

Hypothesis testing for Stochastic PDEs

Igor Cialenco

Department of Applied Mathematics
Illinois Institute of Technology

`igor@math.iit.edu`

Joint work with Liaosha Xu

Research partially funded by NSF grants DMS-1211256

Asymptotical Statistics of Stochastic Processes X
Le Mans, 17 - 20 March, 2015

- Overview of Parameter Estimation for SPDEs
- Hypothesis testing
- Error estimation
- Numerical results

Bibliography

I. Cialenco, L. Xu, Hypothesis testing for stochastic PDEs driven by additive noise, *Stochastic Processes and their Applications*, vol. 125, Issue 3, March 2015, pp. 819-866.

I. Cialenco, L. Xu, A note on error estimation for hypothesis testing problems for some linear SPDEs, *Stochastic Partial Differential Equations: Analysis and Computations*, September 2014, vol. 2, No 3, pp. 408-431.

Let $(\cdot, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ be a stochastic basis with usual assumptions. Consider heat equation driven by additive noise:

$$dU(t, x) - \theta U(t, x)dt = \sigma \sum_{k=1}^{\infty} \lambda_k^{-\gamma} h_k(x) dw_k(t), \quad (0.1)$$

where $x \in G$, G is a bounded domain in \mathbb{R}^d , $t \in [0, T]$;

$\theta > 0$ is the parameter of interest; $\gamma \geq 0$, $\sigma \in \mathbb{R} \setminus \{0\}$ are known.

- zero initial conditions and boundary values;
- $\{w_k(t)\}_{k \in \mathbb{N}}$ are independent Brownian motions;
- Δ is the Laplace operator on G with zero boundary condition $U = \sum_{j=1}^d \frac{\partial^2 U}{\partial x_j^2}$;
- $\{h_k\}$ are the eigenfunctions of Δ in $L^2(G)$, and $\{\rho_k\}$ are the corresponding eigenvalues; $\lambda_k = \sqrt{-\rho_k} \sim k^{1/d}$;

Theorem (Existence and Uniqueness)

If $(\gamma - s)/d > 1/2$, then the SPDE (0.1) has a unique solution (weak in PDE sense, and strong in probability sense)

$$U \in L_2(\times [0, T]; H^{s+1}(G)) \cap L^2(\ ; C((0, T); H^s(G))).$$

PART I: Hypothesis Testing for SPDEs

The Problem: Simple Hypothesis

$$dU(t, x) - \theta U(t, x)dt = \sigma \sum_{k=1}^{\infty} \lambda_k^{-\gamma} h_k(x) dw_k(t).$$

Assume that θ can take only two values $\{\theta_0, \theta_1\}$.

Consider a simple hypothesis:

$$\mathcal{H}_0 : \theta = \theta_0,$$

$$\mathcal{H}_1 : \theta = \theta_1.$$

For simplicity, assume $\theta_1 > \theta_0$ and $\sigma > 0$.

Construction of the Test

$$dU(t, x) - \theta U(t, x)dt = \sigma \sum_{k=1}^{\infty} \lambda_k^{-\gamma} h_k(x) dw_k(t), \quad U(0, x) = 0.$$

- The k -th Fourier coefficient $u_k(t) = \langle U(t, x), h_k(x) \rangle$ is given by

$$\begin{aligned} du_k &= -\theta \lambda_k^2 u_k dt + \sigma \lambda_k^{-\gamma} dw_k(t), \quad u_k(0) = 0, \\ u_k(t) &= \sigma \lambda_k^{-\gamma} \int_0^t e^{-\theta \lambda_k^2 (t-s)} dw_k, \quad k \geq 1. \end{aligned}$$

- Let $\mathbb{P}_\theta^{N,T}(\cdot) = \mathbb{P}(U_T^N \in \cdot)$ be the measure on $C([0, T]; \mathbb{R}^N)$ generated by $U_T^N(t) = (u_1, \dots, u_N)$ up to time T .

