LECTURE 12 PARAMETER ESTIMATION FOR STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS

1. PARAMETER ESTIMATION FOR SODES

Consider the estimation of the parameter $\theta$ in the following scalar equation

\begin{equation}
    dX(t) = \theta f(X(t)) \, dt + \sigma b(X(t)) \, dw(t),
\end{equation}

where $w(t)$ is a standard Brownian motion and $a(\cdot), b(\cdot)$ are suitable real-valued functions such that a strong solution $X$ is well-posed.

Usually, the solution $X(t)$ is observed at some realizations $X(t_i, \omega_j), i = 0, 1, 2, \ldots, N$ and $j = 1, 2, \ldots, M$.

**1.1. Estimating $\sigma$ using the quadratic variation argument.** We can apply the quadratic variation of both sides of the equation and obtain

\begin{equation}
    \sigma^2 = \frac{\langle X \rangle_T}{\int_0^T b^2(X(t)) \, dt}.
\end{equation}

Thus with given observations, we obtain

\begin{equation}
    \hat{\sigma}^2 = \frac{1}{M} \sum_{j=1}^M \frac{\sum_{i=0}^{N-1} |X(t_{i+1}, \omega_j) - X(t_i, \omega_j)|^2}{\sum_{i=0}^{N-1} b^2(X(t_i, \omega_j))(t_{i+1} - t_i)}.
\end{equation}

It can be shown that (supposing uniform time step sizes)

\begin{equation}
    \sqrt{N} \left( \frac{\langle X \rangle_T}{\int_0^T b^2(X(t)) \, dt} - \sum_{i=0}^{N-1} |X(t_{i+1}) - X(t_i)|^2 \right) \rightarrow N(0, 2).
\end{equation}

We then apply Slutsky’s theorem to obtain the convergence of this estimator.

We will focus only on the estimation of the parameter $\theta$ as the estimation of $\sigma$ is independent of estimating $\theta$.

**1.2. Estimating $\theta$ using the maximum likelihood method.** Consider the following two scalar diffusion processes driven by the same Brownian motion.

\begin{align*}
    dX &= A(t, X(t)) \, dt + \sigma(t, X(t)) \, dw(t), \quad X(0) = X_0, \\
    dY &= a(t, X(t)) \, dt + \sigma(t, Y(t)) \, dw(t), \quad Y(0) = X_0.
\end{align*}

Here the functions $A, a$ and $\sigma$ satisfy the conditions to ensure existence of a unique strong solution. Assume that the initial conditions are independent of $w(t)$ and $\sigma \geq \sigma_0 > 0$.

Let

\begin{equation}
    B(t, x) = \frac{A(t, x) - a(t, x)}{\sigma(t, x)},
\end{equation}

then by Girsanov’s theorem (e.g., Theorem 8.6.8 of [Øksendal, 2003]),

\begin{equation}
    \tilde{w}(t) = - \int_0^t B(s, Y(s)) \, ds + w(t), \quad 0 \leq t \leq T.
\end{equation}
is a standard Brownian motion under $P_T^Y$, where the measure $P_T^Y$ is defined by

$$P_T^Y(A) = \mathbb{E}^P[1_A \exp(\int_0^T B(t, Y(t)) \, dw - \frac{1}{2} \int_0^T B^2(t, Y(t)) \, dt)].$$

Here $\mathbb{P}$ or $P_T^X$ is the measure generated by the process $X$ on $C([0, T]; \mathbb{R})$. We can write $Y(t)$ as

$$dY = A(t, Y(t)) \, dt + \sigma(t, Y(t)) \, d\tilde{w}(t).$$

Then we may define the likelihood (Radon-Nykodym derivative, see e.g. [Liptser and Shiryaev, 2001])

$$\frac{dP_T^X}{dP_T^T} = \exp(\int_0^T B(t, Y(t)) \, dw - \frac{1}{2} \int_0^T B^2(t, Y(t)) \, dt)$$

(1.10) $$= \exp(\int_0^T \frac{A(t, Y(t)) - a(t, Y(t))}{\sigma^2(t, Y(t))} \, dY(t) - \frac{1}{2} \int_0^T \frac{A^2(t, Y(t)) - a^2(t, Y(t))}{\sigma^2(t, Y(t))} \, dt)$$

