Recent Advances in Statistical Inference for Stochastic PDEs

Igor Cialenco
Department of Applied Mathematics
Illinois Institute of Technology
cialenco@iit.edu

Conference On Ambit Fields And Related Topics
Aarhus University, Denmark

August 15, 2016

Outline of the talk

- **Introduction:** Parameter Estimation for SODE vs SPDE.
- **Part I:** Parameter Estimation for Stochastic PDEs.

  Maximum Likelihood Estimators
  Trajectory Fitting Estimators
- **Part II:** Hypothesis testing

Stochastic ODE: Estimating Drift θ, with σ known

$$dX(t) = \theta X(t)dt + \sigma X(t)dw(t), \quad t \geq 0$$

Problem

Assuming that one sample path $X(\omega, t)$, $t \in [0, T]$, is observed, find/estimate the parameters θ and σ.

θ**:** Girsanov Theorem (change of drift) $\mapsto$ find the Likelihood Ratio $\mapsto$ Maximize $d\mathbb{P}/d\mathbb{P}_0$ $\mapsto$ find MLE

$$\widehat{\theta}_t = \frac{1}{t}\int_0^t \frac{dX(s)}{X(s)} = \frac{1}{t}\log\frac{X(t)}{X(0)} - \frac{\sigma^2}{2}$$

$$\boxed{\hat{\theta}_t \to \theta \ , \quad t \to \infty}$$

σ**:** Quadratic Variation $\langle X\rangle_t = \sigma^2 \int_0^t X_s^2 ds \rightsquigarrow \sigma = \sqrt{\langle X\rangle_t / \int_0^t X_s^2 ds}$

Stochastic ODE: conclusion

- the drift θ - approximated.
  Regular model
  1) $\frac{d\mathbb{P}_\theta}{d\mathbb{P}_0}$ exists; 2) has a special form (LAN)
  Same procedure for all
  Find MLE by maximizing likelihood ratio

- the volatility σ - exactly.
  Singular model otherwise
  Individual approach
  In particular, if $\mathbb{P}_{\sigma_1} \perp \mathbb{P}_{\sigma_2}$ for $\sigma_1 \neq \sigma_2$, then one may find σ exactly

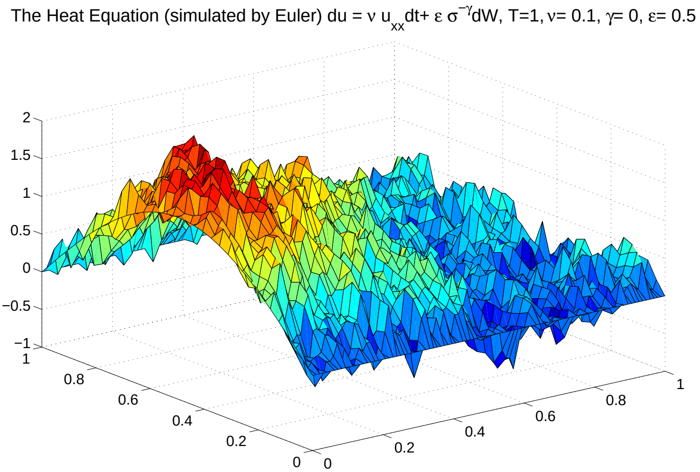
What do we have for SPDEs?

Reference example of SPDE to keep in mind:

$$du(t,x) = \theta u_{xx}dt + \sigma dW(t,x), \quad t \ge 0,\ x \in [0,\pi],$$

with zero boundary conditions and $dW(t,x) = \sum_{k=1}^{\infty} \sin(kx)dw_k(t)$.

What do we have for SPDEs? Mostly singular problems.



Explore the singularity and try to find the exact value (or as a limit of regular models) of the drift/viscosity coefficient.