Observable: First N Fourier coefficients, for all $t \in [0, T]$,

$$U_T^N = ((u_1(t), \dots, u_N(t)), t \in [0, T]).$$

In this case, the Likelihood Ratio has the form

$$L(\theta_0, \theta_1, U_T^N) = \frac{\mathbb{P}_{\theta_1}^{N,T}}{\mathbb{P}_{\theta_0}^{N,T}} = \exp \left(-(\theta_1 - \theta_0) \sigma^{-2} \sum_{k=1}^N \lambda_k^{2+2\gamma} \right. \\ \left. \times \left(\int_0^T u_k(t) du_k(t) + \frac{1}{2} (\theta_1 + \theta_0) \lambda_k^2 \int_0^T u_k^2(t) dt \right) \right),$$

and the Maximum Likelihood Estimator is

$$\widehat{\theta}_T^N = - \frac{\sum_{k=1}^N \lambda_k^{2+2\gamma} \int_0^T u_k(t) du_k(t)}{\sum_{k=1}^N \lambda_k^{4+2\gamma} \int_0^T u_k^2(t) dt}, \quad N \in \mathbb{N}, T > 0.$$

Two Asymptotic Regimes:

Large times, $T \rightarrow \infty$;

Large number of Fourier modes, $N \rightarrow \infty$.

Theorem

Assume that $2\gamma > d$. Then,

$$\lim_{N \rightarrow \infty} \widehat{\theta}_T^N = \lim_{T \rightarrow \infty} \widehat{\theta}_T^N = \theta, \quad a.e.$$

and

$$\lim_{N \rightarrow \infty} N^{1/d + \frac{1}{2}} (\widehat{\theta}_T^N - \theta) \stackrel{d}{=} \mathcal{N}\left(0, \frac{(4/d + 2)\theta}{\varpi \sigma^2 T}\right).$$

$$\lim_{T \rightarrow \infty} \sqrt{T} (\widehat{\theta}_T^N - \theta) \stackrel{d}{=} \mathcal{N}(0, 2\theta/M),$$

where $M = \sum_{k=1}^N \lambda_k^2$.

- Looking for rejection region $R \in \mathcal{B}(C([0, T]; \mathbb{R}^N))$.
- **Type I error** = $\mathbb{P}_{\theta_0}^{N, T}(R)$;
- **Type II error** = $1 - \mathbb{P}_{\theta_1}^{N, T}(R)$, and **power of the test** = $\mathbb{P}_{\theta_1}^{N, T}(R)$
- Define the class of test

$$\mathcal{K}_\alpha := \left\{ R \in \mathcal{B}(C([0, T]; \mathbb{R}^N)) : \mathbb{P}_{\theta_0}^{N, T}(R) \leq \alpha \right\}.$$

with $\alpha \in (0, 1)$ being the significant level, fixed in what follows.

Definition

We say that a rejection region $R^* \in \mathcal{K}_\alpha$ is **the most powerful in the class** \mathcal{K}_α if

$$\mathbb{P}_{\theta_1}^{N, T}(R) \leq \mathbb{P}_{\theta_1}^{N, T}(R^*), \quad \text{for all } R \in \mathcal{K}_\alpha.$$

Neyman-Pearson Lemma

Theorem (C. and Xu, '14)

Let c_α be a real number such that

$$\mathbb{P}_{\theta_0}^{N,T}(L(\theta_0, \theta_1, U_T^N) \geq c_\alpha) = \alpha.$$

Then,

$$R^* := \{U_T^N : L(\theta_0, \theta_1, U_T^N) \geq c_\alpha\},$$

is the most powerful rejection region in the class \mathcal{K}_α .

The Difficulty: c_α has no explicit formula for finite T and N .

The Difficulty: c_α has no explicit formula for finite T and N .

We take/suggest an **“Asymptotic Method”**

- (1) Fix N , let $T \rightarrow \infty$;
- (2) Fix T , let $N \rightarrow \infty$.

Inspired by large times Asymptotic Method for finite dimensional ergodic diffusion processes developed by Kutoyants '74, and '04.

In this talk we focus on case (2), asymptotics in number of Fourier modes;

For case (1) see [CX'14].