(1.11) $$= \exp(\int_0^T \frac{A(t, X(t)) - a(t, X(t))}{\sigma^2(t, X(t))} \, dX(t) - \frac{1}{2} \int_0^T \frac{A^2(t, X(t)) - a^2(t, X(t))}{\sigma^2(t, X(t))} \, dt).$$

For the problem of estimating $\theta$, let $A(t, x) = \theta f(x)$ and $a(t, x) = \theta_0 f(x)$. Then by taking the logarithm of the likelihood and letting the derivative of the log-likelihood be zero, we have

$$\hat{\theta} = \frac{\int_0^T \frac{f(X(t))}{\sigma^2(X(t))} \, dX(t)}{\int_0^T \frac{f^2(X(t))}{\sigma^2(X(t))} \, dt}.$$  

(1.12)

Multiplying $f(X)/\sigma^2(X)$ over both sides of the equation of $X$, we obtain that

$$\theta = \frac{\int_0^T \frac{f(X(t))}{\sigma^2(X(t))} \, dX(t) + \int_0^T \frac{f(X(t))}{\sigma(X(t))} \, dw(t)}{\int_0^T \frac{f^2(X(t))}{\sigma^2(X(t))} \, dt}.$$  

(1.13)

Thus, we have

$$\hat{\theta} - \theta = \frac{\int_0^T \frac{f(X(t))}{\sigma(X(t))} \, dw(t)}{\int_0^T \frac{f^2(X(t))}{\sigma^2(X(t))} \, dt}.$$  

(1.14)

Under certain conditions on $f$ and $\sigma$, we may obtain that $\hat{\theta} - \theta$ converges in distribution to a normal random variable with zero mean and a certain variance. For convergence and its rate, we will defer discussions to the next section, where we will use the Ornstein–Uhlenbeck process as an example.

1.3. **Estimating $\theta$ using the Gaussian (or nonlinear) filter.** The parameter $\theta$ can be also estimated using the filtering technique.

**Example 1.1 (Parameter estimation).**

$$dY = \theta G(t) \, dt + D(t) \, dB(t), Y(0) = 0,$$

Suppose that we have $Y(t)$, $G(t)$, $D(t)$, we want to find out what $\theta$ is.

---

1Here we then have a weak solution $(Y(t), \tilde{w}(t))$ to Equation (1.1).
Solution. Here \(d\theta = 0\) (state model) and \(Y\) is the observation. Then from Lecture 5, we have
\[
\frac{dS}{dt} = -\frac{G^2(t)}{D^2(t)} S(t),
\]
and
\[
\hat{\theta}_t = \mathbb{E}[\theta|F_t^Y] = \frac{\hat{\theta}_0 S_0^{-1} + \int_0^t G(s) D^{-2}(s) dY(s)}{S_0^{-1} + \int_0^t G^2(s) D^{-2}(s) ds}.
\]
See Lecture 5 for the derivation of this formula.

**Remark 1.2.** Let \(S_0^{-1} = 0\) gives the maximum likelihood estimation.

1.4. **A least-square method.** We may also consider the following problem with multiple parameters
\[
\frac{dX}{dt} = \sum_{j=1}^N \theta_j f_j(x) \, dt + \sum_{j=1}^N \sqrt{\theta_j} g_j(x) \, dw_j(t), \quad \theta_j > 0. \quad X(0) = x.
\]
Suppose that \(\phi \in C_b^2(\mathbb{R}^d)\), then we have from Ito’s formula that
\[
\mathbb{E}[\phi(X(t))] - \mathbb{E}[\phi(X(0))] = \int_0^t \mathcal{L}\phi(X(s)) \, ds,
\]
where the operator \(\mathcal{L}\) is the generator of the process \(X(t)\) and
\[
\mathcal{L}v = \sum_{j=1}^n \theta_j f_j^\top \nabla v + \frac{1}{2} \text{Tr}(\sum_{j=1}^n \theta_j g_j g_j^\top H_X v).
\]
Denote \(\mathcal{L}_j = f_j^\top \nabla v + \frac{1}{2} \text{Tr}(g_j g_j^\top \nabla v)\). Then
\[
\sum_{j=1}^n \theta_j \int_0^t \mathcal{L}_j \phi(X(s)) \, ds = \mathbb{E}[\phi(X(t))] - \mathbb{E}[\phi(x)].
\]
This equation is under-determined. If we can obtain samples for various \(x\), e.g., \(x_i \in \mathbb{R}^d, i = 1, 2, \ldots, M\). We then obtain an equation of the form \(A\theta = b\), where \(A_{i,j} = \int_0^t \mathcal{L}_j \phi(X_{0,x_i}(s)) \, ds\) and \(b_i = \mathbb{E}[\phi(X_{0,x_i}(t))] - \mathbb{E}[\phi(x_i)]\). We can then obtain a least-square solution \(\hat{\theta} = A^+ b\), where \(A^+\) is the pseudoinverse of \(A\).