- *additive noise*: Huebner-Khasminskii-Rozovskii '92, '95
- *Bayesian*: Bishwal ('02)
- *Several parameters*: Huebner ('97)
- *Discrete-time observations*: Piterbarg-Rozovskii ('97)
  $q = \frac{2(m_1 - 2m)}{d} \geq 1$, Markussen '03
- $\theta(t)$*-random*: Lototsky ('04)
- *Small noise*: Huebner ('97), Ibragimov-Khasminskii ('98,'99)
- *"almost" diagonalizable model*: Rozovskii-Lototsky ('97, '01)
- additive fractional noise: IgC, Lototsky, Pospisil ('09)
- multiplicative noise: IgC and Lototsky ('08), IgC ('10)
- nonlinear SPDE: IgC and Glatt-Holtz ('11)
- Hypothesis testing: IgC and Xu ('14, '15)
- Non-MLE / Trajectory fitting estimators: IgC, Gong, Huang ('16)

Part I(a): Maximum Likelihood Estimators

$$dU(t) + \theta AU(t)dt + F(U)dt = \sigma dW(t), \quad U(0) = U_0$$

- given stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$
- assume that $U(\omega, t)$ belongs to some "suitable" Hilbert space $\mathcal{H}$; in particular $U = U(\omega, t, x)$
- A a linear, selfadjoint, positive-defined (think Laplace$^\beta$) in $\mathcal{H}$ with eigenfunctions $\{h_k\}_{k\geq 1}$ CONS in $\mathcal{H}$
- $\sigma dW(t) = \sum_{k\geq 1} \sigma_k h_k dW_k(t)$, $W_k, k \in \mathbb{N}$ ind. Brownian Motions
- F maybe nonlinear; σ **known**
- U observed for all $t \in [0, T]$ - **continuous observations**

Goal:

Find estimators $\hat{\theta}(\omega)$, $\omega \in \Omega$, for parameters θ by **observing a single outcome** $U = U(\omega, t) \in \mathcal{H}$ over a finite time horizon $t \in [0, T]$.

Formal Procedure to Derive an Estimator

- Project the full system down to N dimensions $P_N(\mathcal{H}) = \mathcal{H}_N \simeq \mathbb{R}^N$

$$dU^N + (\theta A U^N + \Psi_N)dt = P_N \sigma dW, \quad U(0) = U_0$$

- Let $\mathbb{P}_\theta^{N,T}(\cdot) = \mathbb{P}(U^N \in \cdot)$ be the measure on $C([0,T];\mathbb{R}^N)$ generated by U^N;
  $\mathbb{P}_\theta^T$ be the measure generated by U on $C([0,T];\mathcal{H})$.
- Usually (at least in linear case), we can prove that $\mathbb{P}_{\theta_1}^{N,T} \sim \mathbb{P}_{\theta_2}^{N,T}$
  Hence, get MLE type estimators $\widehat{\theta}_{N,T}$.

Formal Procedure to Derive an Estimator

- Project the full system down to N dimensions $P_N(\mathcal{H}) = \mathcal{H}_N \simeq \mathbb{R}^N$

$$dU^N + (\theta A U^N + \Psi_N)dt = P_N \sigma dW, \quad U(0) = U_0$$

- Let $\mathbb{P}_\theta^{N,T}(\cdot) = \mathbb{P}(U^N \in \cdot)$ be the measure on $C([0,T];\mathbb{R}^N)$ generated by U^N;
  $\mathbb{P}_\theta^T$ be the measure generated by U on $C([0,T];\mathcal{H})$.
- Usually (at least in linear case), we can prove that $\mathbb{P}_{\theta_1}^{N,T} \sim \mathbb{P}_{\theta_2}^{N,T}$
  Hence, get MLE type estimators $\widehat{\theta}_{N,T}$.