Asymptotic Method in Number of Fourier modes N

Define a new class

$$\tilde{\mathcal{K}}_\alpha := \left\{ (R_N)_{N \in \mathbb{N}} : R_N \in \mathcal{B}(C([0, T]; \mathbb{R}^N), \limsup_{N \rightarrow \infty} \mathbb{P}_{\theta_0}^{N, T}(R_N) \leq \alpha \right\},$$

where T is fixed, and α is the **“Asymptotic Significant Level”**.

Asymptotic Method in Number of Fourier modes N

Define a new class

$$\tilde{\mathcal{K}}_\alpha := \left\{ (R_N)_{N \in \mathbb{N}} : R_N \in \mathcal{B}(C([0, T]; \mathbb{R}^N)), \limsup_{N \rightarrow \infty} \mathbb{P}_{\theta_0}^{N, T}(R_N) \leq \alpha \right\},$$

where T is fixed, and α is the “**Asymptotic Significant Level**”.

Goal:

We want to find a rejection region $(\tilde{R}_N)_{N \in \mathbb{N}}$ such that

$$\lim_{N \rightarrow \infty} \mathbb{P}_{\theta_0}^{N, T}(\tilde{R}_N) = \alpha.$$

Attempt:

We still try Likelihood Ratio test. Then, what is c_α ?

$$\begin{aligned} & \mathbb{P}_{\theta_0}^{N,T}(L(\theta_0, \theta_1, U_T^N) \geq \tilde{c}_\alpha(N)) \\ &= \mathbb{P}_{\theta_0}^{N,T} \left(X^N - Y^N \geq \frac{\sqrt{8\theta_0^3} \ln \tilde{c}_\alpha(N)}{\sqrt{TM}(\theta_1^2 - \theta_0^2)} + \frac{\sqrt{2\theta_0 TM}(\theta_1 - \theta_0)}{2(\theta_1 + \theta_0)} - \frac{(\theta_1 - \theta_0)N}{\sqrt{8\theta_0 TM}(\theta_1 + \theta_0)} \right), \end{aligned}$$

where

$$X^N := \frac{\sqrt{\theta_0}(\theta_1 - \theta_0)}{\sigma^2 \sqrt{2TM}(\theta_1 + \theta_0)} \left(\sum_{k=1}^N \lambda_k^{2+2\gamma} u_k^2(T) - \frac{\sigma^2}{2\theta_0} N \right),$$

$$Y^N := \frac{\sqrt{2\theta_0}}{\sigma \sqrt{TM}} \sum_{k=1}^N \lambda_k^{2+\gamma} \int_0^T u_k dw_k,$$

$$M := \sum_{k=1}^N \lambda_k^2.$$

For $\delta \in \mathbb{R}$, by splitting argument, we get

$$\mathbb{P}_{\theta_0}^{N,T}(L(\theta_0, \theta_1, U_T^N) \geq \tilde{c}_\alpha(N)) \leq \mathbb{P}_{\theta_0}^{N,T}(X^N \geq \delta) + \mathbb{P}_{\theta_0}^{N,T}\left(-Y^N \geq \frac{\sqrt{8\theta_0^3} \ln \tilde{c}_\alpha(N)}{\sqrt{TM}(\theta_1^2 - \theta_0^2)} + \frac{\sqrt{2\theta_0 TM}(\theta_1 - \theta_0)}{2(\theta_1 + \theta_0)} - \frac{(\theta_1 - \theta_0)N}{\sqrt{8\theta_0 TM}(\theta_1 + \theta_0)} - \delta\right).$$

Take $\delta = N^{-1/2d}$, and deduce that $\mathbb{P}_{\theta_0}^{N,T}(X^N \geq \delta) \rightarrow 0$, $N \rightarrow \infty$.

Use the fact that $Y^N \xrightarrow{d} \mathcal{N}(0, 1)$ as $N \rightarrow \infty$, and find

Reasonable candidate for the threshold constant

$$\hat{c}_\alpha(N) = \exp\left(-\frac{(\theta_1 - \theta_0)^2 TM}{4\theta_0} + \frac{(\theta_1 - \theta_0)^2 N}{8\theta_0^2} - \frac{\sqrt{TM}(\theta_1^2 - \theta_0^2)}{\sqrt{8\theta_0^3}} q_\alpha\right).$$

Theorem (C. and Xu)

Suppose

$$\widehat{R}_N := \{U_T^N : L(\theta_0, \theta_1, U_T^N) \geq \widehat{c}_\alpha(N)\}, \quad \text{for all } N,$$

where \widehat{c}_α is given above.

