} \sum_{i=1}^{n} \langle Q^{-1/2} v(\theta, X_{t_{i-1}}),\ \Delta W_i \rangle = \frac{1}{T} \sum_{m=1}^{\infty} \sum_{i=1}^{n} \langle Q^{-1/2} a_m(X_{t_{i-1}}) e^{\pi j m \theta},\ \Delta W_i \rangle$$

where

$$|a_m(x)| \le c_m |x|, \quad \sum_{m=1}^{\infty} m^{1+\gamma} c_m^4 < \infty.$$

Let

$$A_{m,n}(s) := \sum_{i=1}^{n} a_m(X_{t_{i-1}}) I_{(t_{i-1} - t_i]}(s)$$

where $I_{(t_{i-1} - t_i]},\ i = 1, 2, ..., n$ are indicator functions. Then

$$\sum_{i=1}^{n} \langle Q^{-1/2} a_m(X_{t_{i-1}}),\ \Delta W_i \rangle = \int_0^T \langle Q^{-1/2} A_{m,n}(s),\ dW_s \rangle$$

By exponential inequality for martingales, we have

$$P\left\{ \int_0^T A_{m,n}(s) dW_s - \frac{\alpha}{2} \int_0^T |Q^{-1/2} A_{m,n}|^2 ds > \beta \right\} \le e^{-\alpha\beta}$$

for any $\alpha, \beta > 0$. Thus

$$P\left\{ \frac{1}{T} \int_0^T \langle Q^{-1/2} A_{m,n}(s), dW_s \rangle > \frac{\beta}{T} + \frac{\alpha}{2T} \int_0^T |Q^{-1/2} A_{m,n}|^2 ds \right\} \le e^{-\alpha\beta}$$

and

$$P\left\{\left|\frac{1}{T}\int_0^T \langle Q^{-1/2}A_{m,n}(s), dW_s\rangle\right| > \frac{\beta}{T} + \frac{\alpha h}{2T}\sum_{i=1}^n |Q^{-1/2}a_m(X_{t_{i-1}})|^2\right\} \le 2e^{-\alpha\beta}.$$

Since

$$\frac{h}{T}\sum_{i=1}^n |Q^{-1/2}a_m(X_{t_{i-1}})|^2 \le c_m^2 \frac{h}{T}\sum_{i=1}^n |Q^{-1/2}X_{t_{i-1}}|^2$$

and by ergodicity

$$\frac{h}{T}\sum_{i=1}^n |Q^{-1/2}X_{t_{i-1}}|^2 \to E|Q^{-1/2}X_0|^2 > 0 \text{ a.s.},$$

there exists a random variable V such that

$$\frac{h}{T}\sum_{i=1}^n |Q^{-1/2}X_{t_{i-1}}|^2 < V \text{ a.s.}$$

for all $T > 0, n = 1, 2, \ldots.$ where $P(V < \infty) = 1$.

Denote

$$Z_{m,n} := \frac{1}{t_n}\int_0^{t_n} \langle Q^{-1/2}A_{m,n}(s),\ dW_s\rangle.$$

Recall that $T = t_n$. Choose

$$\alpha := \frac{m^a}{t_n^\delta}, \quad \beta := \frac{t_n^\gamma}{m^b},$$

where $\delta < \gamma < 1$ and $\frac{1}{2} < b < \frac{1+\gamma}{2}$.

Then

$$P\left(|Z_{m,n}| > \frac{1}{t_n^{1-\gamma}m^b} + \frac{m^a c_m^2 V}{2t_n^\delta}\right) < 2e^{-m^{a-b}t_n^{\gamma-\delta}}.$$

This

$$\begin{aligned}
&P\left(\sum_{m=1}^\infty Z_{m,n}^2 > \sum_{m=1}^\infty \left(\frac{1}{t_n^{1-\gamma}m^b} + \frac{m^a c_m^2 V}{2t_n^\delta}\right)^2\right) \\
&\le \sum_{m=1}^\infty P\left(Z_{m,n}^2 > \left(\frac{1}{t_n^{1-\gamma}m^b} + \frac{m^a c_m^2 V}{2t_n^\delta}\right)^2\right) \\
&= \sum_{m=1}^\infty P\left(|Z_{m,n}| > \frac{1}{t_n^{1-\gamma}m^b} + \frac{m^a c_m^2 V}{2t_n^\delta}\right) \\
&\le 2\sum_{m=1}^\infty e^{-m^{a-b}t_n^{\gamma-\delta}} \\
&\le 2e^{-t_n^{\gamma-\delta}}\sum_{m=1}^\infty e^{-m^{a-b}}.
\end{aligned}$$