Reasonable ansatz:

$$\widehat{\theta}_{N,T} \xrightarrow[N\to\infty]{} \theta$$

Formal Procedure to Derive an Estimator in Nonlinear Case

- Formally treat $\Psi_N = P_N F(U)$ as an external and known quantity (independent of θ)
- Assume that $P_N\sigma$ is invertible on H_N
- Take $G := P_N\sigma(U)(P_N\sigma(U))^*$ and assume it commutes with A
- For a reference values θ_0, apply (formally) Girsanov Theorem and get the 'Likelihood Ratio' (Radon-Nikodym derivative) $d\mathbb{P}_\theta^{N,T}/d\mathbb{P}_{\theta_0}^{N,T}$
- Maximize the Log-Likelihood Ratio $\widetilde{\theta}_{N,T}(\omega) := \arg\max_\theta d\mathbb{P}_\theta^{N,T}/d\mathbb{P}_{\theta_0}^{N,T}(\omega)$

$$
\begin{aligned}
\frac{d\mathbb{P}_{\theta}^{N,T}}{d\mathbb{P}_{\theta_0}^{N,T}} = \exp\Big[\int_0^T (\theta-\theta_0)\langle AU^N, GdU^N(t)\rangle \\
+ \frac{1}{2}\int_0^T (\theta^2-\theta_0^2)\langle AU^N, GAU^N dt\rangle \\
+ \int_0^T (\theta-\theta_0)\langle AU^N, G\psi^N dt\rangle\Big],
\end{aligned}
$$

$$
\widetilde{\theta}_N = -\frac{\int_0^T \langle AU^N, GdU^N\rangle + \int_0^T \langle AU^N, GP_N\mathbf{F(U)}\rangle dt}{\int_0^T \langle AU_N, GAU^N\rangle dt}
$$

$$\frac{d\mathbb{P}_{\theta}^{N,T}}{d\mathbb{P}_{\theta_0}^{N,T}} = \exp\Big[\int_0^T (\theta - \theta_0)\langle AU^N, GdU^N(t)\rangle$$

$$+ \frac{1}{2}\int_0^T (\theta^2 - \theta_0^2)\langle AU^N, GAU^N dt\rangle$$

$$+ \int_0^T (\theta - \theta_0)\langle AU^N, G\psi^N dt\rangle\Big],$$

$$\widetilde{\theta}_N = -\frac{\int_0^T \langle AU^N, GdU^N\rangle + \int_0^T \langle AU^N, GP_N\mathbf{F(U)}\rangle dt}{\int_0^T \langle AU_N, GAU^N\rangle dt}$$

Main Idea #1: Modified MLE

$$\widetilde{\theta}_N = -\frac{\int_0^T A^{1+\rho_1}U_N G_N^{\rho_2} dU_N + \int_0^T A^{1+\rho_1}U_N G_N^{\rho_2} P_N F(U))dt}{\int_0^T A^{1+\rho_1}U_N G_N^{\rho_2} AU_N dt}$$

for some ρ_1, ρ_2.

Motivated by MLE type estimator

$$\hat{\theta}_{1,N} = -\frac{\int_0^T A^{1+\rho_1} U_N G_N^{\rho_2} dU_N + \int_0^T A^{1+\rho_1} U_N G_N^{\rho_2} P_N F(U))dt}{\int_0^T A^{1+\rho_1} U_N G_N^{\rho_2} A U_N dt},$$

$$\hat{\theta}_{2,N} = -\frac{\int_0^T A^{1+\rho_1} U_N G_N^{\rho_2} dU_N + \int_0^T A^{1+\rho_1} U_N G_N^{\rho_2} P_N F(U_N))dt}{\int_0^T A^{1+\rho_1} U_N G_N^{\rho_2} A U_N dt},$$

$$\hat{\theta}_{3,N} = -\frac{\int_0^T A^{1+\rho_1} U_N G_N^{\rho_2} dU_N}{\int_0^T A^{1+\rho_1} U_N G_N^{\rho_2} A U_N dt}.$$

Choose ρ_1, ρ_2 such that we can prove

$$\widehat{\theta}_{i,N} \longrightarrow \theta, \qquad \text{as } N \to \infty,$$

for $i = 1, 2, 3$.