Then, the rejection region $(\widehat{R}_N)_{N \in \mathbb{N}} \in \widetilde{\mathcal{K}}_\alpha$, and moreover

$$\lim_{N \rightarrow \infty} \mathbb{P}_{\theta_0}^{N,T}(\widehat{R}_N) = \alpha.$$

The Next Question:

Is $(\widehat{R}_N)_{N \in \mathbb{N}}$ (asymptotically) the 'most powerful' test in $\widetilde{\mathcal{K}}_\alpha$?

The Next Question:

Is $(\widehat{R}_N)_{N \in \mathbb{N}}$ (asymptotically) the 'most powerful' test in $\widetilde{\mathcal{K}}_\alpha$?

Short answer: No!

The Next Question:

Is $(\widehat{R}_N)_{N \in \mathbb{N}}$ (asymptotically) the 'most powerful' test in $\widetilde{\mathcal{K}}_\alpha$?

Short answer: No!

How does the power of this test $\mathbb{P}_{\theta_1}^{N,T}(\widehat{R}_N)$ behave?

Definition

We say that a rejection region $(R_N^*) \in \mathcal{K}_\alpha$ is **asymptotically the most powerful** in the class \mathcal{K}_α if

$$\liminf_{N \rightarrow \infty} \frac{1 - \mathbb{P}_{\theta_1}^{N,T}(R_N)}{1 - \mathbb{P}_{\theta_1}^{N,T}(R_N^*)} \geq 1, \quad \text{for all } (R_N) \in \mathcal{K}_\alpha^*.$$

Similarly, we define asymptotically the most powerful rejection regions for a different given class of tests.

Theorem (C. and Xu)

(Criterion for Most Powerful Test)

Consider the rejection region of the form

$$R_N^* = \{U_T^N : L(\theta_0, \theta_1, U_T^N) \geq c_\alpha^*(N)\},$$

where $c_\alpha^*(N)$ is a function of N such that, $c_\alpha^*(N) > 0$ for all $N \in \mathbb{N}$, and

$$\lim_{N \rightarrow \infty} \mathbb{P}_{\theta_0}^{N,T}(R_N^*) = \alpha,$$

$$\lim_{N \rightarrow \infty} \frac{c_\alpha^*(N)}{1 - \mathbb{P}_{\theta_1}^{N,T}(R_N^*)} < \infty.$$

Then (R_N^*) is asymptotically the most powerful in \mathcal{K}_α^* .

Is the test \widehat{R}_N asymptotically the most powerful in the class $\widetilde{\mathcal{K}}_\alpha$?

We need to control the power of the test $\mathbb{P}_{\theta_1}^{N,T}(\widehat{R}_N)$, as $N \rightarrow \infty$.

- Calculate the Cumulant Generating Function of the Log-Likelihood ratio $m(\epsilon) = \mathbb{E}[\exp(\epsilon \ln L(\theta_0, \theta_1, U_T^N))]$

$$m(\epsilon) = \exp \left[-\frac{1}{2} \sum_{k=1}^N \ln(\cosh(\gamma_k T) - p \sinh(\gamma_k T)) + \frac{\epsilon(\theta_1 - \theta_0) + \theta_1}{2} MT \right].$$

Similar to Gapeev and Kuchler [2008]

- Use Feynman-Kac Formula to derive a PDE
- Make some transforms and guess the solution