Hence

$$\sum_{n=1}^\infty P\left(\sum_{m=1}^\infty Z_{m,n}^2 > \sum_{m=1}^\infty \left(\frac{1}{t_n^{1-\gamma}m^b} + \frac{m^a c_m^2 V}{2t_n^\delta}\right)^2\right)$$

$$\leq 2\sum_{n=1}^{\infty} e^{-t_n^{1-\gamma}} \sum_{m=1}^{\infty} e^{-m^{a-b}} < \infty$$

since $\gamma - \delta > 0$ and $a - b > 0$. The above implies

$$\sum_{n=1}^{\infty} P\left(\sum_{m=1}^{\infty} Z_{m,n}^2 > \frac{2}{t_n^{2(1-\gamma)}} \sum_{m=1}^{\infty} m^{-2b} + \frac{V^2}{t_n^{2\delta}} \sum_m m^{2a} c_m^4\right) < \infty.$$

By Borel-Cantelli lemma,

$$\sum_{m=1}^{\infty} \left(\frac{1}{t_n} \sum_{i=1}^{n} \langle Q^{-1/2} a_m(X_{t_{i-1}}),\ \Delta W_i\rangle\right)^2 \longrightarrow 0 \text{ a.s. as } n \to \infty.$$

Thus $\frac{1}{T} M_{n,T}(\theta) \to 0$ almost surely as $T \to \infty$ and $T/n \to 0$. This completes the proof of part a).

b) Next we show that $\frac{1}{T} N_{n,T}(\theta) \to 0$ almost surely as $T \to \infty$ and $T/n \to 0$.

For $m > 0$, we have

$$\begin{aligned}
& E\left\{\sup_{\theta\in\Theta} \left| \frac{1}{T} \sum_{i=1}^{n} Q^{-1/2} \int_{t_{i-1}}^{t_i} [f(\theta_0, X_s) - f(\theta_0, X_{t_{i-1}})] v(\theta, X_{t_{i-1}}) ds \right|^{2m}\right\} \\
= \ & E\left\{\sup_{\theta\in\Theta} \left| \frac{1}{T} \int_0^T G_n(s) ds \right|^{2m}\right\}.
\end{aligned}$$

where $G_n(s) = \sum_{i=1}^{n} Q^{-1/2} \int_{t_{i-1}}^{t_i} [f(\theta_0, X_s) - f(\theta_0, X_{t_{i-1}})] v(\theta, X_{t_{i-1}})$ if $t_{i-1} \leq s \leq t_i$.

Hölder's inequality implies that

$$\begin{aligned}
& E\left\{\sup_{\theta\in\Theta} \left| \frac{1}{T} \int_0^T G_n(s) ds \right|^{2m}\right\} \\
\leq \ & T^{-2m} E\left\{\sup_{\theta\in\Theta} T^{2m-1} \int_0^T |G_n(s)|^{2m} ds\right\} \\
\leq \ & T^{-2m} E\left(\sup_{\theta\in\Theta} T^{2m-1} \sum_{i=1}^{n} Q^{-1/2} \int_{t_{i-1}}^{t_i} |f(\theta_0, X_s) - f(\theta_0, X_{t_{i-1}})|^{2m} |v(\theta, X_{t_{i-1}})|^{2m} ds\right) \\
& \leq T^{-1} U_m \sum_{i=1}^{n} Q^{-1/2} \int_{t_{i-1}}^{t_i} E(|f(\theta_0, X_s) - f(\theta_0, X_{t_{i-1}})|^{2m} |C(X_{t_{i-1}})|^{2m} ds)
\end{aligned}$$

by condition (D2) where $U_m := \sup_{\theta\in\Theta} |\theta - \theta_0|^{2m} < \infty$.

By Cauchy-Schwarz's inequality the above term is

$$\begin{aligned}
& \leq T^{-1} U_m \sum_{i=1}^{n} Q^{-1/2} \int_{t_{i-1}}^{t_i} (E|f(\theta_0, X_s) - f(\theta_0, X_{t_{i-1}})|^{4m})^{1/2} (E(C(X_{t_{i-1}})|^{4m})^{1/2} ds \\
\leq \ & T^{-1} U_m K^{2m}(\theta_0) (E|C(Q^{-1/2}(X_0))|^{4m})^{1/2} \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} (E|Q^{-1/2}(X_s - X_{t_{i-1}})|^{4m})^{1/2} ds
\end{aligned}$$

by condition (D2). Since $E|Q^{-1/2}(X_t - X_s)|^{2m} \le M(t-s)^m$, from Gikhman and Skorohod (1975, p.48), the above term

$$
\begin{aligned}
&\le T^{-1}U_m K^{2m}(\theta_0)(E|C(Q^{-1/2}X_0)|^{4m})^{1/2}M^{1/2}\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}(s-t_{i-1})^m ds \\
&= U_m K^{2m}(\theta_0)(E|C(Q^{-1/2}(X_0))|^{4m}M)^{1/2}T^{-1}\sum_{i=1}^{n}\frac{(\Delta t_i)^{m+1}}{m+1} \\
&\le \frac{U_m K^{2m}(\theta_0)}{m+1}(E|C(Q^{-1/2}(X_0))|^{4m}M)^{1/2}h^m n^{-m/2}, \quad m > 4.
\end{aligned}
$$