2. **Parameter estimation for bilinear SPDEs**

**Example 2.1** (Stochastic heat equation driven by additive noise). Consider the following evolution equation
\[
du(t, x) - \theta \Delta u(t, x) \, dt = \sigma \, dW^Q(t, x), \quad t > 0, \quad x \in \mathbb{R}^d
\]
with zero initial condition \(u(0, x) = 0\), and \(\Delta u = \sum_{k=1}^d \partial^2_{x_i} u\) denotes the Laplace operator, \(\theta, \sigma \in \mathbb{R}_+\). Here \(G\) is a smooth bounded domain in \(\mathbb{R}^d\) and we consider zero Dirichlet boundary condition.

If \(-\Delta\) and \(Q\) have only point spectrum and the same set of eigenfunctions (in this case, the equation is called a fully diagonalizable equation), say \(Qe_k = q_k e_k\) and \(-\Delta e_k = \nu_k e_k\), then we have
\[
du_k(t) + \theta \nu_k u_k \, dt = \sigma \sqrt{q_k} du_k(t), \quad k \geq 1.
\]
Here \(u_k = (u, e_k)\) and \(\{e_k\}\) is a complete orthonormal system in \(L^2(G)\).
From the discussion in the last section, the MLE gives
\begin{equation}
\hat{\theta}_N^k = -\frac{\sigma \int_0^T \nu_k^\alpha q_k^\rho u_k(t) \, dt}{\int_0^T \nu_k^{1+\alpha} q_k^\rho u_k^2(t) \, dt} = \theta - \frac{\sigma \int_0^T \nu_k^\alpha q_k^{\rho+1/2} u_k(t) \, dt}{\int_0^T \nu_k^{1+\alpha} q_k^\rho u_k^2(t) \, dt},
\end{equation}
where \(\alpha = 1\) and \(\rho = -1\). But the derivation of an estimation of \(\theta\) (1.13) shows that we can take arbitrary \(\alpha \in \mathbb{R}\) and \(\rho \in \mathbb{R}\).

As we have an explicit form of the difference \(\hat{\theta}_N^k - \theta\)
\begin{equation}
\hat{\theta}_N^k - \theta = -\frac{\sigma \int_0^T \nu_k^\alpha q_k^{\rho+1/2} u_k(t) \, dt}{\int_0^T \nu_k^{1+\alpha} q_k^\rho u_k^2(t) \, dt},
\end{equation}
we may work out the convergence and its rate which is stated in the following theorem; see details of calculation in Chapter 6.1 of [Lototsky and Rozovsky, 2017].

**Theorem 2.2.** Assume \(q_k \equiv 1\) and \(\alpha = 1\) and \(\rho = -1\). For each \(k \geq 1\), the estimator \(\theta_N^k\) is asymptotically normal with rate \(\sqrt{T}\).
\begin{equation}
\lim_{T \to \infty} \sqrt{T}(\hat{\theta}_N^k(T) - \theta) \overset{d}{=} N(0, \frac{2\theta}{k^2}).
\end{equation}
For each \(T > 0\),
\begin{equation}
\lim_{k \to \infty} k(\hat{\theta}_N^k(T) - \theta) \overset{d}{=} N(0, \frac{2\theta}{T}).
\end{equation}

**Exercise 2.3.** Show the convergence of \(\hat{\theta}_N^k\) to \(\theta\) is also in probability.

We need \((\hat{\theta}_N^k)^+ = \hat{\theta}_N^k \vee 0\) to ensure that the estimator of \(\theta\) is positive.