- Sharp Large Deviations (see Lin'kov 1999, and Bercu and Roualt 2001 for $T \rightarrow \infty$)

$$\mathbb{L}_T(\epsilon) = T^{-1} \ln \mathbb{E}_{\theta_1} \exp \epsilon \ln L(\theta_0, \theta_1, U_T^N) = \mathbb{L}(\epsilon) + T^{-1} \mathbb{H}(\epsilon) + T^{-1} \mathbb{R}_T(\epsilon),$$

$$\mathbb{L}_N(\epsilon) = M^{-1} \ln \mathbb{E}_{\theta_1} \exp \epsilon \ln L(\theta_0, \theta_1, U_T^N) = \mathbb{L}(\epsilon) + NM^{-1} \mathbb{H}(\epsilon) + M^{-1} \mathbb{R}_N(\epsilon),$$

$$M = \sum_{k=1}^N \lambda_k^2 N^{2/d+1}.$$

Corollary

After a series of technical results, we get

$$\widehat{c}_\alpha(N) / \left(1 - \mathbb{P}_{\theta_1}^{N,T}(\widehat{R}_N)\right) \sim \sqrt{M} \sim N^{1/d+1/2} \rightarrow \infty,$$

and hence \widehat{R}_N is NOT asymptotically most powerful in the class $\widetilde{\mathcal{K}}_\alpha$.

Corollary

After a series of technical results, we get

$$\widehat{c}_\alpha(N) / \left(1 - \mathbb{P}_{\theta_1}^{N,T}(\widehat{R}_N)\right) \sim \sqrt{M} \sim N^{1/d+1/2} \rightarrow \infty,$$

and hence \widehat{R}_N is NOT asymptotically most powerful in the class $\widetilde{\mathcal{K}}_\alpha$.

Solution: a new class of tests

Similar remark is true for T :

$$c_\alpha^\sharp(T) = 1 - \mathbb{P}_{\theta_1}^{N,T}(R_T^\sharp) \sim \bar{T}, \quad \text{as } T \rightarrow \infty,$$

where

$$R_T^\sharp = \{U_T^N : L(\theta_0, \theta_1, U_T^N) \leq c_\alpha^\sharp(T)\}, \quad \text{for all } T,$$

$$c_\alpha^\sharp(T) = \exp \left\{ -\frac{(\theta_1 - \theta_0)^2}{4\theta_0} MT - \frac{\theta_1^2 - \theta_0^2}{2\theta_0} \frac{MT}{2\theta_0} \right\} q_\alpha.$$

Control of Type I error

Proposition (C. and Xu)

For any $x \in \mathbb{R}$, we have the following expansion,

$$\begin{aligned} \mathbb{P}_{\theta_0}^{N,T} (I^N \leq x) = & \Phi(x) + \delta_1(x)M^{-1/2} + \delta_2(x)NM^{-1} \\ & + \mathfrak{R}_N^\delta(x) \left(M^{-1} + NM^{-3/2} + N^2M^{-2} \right), \end{aligned}$$

where

$$I^N = -\frac{\sqrt{8\theta_0^3 \ln L(\theta_0, \theta_1, U_T^N)}}{\sqrt{TM}(\theta_1^2 - \theta_0^2)} - \frac{\sqrt{2\theta_0 TM}(\theta_1 - \theta_0)}{2(\theta_1 + \theta_0)} + \frac{(\theta_1 - \theta_0)N}{\sqrt{8\theta_0 TM}(\theta_1 + \theta_0)},$$

and $\Phi(\cdot)$ is the distribution of a standard Gaussian random variable, and $\delta_1(\cdot)$ and $\delta_2(\cdot)$ are bounded smooth functions and all their derivatives are bounded, and $\mathfrak{R}_N(\cdot)$ is uniformly bounded in x .

Main result

Define

$$\hat{\mathcal{K}}_\alpha := \left\{ (R_N) : \limsup_{N \rightarrow \infty} \left(\mathbb{P}_{\theta_0}^{N,T}(R_N) - \alpha \right) \sqrt{M} \leq \hat{\alpha}_1 \right\}.$$

Note that $\hat{R}_N = \{U_T^N : I^N \leq q_\alpha\}$, so by the previous lemma we have

$$\lim_{N \rightarrow \infty} \left(\mathbb{P}_{\theta_0}^{N,T}(\hat{R}_N) - \alpha \right) \sqrt{M} = \hat{\alpha}_1,$$

where $\hat{\alpha}_1$ is an explicit constant.