Chebyshev's inequality and the above implies that for any $\epsilon > 0$,

$$\sum_{n=1}^{\infty} P\left\{\sup_{\theta\in\Theta}\left|\frac{1}{T}\sum_{i=1}^{n} Q^{-1/2}\int_{t_{i-1}}^{t_i}[f(\theta_0, X_s) - f(\theta_0, X_{t_{i-1}})]v(\theta, X_{t_{i-1}})ds\right| > \epsilon\right\} < \infty.$$

Hence Borel-Cantelli lemma yields the result. This completess the proof of part b). By the same method along with ergodicity, it can be shown that

$$\frac{1}{T}R_{n,T}(\theta) \to E|Q^{-1/2}(f(\theta, X_0) - f(\theta_0, X_0))|^2 \ \ \mathrm{P-a.s.} \ \ \text{as} \ \ T\to\infty \ \ \text{and} \ \ T/n \to 0.$$

This completes the proof of the theorem. □

In order to prove asymptotic normality we need the following lemma from Bishwal (2008).

Lemma 3.1 Let ξ, ζ and η be any three random variables on a probability space (Ω, F, P) with $P(\eta > 0) = 1$. Then, for any $\epsilon > 0$, we have

$$(a)\ \sup_{x\in\mathbb{R}}|P\{\xi+\zeta \le x\} - \Phi(x)| \le \sup_{x\in\mathbb{R}}|P\{\xi \le x\} - \Phi(x)| + P(|\zeta| > \epsilon) + \epsilon,$$

$$(b)\ \sup_{x\in\mathbb{R}}|P\{\frac{\xi}{\eta} \le x\} - \Phi(x)| \le \sup_{x\in\mathbb{R}}|P\{\xi \le x\} - \Phi(x)| + P\{|\eta - 1| > \epsilon\} + \epsilon.$$

Next theorem gives the asymptotic normality of the AMLE.

Theorem 3.2 (Asymptotic Normality)

$$\sqrt{TI(\theta_0)}(\hat{\theta}_{n,T} - \theta_0) \to N(0,1) \text{ in distribution as } T\to\infty \text{ and } T/\sqrt{n} \to 0.$$

Proof. By Taylor expansion of the log-likelihood, we have

$$l'_{n,T}(\hat{\theta}_{n,T}) = l'_{n,T}(\theta_0) + (\hat{\theta}_{n,T} - \theta_0)l''_{n,T}(\bar{\theta}_{n,T})$$

where $\left|\bar{\theta}_{n,T}-\theta\right| \leq \left|\hat{\theta}_{n,T}-\theta_0\right|$. Since $l'_{n,T}(\hat{\theta}_{n,T})=0$, hence we have

$$\sqrt{TI(\theta_0)}(\hat{\theta}_{n,T}-\theta_0) = -\frac{\frac{1}{\sqrt{TI(\theta_0)}}l'_{n,T}(\theta_0)}{\frac{1}{TI(\theta_0)}l''_{n,T}(\bar{\theta}_{n,T})} = -\frac{\frac{1}{\sqrt{TI(\theta_0)}}\sum_{i=1}^n\langle Q^{-1/2}f'(\theta_0,X_{t_{i-1}}),\Delta W_i\rangle}{\frac{1}{TI(\theta_0)}\sum_{i=1}^n Q^{-1}f''(\bar{\theta}_{n,T},X_{t_{i-1}})}$$
$$=: \frac{M_{n,T}}{V_{n,T}}.$$

Note that

$$V_{n,T} = \frac{1}{TI(\theta_0)}\sum_{i=1}^n Q^{-1}f''(\bar{\theta}_{n,T},X_{t_{i-1}})\Delta t_i = \frac{1}{TI(\theta_0)}\sum_{i=1}^n |Q^{-1/2}f'\bar{\theta}_{n,T},X_{t_{i-1}}|^2\Delta t_i.$$

But $E(I_T-1)^2 \leq CT^{-1}$ (see Altmeyer and Chorowski (2018)). It can be shown that $E[(V_{n,T}-I_T)^2 \leq C\frac{T}{n}$. Hence

$$E(V_{n,T}-1)^2 = E[(V_{n,T}-I_T)+(I_T-1)]^2 \leq C(T^{-1}\bigvee\frac{T}{n}).$$

Since

$$\begin{aligned}
&\sup_{x\in\mathbb{R}}\left|P_\theta\left\{M_{n,T}\leq x\right\}-\Phi(x)\right| \\
\leq\ &\sup_{x\in\mathbb{R}}\left|P_\theta\left\{M_{n,T}\leq x\right\}-\Phi(x)\right| + P_\theta\left\{|R_{n,T}-M_T|\geq\epsilon\right\}+\epsilon \\
\leq\ &C(T^{-1/2}\textstyle\bigvee\frac{T^2}{n}) + \epsilon^{-2}E\left|R_{n,T}-M_T\right|^2+\epsilon \\
\leq\ &(T^{-1/2}\textstyle\bigvee\frac{T^2}{n}) + \epsilon^{-2}C(T^{-1}\bigvee\frac{T^3}{n^2})+\epsilon.
\end{aligned}$$