$$\hat{\theta}_{2,N} = \theta + \frac{\int_0^T \langle A^{1+\rho_1} U^N, G^{\rho_2} \sum_{j=1}^N \sigma_j(U)\Phi_j dW_j(t)\rangle}{\int_0^T A^{1+\rho_1} U_N G_N^{\rho_2} A U_N dt} + \frac{\int_0^T \langle A^{1+\rho_1} U^N, G^{\rho_2}(F^N(U) - F^N(U^N))\rangle dt}{\int_0^T A^{1+\rho_1} U_N G_N^{\rho_2} A U_N dt}$$

$$\hat{\theta}_{3,N} = \theta + \frac{\int_0^T \langle A^{1+\rho_1} U^N, G^{\rho_2} \sum_{j=1}^N \sigma_j(U)\Phi_j dW_j(t)\rangle}{\int_0^T A^{1+\rho_1} U_N G_N^{\rho_2} A U_N dt} + \frac{\int_0^T \langle A^{1+\rho_1} U^N, G^{\rho_2} F^N(U^N)\rangle dt}{\int_0^T A^{1+\rho_1} U_N G_N^{\rho_2} A U_N dt}$$

- Need to show that each of ‘the excess term converge to zero’
- Successfully applied to:
  - Stochastic linear parabolic SPDE, additive noise
  - Stochastic Navier-Stokes Equations, 2D, additive noise

PART I(B): Trajectory Fitting Estimators

I. Cialenco, R. Gong and Y. Huang, *Trajectory Fitting Estimators for SPDEs Driven by Additive Noise* submitted for publication, 2016 http://arxiv.org/abs/1607.04912

Trajectory fitting estimators (TFE) for SDEs

The observed process $S(\theta) := \{S(t;\theta)\}_{t\geq 0}$ follows the dynamics

$$dS(t;\theta) = b(\theta, S(t;\theta))dt + \sigma(S(t;\theta))\, dB(t),$$

where B is an 1d standard Brownian motion, and θ is the parameter of interest. Let $F : \mathbb{R} \to \mathbb{R},\ F \in C^2$; by Itô's formula,

$$F(S(t;\theta)) = F(S_0) + \int_0^t \Big(F'(S(s))b\,(\theta, S(s)) + \frac{1}{2}F''(S(s))\sigma^2(S(s))\Big)\,ds + \int_0^t F'(S(s))\sigma(S(s))\,dB(s).$$

For any $\theta \in \Theta$ and $t \in [0,T]$, consider an *artificial trajectory*

$$\widetilde{F}(t;\theta) := F(S_0) + \int_0^t \Big(F'(S(s))b\,(\theta, S(s)) + \frac{1}{2}F''(S(s))\sigma^2(S(s))\Big)\,ds.$$

TFE for SDEs; continued

The *trajectory fitting estimator* $\widetilde{\theta}_T$ of θ is defined as the solution to the minimization problem

$$\widetilde{\theta}_T := \arg\inf_{\theta \in \Theta} \int_0^T \left(F(S(t;\theta)) - \widetilde{F}(t;\theta)\right)^2 dt.$$

The choice of F depends on the underlying models to insure the desired asymptotic properties of the estimator; e.g. $F(x) = x^2$.

For ergodic, finite dimensional diffusion processes, one can prove that $\widetilde{\theta}_T \to \theta$, as $T \to \infty$.

Goal:

- Can we derive tractable TFEs for SPDEs?
- Study the asymptotic properties of TFEs as number of Fourier modes N increases.

TFE for SPDEs

$(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P})$ a stochastic basis;
We consider the evolution equation, in a separable Hilbert space H

$$du(t) + \theta \mathcal{A} u(t)\, dt = \sigma\, dW(t), \quad u_0 \in H,$$

where $\mathcal{A}_1$ is a linear operators on H, $W := \{W(t)\}_{t\geq 0}$ is a cylindrical Brownian motion in H

Continuous-time observation framework of first N Fourier modes on a finite time interval $t \in [0, T]$.