Above we use the information of one single frequency \(k\), averaging over \(k = 1, 2, \ldots, N\) leads to another estimation and we expect higher-order convergence of the averaging as \(N\) modes of solutions have more information. Denote
\begin{equation}
\hat{\theta}_N = \sum_{k=1}^n \frac{a_k}{\sum_{k=1}^n a_k} \hat{\theta}_N^k, \quad \hat{\theta}_N^k = \frac{b_k}{a_k}, \quad a_k = \int_0^T \nu_k^{1+\alpha} q_k^\rho u_k^2(t) \, dt, \quad b_k = -\sigma \int_0^T \nu_k^\alpha q_k^\rho u_k(t) \, dt.
\end{equation}
The estimator \(\hat{\theta}_N\) can be written as
\begin{equation}
\hat{\theta}_N = -\frac{\sigma \sum_{k=1}^N \int_0^T \nu_k^\alpha q_k^\rho u_k(t) \, dt}{\sum_{k=1}^N \int_0^T \nu_k^{1+\alpha} q_k^\rho u_k^2(t) \, dt} = \theta - \frac{\sigma \sum_{k=1}^N \int_0^T \nu_k^\alpha q_k^{\rho+1/2} u_k(t) \, dt}{\sum_{k=1}^N \int_0^T \nu_k^{1+\alpha} q_k^\rho u_k^2(t) \, dt}.
\end{equation}

When \(\alpha = 1\), \(\rho = -1\), \(q_k \equiv 1\), the proof convergence is sketched as follows. By the ergodicity of the process \(u_k\),
\begin{equation}
\frac{1}{T} \int_0^T \sum_{k=1}^N q_k^{-1} \nu_k^2 u_k^2(t) \, dt \to \sum_{k=1}^N q_k^{-1} \frac{\sigma^2}{2\mu_k(\theta)}, \quad T \to \infty.
\end{equation}
We can also show that \(\sigma \sum_{k=1}^N \int_0^T \nu_k q_k^{-1} u_k(t) \, dt\) converges in distribution to a normal random variable with mean zero (see Chapter 6.3 of [Lototsky and Rozovsky, 2017] for a proof) using the martingale central limit theorem A.2 in Appendix.

**Theorem 2.4.** Assume that, \(\nu_k \to \infty\), \(q_k \to 0\), and \(\alpha, \rho\) are such that
\begin{equation}
\nu_n^{\alpha-1} q_n^{2\rho+2} \leq M, \quad n \geq 1, \quad \sum_{k \geq 1} \nu_k^\alpha q_k^{2\rho+2} = \infty,
\end{equation}
for some \( M \in \mathbb{R} \). Then, \( \hat{\theta}_N \) is a consistent estimator of \( \theta \), i.e. \( \hat{\theta}_N \to \theta \) with probability one. Moreover, if
\[
\sum_{k \geq 1} \nu_k^{2\alpha-1} q_k^{4\rho+4} = \infty,
\]
then, \( \hat{\theta}_N \) is also asymptotically normal
\[
\frac{\sum_{k=0}^{N} \nu_k^\alpha q_k^{2\rho+2}}{\sqrt{\sum_{k=0}^{N} \nu_k^{2\alpha-1} q_k^{4\rho+4}}} (\hat{\theta}_N - \theta) \xrightarrow{d_{N \to \infty}} \mathcal{N}(0, \frac{2\theta}{T}).
\]

When \( q_k \equiv 1 \) and \( \nu_k = k^2 \) \((G = (0, \pi) \subseteq \mathbb{R}^1)\), then the convergence rate is
\[
\sqrt{\sum_{k=1}^{N} \nu_k^{2\alpha-1}} \approx N^{2(2\alpha-1)+1/2}N^{-(2\alpha+1)} = N^{-\frac{3}{2}}.
\]

**Exercise 2.5.** What is the convergence rate for \( \theta_N \) when \( T \to \infty \)?

**Example 2.6** (Stochastic heat equation driven by additive noise).
\[
\begin{aligned}
\frac{d}{dt}\left(u(t,x) - (A_0 u(t,x) + \theta A_1 u(t,x))\right) &= \sigma dW^Q(t,x), \quad (t,x) \in (0,T] \times (0,\pi),
\end{aligned}
\]
with zero initial and boundary conditions, and \( \theta \in \mathbb{R} \) and \( \sigma \in \mathbb{R}^+ \).