we have

$$\begin{aligned}
&\sup_{x\in\mathbb{R}}\left|P_\theta\left\{\sqrt{TI(\theta)}(\hat{\theta}_{n,T}-\theta)\leq x\right\}-\Phi(x)\right| \\
=\ &\sup_{x\in\mathbb{R}}\left|P_\theta\left\{\frac{M_{n,T}}{V_{n,T}}\leq x\right\}-\Phi(x)\right| \\
=\ &\sup_{x\in\mathbb{R}}\left|P_\theta\left\{M_{n,T}\leq x\right\}-\Phi(x)\right| + P_\theta\left\{|V_{n,T}-1|\geq\epsilon\right\}+\epsilon \\
\leq\ &C(T^{-1/2}\textstyle\bigvee\frac{T^2}{n}) + \epsilon^{-2}C(T^{-1}\bigvee\frac{T^2}{n^2})+\epsilon.
\end{aligned}$$

Choosing $\epsilon = T^{-1/2}$ completes the proff of the theorem. $\square$

4. Examples

1) Non-linear Case: Let $(\Omega, F, \{F_t\}_{t\geq 0}, P)$ be a stochastic basis with the usual assumptions on which countably many independent standard Brownian motions $W_k = \{W_k(t), t\geq 0\}, k=1,2,\ldots$ are defined. Let G be a domain in $\mathbb{R}^d$. The space time Gaussian white noise $\dot{W}=\dot{W}(t,x)$ on G is a collection of zero mean Gaussian random variables $\dot{W}[g_1], g_1\in L_2((0,\infty)\times G)$ such that

$$E(\dot{W}[g_1]\dot{W}[g_2]) = \int_0^\infty\int_G g_1(t,x)g_2(t,x)dxdt.$$

Given an orthonormal basis $\{h_k = h_k(x), k \geq 1\}$ in $L_2(G)$, the process $\dot{W}$ can be written as a formal sum

$$\dot{W}(t,x) = \sum_{k=1}^{\infty} h_k(x)\dot{W}_k(t).$$

Similarly,

$$W(t,x) = \sum_{k=1}^{\infty} h_k(x)W_k(t)$$

is called cylindrical Brownian motion in $L_2(G)$. For a square integrable function $g_1 = g_1(t,x)$,

$$\int_0^t \int_G g_1(s,y)W(ds,dy) = \sum_{k=1}^{\infty} \left(\int_0^t \int_G g_1(s,y)h_k(y)dy \right) dW_k(s).$$

The covariance function cylindrical Brownian motion W is given by

$$EW(t_1,\zeta_1)W(t_2,\zeta_2) = (t_1 \wedge t_2)(\zeta_1 \wedge \zeta_2), \quad t_i \geq 0, \ \zeta_i \in [0,1], \ i = 1,2.$$

Consider the nonlinear SPDE where $H = L^2(0,1)$ and

$$A_0 x(\zeta) = \frac{\partial}{\partial \zeta}\left(a \frac{\partial x}{\partial \zeta} \right)(\zeta) + b(\zeta)\frac{\partial x}{\partial \zeta}(\zeta) + c_0(\zeta)x(\zeta)$$

endowed with the Dirichlet boundary conditions. We assume that a is Lipschitz on $[0,1]$, $b, c_0 \in L^\infty(0,1)$ and $a^2(x) \geq m > 0, \ \zeta \in [0,1]$. Under these assumptions, A generates an analytic C_0-semigroup of contractions on $L^2(0,1)$. It is also known that A generates a semigroup of Hilbert-Schmidt operators on $L^2(0,1)$ and for all $T > 0$, $\int_0^T \|S_t\|_{HS}^2 ds < \infty < \infty$. The SPDE driven by space-time cylindrical Wiener process has unique solution.

The hypotheses (D2), (D3), (D4) and (D5) are satisfied in the present case and therefore the SPDE

$$\frac{\partial X}{\partial t}(t,\zeta) = \frac{\partial}{\partial \zeta}\left(a \frac{\partial X}{\partial \zeta} \right)(t,\zeta) + b(\zeta)\frac{\partial X}{\partial \zeta}(t,\zeta) + c_0(\zeta)X(t,\zeta) + f_0(\theta_0, X(t,\zeta)) - u_t(\zeta) + \frac{\partial^2 W}{\partial t \partial \zeta}$$

$$X(0,\zeta) = x(\zeta), \zeta \in [0,1], X(t,0) = X(t,1) = 0$$

has a unique solution. As a consequence, we obtain the strong consistency of the MLE θ_t of the parameter θ_0 provided the family of measures $\{\mu_t : t \geq 0\}$ is relatively compact on $L^2(0,1)$. In particular, this condition is satisfied if $u_t(\zeta) = K(\theta_t, X(t,\zeta))$ where the function $K : \Theta \times \mathbb{R} \to \mathbb{R}$ is uniformly bounded and continuous.