Main result

Define

$$\hat{\mathcal{K}}_\alpha := \left\{ (R_N) : \limsup_{N \rightarrow \infty} \left(\mathbb{P}_{\theta_0}^{N,T}(R_N) - \alpha \right) \sqrt{M} \leq \hat{\alpha}_1 \right\}.$$

Note that $\hat{R}_N = \{U_T^N : I^N \leq q_\alpha\}$, so by the previous lemma we have

$$\lim_{N \rightarrow \infty} \left(\mathbb{P}_{\theta_0}^{N,T}(\hat{R}_N) - \alpha \right) \sqrt{M} = \hat{\alpha}_1,$$

where $\hat{\alpha}_1$ is an explicit constant.

Theorem (C. and Xu)

The rejection region (\hat{R}_N) is asymptotically the most powerful in $\hat{\mathcal{K}}_\alpha$.

PART II: **Error Control and Numerical Results**

New Tests for Error Control, $N \rightarrow \infty$

Theorem (C. and Xu)

Consider the test statistics of the form

$$R_N^0 = \{U_T^N : \ln L(\theta_0, \theta_1, U_T^N) \geq \zeta_0 M\},$$

where ζ_0 is given by an explicit formula of the form $-\frac{(\theta_1 - \theta_0)^2}{4\theta_0} T + O(N^{-1/2-1/d})$. If $N \geq N_0$ (N_0 has explicit formula), then the Type I and Type II errors have the following bound estimates

$$\begin{aligned} \mathbb{P}_{\theta_0}^{N,T} (R_N^0) &\leq (1 + \varrho)\alpha, \\ 1 - \mathbb{P}_{\theta_1}^{N,T} (R_N^0) &\leq (1 + \varrho) \exp\left(-\frac{(\theta_1 - \theta_0)^2}{16\theta_0^2} MT\right), \end{aligned}$$

where ϱ denotes a given threshold of error tolerance.

Table: Type I errors for various N

| N | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 |
|--|-------|-------|-------|-------|-------|-------|-------|-------|
| $\mathbb{P}_{\theta_0}^{N,T} (R_N^0)$ | 0.007 | 0.012 | 0.010 | 0.017 | 0.012 | 0.014 | 0.010 | 0.013 |
| $\mathbb{P}_{\theta_0}^{N,T} (R_N^\#)$ | 0.006 | 0.037 | 0.039 | 0.053 | 0.040 | 0.039 | 0.054 | 0.046 |

Other parameters: $\alpha = 0.05$, $\theta_0 = 0.1$, $\theta_1 = 0.2$, $T = 1$, $\varrho = 0.1$, $d = \sigma = 1$, $\gamma = 0$

Table: $T = T_b^1$ given by theoretical results and Type I error for various α .

| α | 0.1 | 0.05 | 0.01 | 0.005 |
|--------------------------------------|-------|-------|--------|--------|
| T_b^1 | 629 | 818 | 1258 | 1447 |
| $\mathbb{P}_{\theta_0}^{N,T}(R_T^0)$ | 0.021 | 0.010 | 0.0025 | 0.0015 |

Other parameters: $\theta_0 = 0.1$, $\theta_1 = 0.2$, $N = 3$, $\varrho = 0.1$, $d = \sigma = 1$, $\gamma = 0$

Table: Type II errors for various T .

| T | 10 | 20 | 30 | 40 | 50 | 60 |
|--|--------|--------|--------|--------|--------|--------|
| $\exp - \frac{(\theta_1 - \theta_0)^2}{16\theta_0} MT$ | 0.4169 | 0.1738 | 0.0724 | 0.0302 | 0.0126 | 0.0052 |
| $1 - \mathbb{P}_{\theta_1}^{N,T} R_T^0$ | 0.7155 | 0.3329 | 0.1148 | 0.0293 | 0.0070 | 0.0012 |
| $1 - \mathbb{P}_{\theta_1}^{N,T} R_T^\dagger$ | 0.7946 | 0.2402 | 0.0457 | 0.0060 | 0.0006 | 0.0002 |

Other parameters: $\alpha = 0.05$, $\theta_0 = 0.1$, $\theta_1 = 0.2$, $N = 3$, $\varrho = 0.1$, $d = \sigma = 1$, $\gamma = 0$

Thank You !

The end of the talk . . .
but not of the story . . .