The procedure of finding \( \theta \) is similar to the one we use in the last example and the estimator is
\[
\hat{\theta}_N = -\frac{\int_0^T (Q^{-1} A_1 u^N, du_N) - \int_0^T (Q^{-1} A_1 u^N, A_0 u^N), dt}{\int_0^T \left\| Q^{-1/2} A_1 u^N \right\|^2 dt},
\]
while
\[
\begin{aligned}
\theta &= -\frac{\int_0^T (Q^{-1} A_1 u^N, du_N) - \int_0^T (Q^{-1} A_1 u^N, A_0 u^N), dt + \sigma \int_0^T Q^{-1} A_1 u^N dW^Q(t)}{\int_0^T \left\| Q^{-1/2} A_1 u^N \right\|^2 dt}.
\end{aligned}
\]

Here \( u^N = P^N u = \sum_{k=1}^{N} u_k e_k \), where \( e_k \)'s are the common eigenfunctions of \( A_0, A_1 \) and the operator \( Q \). If we set \( A_0 e_k = \rho_k e_k \) \((\rho_k > 0)\), \( A_1 e_k = \nu_k e_k \) \((\nu_k > 0)\) and \( Q e_k = q_k e_k \), we then have
\[
\hat{\theta}_N - \theta = \frac{\sigma \int_0^T Q^{-1} A_1 u^N dW^Q(t)}{\int_0^T \left\| Q^{-1/2} A_1 u^N \right\|^2 dt} - \frac{\sigma \sum_{k=1}^{N} \int_0^T \nu_k q_k^{-1} u_k dw_k(t)}{\int_0^T \sum_{k=1}^{N} q_k^{-1} \nu_k q_k^{-1} u_k^2 dt}.
\]

Denote \( I_{N,\theta} = \sigma^2 \sum_{k=1}^{N} \frac{\nu_k^2}{\mu_k(\theta)} \), where \( \mu_k(\theta) = \nu_k + \theta \rho_k \).

**Theorem 2.7.** Let \( q_k \equiv 1 \). Assume that \( \theta \in \mathbb{R} \) and \( u(0,x) = 0 \). Assume also that \( \mu_k(\theta) \to +\infty \) when \( k \to \infty \) and \( \sup_k \frac{|\nu_k|}{\mu_k(\theta)} < +\infty \). When \( \lim_{N \to \infty} I_{N,\theta} = +\infty \), then \( \lim_{N \to \infty} \hat{\theta}_N = \theta \) with probability one.
\[
\lim_{N \to \infty} \sqrt{I_{N,\theta}(\hat{\theta}_N - \theta)} \xrightarrow{d} \mathcal{N}(0, \frac{2}{T}).
\]
When $\lim_{N \to \infty} I_{N, \theta} < +\infty$,

$$\lim_{N \to \infty} \hat{\theta}_N = \frac{\sum_{k=1}^{\infty} \nu_k q_k^{-1} \int_0^T u_k \, dw_k(t)}{\sum_{k=1}^{\infty} \int_0^T q_k^{-1} \nu_k^2 u_k^2 \, dt}.$$ (2.18)

The proof of this theorem can be found at Chapter 6.3 of [Lototsky and Rozovsky, 2017].

2.1. General theory. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ satisfy the usual conditions and let $H$ be a separable Hilbert space equipped with the inner product $(\cdot, \cdot)$ and the norm $\|\cdot\|$. Let’s consider the following evolution equation

$$du(t) + (\theta A + B) u(t) \, dt = (Mu(t) + \sigma) \, dW^Q(t),$$

with the initial condition $u(0) = u_0 \in H$, and $\sigma, \theta > 0$ and

- the leading operator $A$ is a linear, positive defined, self-adjoint operator in $H$;
- $B$ is a linear or nonlinear operator in $H$;
- $M$ is an operator acting in a suitable Hilbert space;
- $W^Q$ is a $Q$-cylindrical Brownian motion in $H$.