2) Linear Case: When $f \equiv 0$ and $u \equiv 0$ in (3.1), a linear SPDE is obtained by

$$dX_t = AX_t dt + Q^{1/2} dW_t, \quad X_0 = x$$

whose solution is given by

$$X_t = S_t x + \int_0^t S_{t-r} Q^{1/2} dW_r$$

which is an H valued process with continuous sample paths, is ergodic and has a unique invariant probability measure $\mu = N(0, Q_\infty)$ where $Q_\infty = \int_0^\infty S_r Q S_r^* dr$ is a trace class operator on H.

Ergodic theorem for the Ornstein-Uhlenbeck process X yields for $t \to \infty$,

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t \phi(X_s) ds = \int_H \phi(y)\mu(dy) \quad \text{a.s.}$$

5. Monte Carlo Method

We consider the linear case. The generalized stochastic forcing term we consider is an additive space-time noise. Formally, we may write

$$\sigma dW = \sum_k \lambda_k^{-\alpha} \phi_k dW_k \tag{5.1}$$

where $\{\phi_k, k \geq 1\}$ are eigenfunctions of the operator, $\{\lambda_k, k \geq 1\}$ represents the associated eigenvalues and $\{W_k, k \geq 1\}$ are one-dimensional independent Brownian motions. We assume α is a real parameter greater than 1 which guarantees some spatial smoothness in the forcing. We may also derive the space-time correlation structure of the noise term

$$E(\sigma dW(x,t)\sigma dW(y,s)) = K(x,y)\delta_{t-s} \tag{5.2}$$

where

$$K(x,y) = \sum_{k\geq 1} \lambda_k^{-2\alpha} \phi_k(x)\phi_k(y). \tag{5.3}$$

Let $u_k(t) = (u_t, \phi_k)$ be the k-th generalized Fourier mode of the solution u_t. Each Fourier mode u_k represents a one dimensional Ornstein-Uhlenbeck process. In order to simulate the trajectories of the solution (Fourier modes), we discretize the SPDE (1.1). Let the Euler-Maruyama scheme of the solution be

$$\widetilde{u}_k^j(t_i) = \widetilde{u}_k^j(t_{i-1}) - \theta\lambda_k^2 \widetilde{u}_k^j(t_{i-1})\Delta T + \sigma\lambda_k^{-\alpha}\xi_{k,i}^j,\ u_k^j(t_i) = u_k(0),\ 1 \leq j \leq l,\ \ 1 \leq i \leq m,$$

$1 \leq k \leq n$ where $\xi_{k,i}^j$ are i.i.d. Gaussian random variables with zero mean and variance $\Delta T = T/m = t_i - t_{i-1}, 1 \leq i \leq m$ and l denotes the number of trials in the Monte Carlo experiment for each Fourier mode. Hence $\widetilde{u}_k^j(t_i)$ is the approximation of $u_k^j(t_i)$ which is the true value of the k-th Fourier mode at time t_i of the j-th trial in the Monte Carlo simulation. In what follows, we will investigate how to approximate the forward AMLE using $\widetilde{u}_k^j(t_i)$ and how the numerical errors are related to m, l, T and n. We obtain error estimates of the corresponding Monte Carlo experiments associated with the Euler-Maruyama scheme, see Glasserman (2004). We consider $d = 1$ with the random forcing term being space time white noise with $\alpha = 0$ and $\sigma = 1$. We assume the special domain $G = [0, \pi]$ and the initial value $u_0 = 0$. In this case $\lambda_k = k$. Using the likelihood ratio (2.3) and Itô formula, we get

$$
\begin{aligned}
&P_{\theta_0}^{n,T}(\sqrt{T}(\hat{\theta}_{n,T} - \theta) \geq \eta) = P_{\theta_0}^{n,T}(|\hat{\theta}_{n,T} - \theta| \geq \eta T^{-1/2}) = P_{\theta_0}^{n,T}(\ln L_T^{\theta_1}(u_n^T) \geq \eta T) \\
&= P_{\theta_0}^{n,T}\left(-\sum_{k=1}^{n} \lambda_k^{2+2\alpha}(\int_0^T u_k(t) du_k(t) + \frac{\theta_1 + \theta_0}{2\theta_0} \int_0^T u_k(t)(\sigma \lambda_k^{\alpha} dW_k(t) \right. \\
&\left. -du_k(t))) \geq \frac{\sigma^2 \eta T}{\theta_1 - \theta_0} \right) \\
&= P_{\theta_0}^{n,T}\left(\sum_{k=1}^{n} \lambda_k^{2+2\alpha}(\frac{\theta_1 - \theta_0}{2} \int_0^T u_k^2(T) - \sigma^2 \lambda_k^{-2\alpha} T) \right. \\
&\left. +(\theta_1 + \theta_0) \int_0^T u_k(t)(\sigma \lambda_k^{\alpha} dW_k(t))) \geq \frac{2\theta_0 \sigma^2 \eta T}{\theta_1 - \theta_0} \right) \\
&= P_{\theta_0}^{n,T}\left(\frac{(\theta_1 - \theta_0)}{2\sigma(\theta_1 + \theta_0)\sqrt{T}} \widetilde{X}_T - \widetilde{Y}_T / \sqrt{T} \geq \frac{2\theta_0 \sigma \Delta\eta \sqrt{T}}{\theta_1^2 - \theta_0^2} \right)
\end{aligned}
$$

where

$$
\eta := -\frac{(\theta_1 - \theta_0)^2}{4\theta_0} M + \frac{(\theta_1 - \theta_0)^2}{2\theta_0^2 T} + \frac{(\theta_1 - \theta_0)^2}{2\theta_0^2} \sqrt{-\theta_0 M T^{-1} \ln \alpha + T^{-2} \ln^2 \alpha},
$$

$$
\Delta\eta := \eta + \frac{(\theta_1 - \theta_0)^2}{4\theta_0} M, \quad \widetilde{X}_T := \sum_{k=1}^{n} \lambda_k^{2+2\alpha} u_k^2(T), \quad \widetilde{Y}_T := \sum_{k=1}^{n} \lambda_k^{2+\alpha} \int_0^T u_k(t) dW_k(t)
$$

and $M = \sum_{k=1}^{n} \lambda_k^2$. We approximate $\widetilde{X}_T$ and $\widetilde{Y}_T$ by

$$
\widehat{\widetilde{X}}^j_{n,T} := \sum_{k=1}^{n} \lambda_k^{2+2\alpha} \widetilde{u}_k^j(t_n)^2, \quad \widehat{\widetilde{Y}}^j_{n,T} := \sum_{k=1}^{n} \lambda_k^{2+\alpha} \sum_{i=1}^{m} \widetilde{u}_k^j(t_{i-1}) \xi_{k,i}^j \tag{5.4}
$$

respectively. Define

$$
\widetilde{R}^{0,j}_{m,T} := \left\{ \frac{\theta_1 - \theta_0}{2\sigma(\theta_1 + \theta_0)\sqrt{M}} \widehat{\widetilde{X}}^j_{m,T} - \widehat{\widetilde{Y}}^j_{m,T} / \sqrt{M} \geq \frac{2\theta_0 \sigma \Delta\eta}{\theta_1^2 - \theta_0^2} \sqrt{M} \right\}. \tag{5.5}
$$

The approximation of $P_{\theta_0}^{n,T}(\sqrt{T}(\hat{\theta}_{n,T} - \theta) \geq \eta)$ is given by

$$
\widetilde{P}_{\theta_0}^{l,m,n,T}(\sqrt{m}(\hat{\theta}_{l,m,n,T} - \theta) \geq \eta) = \frac{1}{l} \sum_{j=1}^{l} I_{\widetilde{R}^{0,j}_{m,T}}. \tag{5.6}
$$

The following theorem gives error estimate on the Monte Carlo simulations:

Theorem 5.1

$$
\sup_{x \in \mathbb{R}} \left| \widetilde{P}_{\theta_0}^{l,m,n,T}(\sqrt{m}(\hat{\theta}_{l,m,n,T} - \theta) \leq x) - P_{\theta_0}^{n,T}(\sqrt{T}(\hat{\theta}_{n,T} - \theta) \leq x) \right| \leq C m^{-1/3} + C l^{-1/2}.
$$

Proof. Let $\theta_1 = \theta_0 + r n^{-1/2}$, $r > 0$. Following Bishwal (2008, Chapter 8), one can show that

$$
E|(\widehat{\widetilde{Y}}^j_{m,T} - \widetilde{Y}_T)/\sqrt{T}|^2 = O(\Delta T), \quad E|\widehat{\widetilde{X}}^j_{m,T} - \widetilde{X}_T| = O(\Delta T).
$$

Consequently, for any $\epsilon > 0$, we have

$$P_{\theta_0}^{l,m,n,T}(\sqrt{m}(\hat{\theta}_{l,m,n,T} - \theta) \geq \eta) \leq P_{\theta_0}^{n,T}\left(\frac{(\theta_1 - \theta_0)}{2\sigma(\theta_1+\theta_0)\sqrt{T}}\widetilde{X}_T - \widetilde{Y}_T/\sqrt{T} \geq \frac{2\theta_0\sigma\Delta\eta\sqrt{T}}{\theta_1^2 - \theta_0^2}\right.$$

$$\left.-\epsilon\right) + P_{\theta_0}^{n,T}|\widetilde{\widetilde{Y}}^j_{m,T} - \widetilde{Y}_T|/\sqrt{T}) \geq \epsilon/2) + P_{\theta_0}^{n,T}\left(\frac{(\theta_1 - \theta_0)}{2\sigma(\theta_1+\theta_0)\sqrt{T}}|\widehat{\widetilde{X}}^j_{m,T} - \widetilde{X}_T| \geq \epsilon/2\right)$$

and

$$P_{\theta_0}^{n,T}\left(\frac{(\theta_1 - \theta_0)}{2\sigma(\theta_1+\theta_0)\sqrt{T}}\widetilde{X}_T - \widetilde{Y}_T/\sqrt{T} \geq \frac{2\theta_0\sigma\Delta\eta\sqrt{T}}{\theta_1^2 - \theta_0^2} - \epsilon\right)$$
$$\leq P_{\theta_0}^{n,T}(\sqrt{T}(\hat{\theta}_{n,T} - \theta) \leq \eta)(1 + C\epsilon).$$