We consider diagonalizable SPDEs (not key assumption but it does simplifies problems)

- the operators $A$ and $Q$ have pure point spectrum, and a common system of eigenfunctions $\{e_k\}_{k \geq 1}$ that forms a complete orthonormal system in $H$.

Thus

$$\hat{W}^Q(t) = \sum_{k=1}^{\infty} \sqrt{q_k} e_k \hat{w}_k(t),$$

where $w_k$ are independent standard Brownian motions.

Estimating $\sigma$ for SPDEs is essentially the same as for SODEs. We will focus on estimating the parameter $\theta \in \Theta \subset \mathbb{R}^+$ and assume that the positive constant $\sigma$ is known.

We apply the projection operator $\mathcal{P}^N$ to (2.19), and obtain

$$du^N + (\theta A^N(t) + \Psi^N(t)) \, dt = \sigma dW^{Q,N}(t), \quad u^N(0) = \mathcal{P}^N u_0,$$ (2.20)

where $\Psi^N = \mathcal{P}^N B(u)$, and $W^{Q,N} = \mathcal{P}^N W^Q$. and $\mathcal{P}^N$ is the projection operator from $H$ onto $H^N = \text{span} \{e_k : k = 1, \ldots, N\}$.

Denote $u_N = \mathcal{P}^N u = \sum_{k=1}^{N} u_k e_k$, where $u_k = (u, e_k), \; k \geq 1$, are the Fourier coefficients (or modes) of the solution $u$ with respect to $\{e_k\}_{k \geq 1}$. Denote by $\mathbb{P}^T_\theta$ the probability measure on $C([0, T]; H^N)$ generated by the solution $u_N^T$ of (2.20). We assume that the family of measures $\{\mathbb{P}^T_\theta(\cdot)\}_{\theta \in \Theta}$ are mutually absolutely continuous (without justification, no need actually). Using the Radon-Nykodym derivative or Likelihood Ratio

$$\frac{d\mathbb{P}^T_\theta}{d\mathbb{P}^T_{\theta_0}} (u_N) = \exp \left( - (\theta - \theta_0) \sigma^{-2} \int_0^T (Q^{-1} A u^N, du^N) - \frac{1}{2} (\theta^2 + \theta_0^2) \sigma^{-2} \int_0^T \|Q^{-1/2} A u^N\|^2 \, dt \right)$$

(2.21)

We can formally compute the maximum likelihood estimator (MLE) for the parameter of interest $\theta$ by maximizing the log-likelihood ratio $\log(\frac{d\mathbb{P}^N_\theta}{d\mathbb{P}^N_{\theta_0}} (u_N))$ with respect to $\theta$, and
obtain the following estimator
\begin{equation}
\hat{\theta}_N^2 = -\frac{\int_0^T (Q^{-1}Au^N, du_N) + \int_0^T (Q^{-1}Au^N, P^N B(u)) \, dt}{\int_0^T \|Q^{-1/2}Au^N\|^2 \, dt}.
\end{equation}

The above methodology can be applied to the following Navier-Stokes equation with additive noise. For details, see [Cialenco and Glatt-Holtz, 2011] or the review paper [Cialenco, 2018].

**Example 2.8** (2D Navier–Stokes Equations forced with additive noise). Consider the equation that describes the flow of a viscous, incompressible fluid
\begin{align}
du - \theta Du \, dt + (u \cdot \nabla u) u \, dt + \nabla P \, dt &= \sigma \, dW^Q(t), \\
\nabla \cdot u &= 0, \\
u(0) &= u_0,
\end{align}
where \( u \) represent the velocity field and \( P \) the pressure. Here we may consider either zero boundary conditions on a bounded domain in \( \mathbb{R}^2 \), or periodic domain with periodic boundary conditions.

For an estimation of multiple parameters using the MLE, see Chapter 6.3 of [Lototsky and Rozovsky, 2017].