From the above results and Chebyshev inequality, we conclude that, for any $x \in \mathbb{R}$,

$$P_{\theta_0}^{l,m,n,T}(\sqrt{m}(\hat{\theta}_{l,m,n,T} - \theta) \leq x)$$
$$\leq P_{\theta_0}^{n,T}(\sqrt{T}(\hat{\theta}_{n,T} - \theta) \leq x)(1 + C\epsilon) + C\epsilon^{-1}E|\widehat{\widetilde{X}}^j_{m,T} - \widetilde{X}_T|/\sqrt{T} + E|(\widehat{\widetilde{Y}}^j_{m,T} - \widetilde{Y}_T)/\sqrt{T}|^2.$$

Similarly, we have

$$P_{\theta_0}^{l,m,n,T}(\sqrt{m}(\hat{\theta}_{l,m,n,T} - \theta) \leq x)$$
$$\geq P_{\theta_0}^{n,T}(\sqrt{T}(\hat{\theta}_{n,T} - \theta) \leq x)(1 - C\epsilon) - C\epsilon^{-1}E|\widehat{\widetilde{X}}^j_{m,T} - \widetilde{X}_T|/\sqrt{T} - E|(\widehat{\widetilde{Y}}^j_{m,T} - \widetilde{Y}_T)/\sqrt{T}|^2.$$

Combining the above two inequalities, we obtain

$$|P_{\theta_0}^{l,m,n,T}(\sqrt{m}(\hat{\theta}_{l,m,n,T} - \theta) \leq x) - P_{\theta_0}^{n,T}(\sqrt{T}(\hat{\theta}_{n,T} - \theta) \leq x)|$$
$$\leq C\epsilon P_{\theta_0}^{n,T}(\sqrt{T}(\hat{\theta}_{n,T} - \theta) \leq x) + C\epsilon^{-1}E|\widehat{\widetilde{X}}^j_{m,T} - \widetilde{X}_T|/\sqrt{T} + E|(\widehat{\widetilde{Y}}^j_{m,T} - \widetilde{Y}_T)/\sqrt{T}|^2.$$

This implies that

$$|P_{\theta_0}^{l,m,n,T}(\sqrt{m}(\hat{\theta}_{l,m,n,T} - \theta) \leq x) - P_{\theta_0}^{n,T}(\sqrt{T}(\hat{\theta}_{n,T} - \theta) \leq x)| \leq C(\Delta T)^{1/3}.$$

It is straightforward to check that for large T

$$\mathrm{Var}\left(\frac{(\theta_1 - \theta_0)}{2\sigma(\theta_1+\theta_0)\sqrt{T}}\widetilde{X}_T - \widetilde{Y}_T/\sqrt{T}\right) \leq C.$$

From here one can show that

$$\mathrm{Var}\left(\frac{(\theta_1 - \theta_0)}{2\sigma(\theta_1+\theta_0)\sqrt{T}}\widehat{\widetilde{X}}^j_{m,T} - \widehat{\widetilde{Y}}^j_{m,T}/\sqrt{T}\right)$$
$$= \mathrm{Var}\left(\frac{(\theta_1 - \theta_0)}{2\sigma(\theta_1+\theta_0)\sqrt{T}}\widetilde{X}_T - \widetilde{Y}_T/\sqrt{T}\right) + O(\Delta T).$$

This implies that the Monte Carlo error simulations can be controlled by $l^{-1/2}$ uniformly with respect to T and m. Therefore we have the following error estimate:

$$\sup_{x\in\mathbb{R}}|\widetilde{P}_{\theta_0}^{l,m,n,T}(\sqrt{m}(\hat{\theta}_{l,m,n,T} - \theta) \leq x) - P_{\theta_0}^{n,T}(\sqrt{T}(\hat{\theta}_{n,T} - \theta) \leq x)| \leq C(\Delta T)^{1/3} + Cl^{-1/2}$$

where l is the number of trials of Monte Carlo simulations. □

REFERENCES

[1] Altmeyer, R., Bretschneider, T., Janak, J. and Reiss, M. (2021) : Nonparametric estimation for linear SPDEs from local measurements, *Annals of Applied Probability* **31** (1), 1-38.

[2] Altmeyer, R., Bretschneider, T., Janak, J. and Reiss, M. (2022) : Parameter estimation in an SPDE model for cell repolarization, *SIAM Journal on Uncertainty Quantification* **10** (1), 179-199.

[3] Altmeyer, R. and Chorowski, J. (2018) : Estimation error for occupation functionals of stationary Markov processes, *Stoch. Process. Appl.* **128** (6), 1830-1848.

[4] Altmeyer, R., Cialenko, I., Pasemann, G. (2023) : Parameter estimation for semilinear SPDEs from local measurements, *Bernoulli* **29** (3), 2035-2061.

[5] Bishwal, J.P.N. (1999) : Bayes and sequential estimation in Hilbert space valued stochastic differential equations, *J. Korean Statistical Society* **28** (1), 93-106.

[6] Bishwal, J.P.N. (2001) : The Bernstein-von Mises theorem and spectral asymptotics of Bayes estimators for parabolic SPDEs, *J. Australian Math. Society* **72** (2), 289-300.

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

[8] Bishwal, J.P.N. (2017) : Hypothesis testing for fractional stochastic partial differential equations with applications to neurophysiology and finance, *Asian Research Journal of Mathematics* **4** (1), 1-24.

[9] Bishwal, J.P.N. (2018) : Bernstein-von Mises theorem and small noise Bayesian asymptotics for parabolic stochastic partial differential equations, *Theory of Stochastic Processes* **23** (1), 6-17.

[10] Bishwal, J.P.N. (2021) : Statistics of SPDEs: From linear to nonlinear, *European Journal of Statistics* **1** (1), 1-57.

[11] Bishwal, J.P.N. (2022a) : Mixingale estimation function for SPDEs with random sampling, *European Journal of Statistics* **2** (1), 1-13.

[12] Bishwal, J.P.N. (2022b) : *Parameter Estimation in Stochastic Volatility Models*, Springer Nature, Cham.

[13] Cialenko, I., Delgado-Vences, F., Kim, H.J. (2020) : Drift estimation for discretely sampled SPDEs, *Stochastic PDEs: Analysis and Computations*, **8**, 895-920.

[14] Cialenko, I., Kim, H.J., Pasemann, G. (2023) : Statistical analysis of discretely sampled semilinear SPDEs: a power varion approach, *Stochastics and Partial Differential Equations: Analysis and Computations*, DOI:10.1007/s40072-022-00285-3.

[15] Clairon, Q. and Samson, A. (2022) : Optimal control for parameter estimation in partially observed hypoelliptic stochastic differential equations, *Computational Statistics* **37**, 2471-2491.

[16] Da Prato, G. and Zabczyk, J. (2014) : *Stochastic Equations in Infinite Dimensions*, Second Ed., Cambridge University Press.

[17] Duncan, T.E., Maslowski, B. and Pasik-Duncan, B. (2000) : Adaptive control for semilinear stochastic systems, *SIAM J. Control Optim.* **38**, 1683-1706.

[18] Glasserman, P. (2004) : *Monte Carlo Methods in Financial Engineering,* Springer, New York.

[19] Goldys, B. and Maslowski, B. (2002) : Parameter estimation for controlled semilinear stochastic systems, *J. Multimariate Analysis* **80**, 322-343.

[20] Hausenblas, E. (2002) : Numerical analysis of semilinear stochastic evolution equation in Banach spaces, *J. Comp. App. Math.* **147**, 485-516.

[21] Hausenblas, E. (2003) : Approximation of semilinear stochastic evolution equation, *Potential Analysis* **18**, 141-186.

[22] Koski, T. and Loges, W. (1985) : Asymptotic statistical inference for a stochastic heat flow problem, *Statist. Prob. Letters* **3**, 185-189.

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

[24] Kriz, P. and Maslowski, B. (2019) : Central limit theorems and minimum-contrast estimators for linear stochastic evoution equations, *Stochastics* **91** (8), 1109-1140.

[25] Le Breton, A. (1976) : On continuous and discrete sampling for parameter estimation in diffusion type processes, *Mathematical Programming Study* **5**,124-144.

[26] Loges, W. (1984) : Girsanov theorem in Hilbert space and an application to the statistics of Hilbert space-valued stochastic differential equations, *Stoch. Process. Appl.* **17**, 243-263.

[27] Oksendal, B. and Sulem, A. (2007) : *Applied Stochastic Control of Jump Diffusions*, Second Edition, Springer-Verlag, Berlin, Hieldelberg.

[28] Pasemann, G., Flemming, S., Alonso, S., Beta, C. and Stannat, W. (2021) : Diffusivity estimation for activator-inhibitor models: theory and application to the intracellular dynamics of the actin cytoskeleton, *Journal of Nonlinear Science* **31** (59), 1-34.

[29] Pasemann, G. and Stannat, W. (2020) : Drift estimation for stochastic reaction-diffusion systems, *Electronic Journal of Statistics* **14**, 547-579.

[30] Peszat, S. and Zabczyk, J. (2007) : *Stochastic Partial Differential Equations with Levy Noise: Evolution Equations Approach*, Cambridge University Press, Cambridge, England.

[31] Stannat, W. and Wessels, L. (2022) : Necessary and sufficient conditions for optimal control of semilinear stochastic partial differential equations, *ArXiv:2112.09639v3*

[32] Walsh, J.B. (1986) : *An Introduction to Stochastic Partial Differential Equations*, Lecture Notes in Mathematics **1180**, 265-439, Springer-Verlag, Berlin.