### 3. The p-variation method

Measurements of the solution \( u(t, x) \) are only at some discrete time points \( t_i \) and/or some discrete spatial points \( x_j \), over one path/realization \( \omega \in \Omega \).

Let \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})\) be a stochastic basis satisfying the usual conditions. We consider the following stochastic partial differential equation on \( H = L^2(G) \), where \( G = (0, \pi) \).

\begin{equation}
\begin{cases}
\partial_t u(t, x) = \theta u_{xx}(t, x) + \sigma \tilde{W}(t, x), & x \in G, \ t > 0, \\
u(0, x) = 0,
\end{cases}
\end{equation}

where \( \theta, \sigma \) are some positive constants, and \( \tilde{W}(t, x) \) is a space-time white noise and we also assume zero boundary conditions \( u(t, 0) = u(t, \pi) = 0, \ t > 0 \). Here \( \theta, \sigma \in \mathbb{R}_+ \) are the (unknown) parameters of interest.

Here we assume that the observation is at a fixed time instant and at discrete space locations. Specifically, for a fixed instant of time \( t > 0 \), and given interval \([a, b] \subset G\), the solution \( u \) is observed at points \( (t, x_j), \, j = 1, \ldots, m \), with \( x_j = a + (b - a)j/m, \ j = 0, 1, \ldots, m \).

Introduce the partition \( \mathcal{T}^m(a, b) = \{a_j \mid a_j = a + (b - a)j/m, \ j = 0, 1, \ldots, m\} \) for the uniform partition of size \( m \) of a given interval \([a, b] \subset \mathbb{R}\). Assume that \( t > 0 \) is a fixed time instant, and consider the partition \( \mathcal{T}^m(a, b) \) of the fixed interval \([a, b] \subset G\). Suppose that the solution \( u \) of (3.1) is observed at the grid points \( \{(t, x_j) \mid x_j \in \mathcal{T}^m(a, b), j = 1, \ldots, m\} \).

Consider the following estimators for \( \theta \) and \( \sigma^2 \) respectively
\begin{align}
\tilde{\theta}_{m,t} := -\frac{(b-a)\sigma^2}{2\sum_{j=1}^m(u(t, x_j) - u(t, x_{j-1}))^2}, \text{ (known } \sigma \text{)} ,
\end{align}
\begin{align}
\tilde{\sigma}_{m,t}^2 := \frac{2\theta}{b-a} \sum_{j=1}^m(u(t, x_j) - u(t, x_{j-1}))^2, \text{ (known } \theta \text{)} .
\end{align}

These estimators are consistent and asymptotically normal.
Theorem 3.1. Assume that \( u \) is the solution of (3.1) and that \( \sigma \) is known. The estimator (3.2) of \( \theta \) is (strongly) consistent, i.e., \( \lim_{m \to \infty} \hat{\theta}_{m,t} = \theta \) with probability one, and asymptotically normal,

\[
\sqrt{m}(\hat{\theta}_{m,t} - \theta) \xrightarrow{d} N(0, 2\theta^2).
\]

Assuming that \( \theta \) is known, the estimator (3.3) is a (strongly) consistent and asymptotically normal estimator of \( \sigma^2 \), with

\[
\sqrt{m}(\hat{\sigma}_{m,t}^2 - \sigma^2) \xrightarrow{d} N(0, 2\sigma^4).
\]

Proof. For every fixed \( t > 0 \), there is a Brownian motion \( B(x) \) on \([0, \pi]\), and a Gaussian process \( R(x) \in C^\infty(0, \pi) \) such that

\[
u(t, x) = \frac{\sigma}{\sqrt{2\theta}} B(x) + R(x), \quad x \in [0, \pi].
\]

Indeed, we may take

\[
B(x) = \xi_0 + \sum_{k \geq 1} \frac{1}{k} \xi_k e_k(x), \quad R(x) = -\frac{\sigma x}{\sqrt{2\theta}} \xi_0 + \frac{\sigma}{\sqrt{2\theta}} \sum_{k \geq 1} \frac{a_k - 1}{k} \xi_k e_k(x),
\]

\[
\xi_k = \sqrt{\frac{2\theta k^2}{(1 - e^{-2\theta k^2})\sigma^2}} u_k(t), \quad a_k = \sqrt{1 - e^{-2\theta k^2}}.
\]

Here \( \xi_k \)'s are i.i.d. standard Gaussian random variables. From Lecture 2, we know that \( B \) is a standard Brownian motion on \([0, \pi]\) and that \( R \) is smooth.

We then can calculate the convergence rate explicitly using the representation (3.6) \( \square \)

Remark 3.2 (fixed location and various observations at discrete time instants). Assume that the solution \( u \) is observed at points \( \{(t_i, x), i = 1, \ldots, n\} \), where \( t_i := c + (d - c)i/n, \ i = 0, 1, \ldots, n, \) and \( t_i \in [c, d] \subset (0, +\infty) \) and a fixed \( x \) from the interior of \( G \). For the equation (3.1), we may consider the following estimators for \( \theta \), and \( \sigma^2 \) respectively,

\[
\hat{\theta}_{n,x} := \frac{3(d - c)\sigma^4}{\pi \sum_{i=1}^n (u(t_i, x) - u(t_{i-1}, x))^4}, \ (\text{known } \sigma),
\]

\[
\hat{\sigma}_{n,x}^2 := \sqrt{\frac{\theta \pi}{3(d - c)} \sum_{i=1}^n (u(t_i, x) - u(t_{i-1}, x))^4}, \ (\text{known } \theta).
\]

These estimators are consistent and asymptotically normal, see [Cialenco and Huang, 2020], where the schemes for estimating \( \theta \) and \( \sigma \) concurrently are also discussed.

For estimations based on Bayesian analysis, interested readers are referred to [Cheng et al., 2018].

Appendix A. Central limit theorems

Theorem A.1. Suppose that \( \sigma_k \in L^2(\Omega, L^2([0, T])) \) is a sequence of independent real-valued predictable process and the strong law of large numbers holds

\[
\lim_{N \to \infty} \frac{\sum_{k=1}^N \int_0^T \sigma_k^2 \, dt}{\sum_{k=1}^N \mathbb{E}[\int_0^T \sigma_k^2 \, dt]} = 1, \ a.s.
\]

Then the central limit theorem holds

\[
\frac{\sum_{k=1}^N \int_0^T \sigma_k \, dw_k(t)}{\left(\sum_{k=1}^N \mathbb{E}[\int_0^T \sigma_k^2 \, dt]\right)^{1/2}}
\]
converges in distribution to a standard normal random variable when $N \to \infty$.

The first conclusion is a variant of the strong law of large numbers, see e.g. Theorem 6.2.11 of [Lototsky and Rozovsky, 2017] and the second conclusion can be proved by the following theorem.

**Theorem A.2** (Martingale central limit theorem, Theorem 5.5.4 of [Liptser and Shiryaev, 1989]). Assume that for $t \geq 0$ and $\varepsilon > 0$, $X_\varepsilon(t)$ and $X(t)$ are real-valued continuous square-integrable martingales. Also, $X(t)$ is a Gaussian process with $X(0) = X(t_0) = 0$. If for some $t_0 > 0$,

(A.3) \[
\lim_{\varepsilon \to 0} \langle X_\varepsilon \rangle(t_0) = \langle X \rangle(t_0), \quad \text{in probability},
\]

Then $\lim_{\varepsilon \to 0} X_\varepsilon(t_0) \overset{d}{=} X(t_0)$.

**APPENDIX B. SOME IMPORTANT TOPICS WE DIDN’T COVER**

- Nonlinear problems with non-globally Lipschitz nonlinear terms (we assume that at most Lipschitz nonlinearity for SPDEs, but we do have non-globally Lipschitz nonlinear terms for SODEs)
- Weak convergence, including convergence in probability, convergence in law etc.
- Spectral methods, finite/spectral element methods for space and/or time discretization (we focus on finite difference schemes)
- Numerical methods for multiscale stochastic partial differential equations
- Theory and numerics for stochastic partial differential equations with Levy noises, see e.g. [Peszat and Zabczyk, 2007];
- Inference of a stochastic process (or random field)
- Large Deviations for SPDEs and their computation
- ...

**REFERENCES**