Parameter of interest $\theta \in \Theta \subset \mathbb{R}_+$.

$$du(t) + \theta \mathcal{A} u(t)\, dt = \sigma\, dW(t), \quad u_0 \in H, \tag{5.1}$$

- The operator $\mathcal{A}$ has only point spectra; the eigenfunctions $\{h_k\}_{k\in\mathbb{N}}$ form a complete, orthonormal system in H; eigenvalues ν_k, $k \in \mathbb{N}$.
- The sequence $\{\nu_k\}_{k\in\mathbb{N}}$ is such that $\lim_{k\to\infty} \nu_k = +\infty$.
- W is a cylindrical Brownian motion in H, and has the following form

$$W(t) = \sum_{k=1}^{\infty} \lambda_k^{-\gamma} h_k w_k(t), \quad t \geq 0,$$

  for some $\gamma \geq 0$, where $\lambda_k := \nu_k^{1/(2m)}$, $k \in \mathbb{N}$, for some $m \geq 0$, and $w_k := \{w_k(t)\}_{t\geq 0}$, $k \in \mathbb{N}$, are independent standard Brownian motions.
- That is: the equation (5.1) is linear, diagonalizable, parabolic, and the solution exists and is unique.

$$du(t) + \theta \mathcal{A} u(t)\, dt = \sigma\, dW(t), \quad u_0 \in H,$$

The unique solution is given by

$$u(t) = \sum_{k=1}^{\infty} u_k(t) h_k, \quad t \geq 0,$$

where, each Fourier mode $u_k,\ k \geq 1$ satisfies the SDE

$$du_k(t) + \theta \nu_k\, u_k(t)\, dt = \sigma \lambda_k^{-\gamma}\, dw_k(t), \quad u_k(0) = (u_0, h_k)_H,$$
$$u_k(t) = e^{-\nu_k \theta t} u_k(0) + \sigma \lambda_k^{-\gamma} e^{-\nu_k \theta t} \int_0^t e^{\nu_k \theta s}\, dw_k(s).$$

We denote by V_k the *artificial trajectory* of u_k, as

$$V_k(t;\theta) := u_k^2(0) + \int_0^t \left(\sigma^2 \lambda_k^{-2\gamma} - 2\nu_k \theta u_k^2(s) \right) ds, \quad k \in \mathbb{N}, \quad t \in [0, T].$$

Definition

The Trajectory Fitting Estimator for the drift parameter θ is defined as

$$\widetilde{\theta}_N = \widetilde{\theta}_N(T) := \arg\inf_{\theta\in\Theta} \sum_{k=1}^{N} \int_0^T \big(V_k(t;\theta) - u_k^2(t)\big)^2\, dt.$$

By direct evaluations, TFE can be computed explicitly

$$\widetilde{\theta}_N = -\frac{\sum_{k=1}^N \nu_k \left(\frac{1}{2}\xi_k^2(T) - u_k^2(0)Y_k(T) - \sigma^2\lambda_k^{-2\gamma}X_k(T)\right)}{2\sum_{k=1}^N \nu_k^2 Z_k(T)},$$

where

$$\xi_k(t) := \int_0^t u_k^2(s)\, ds, \qquad X_k(t) := \int_0^t s\xi_k(s)\, ds,$$
$$Y_k(t) := \int_0^t \xi_k(s)\, ds, \qquad Z_k(t) := \int_0^t \xi_k^2(s)\, ds.$$

TFE: Consistency

Noting that

$$\widetilde{\theta}_N - \theta = -\frac{\sum_{k=1}^N \nu_k \left(\frac{1}{2}\xi_k^2 - u_k^2(0)Y_k - \sigma^2\lambda_k^{-2\gamma}X_k + 2\nu_k\theta Z_k\right)}{2\sum_{k=1}^N \nu_k^2 Z_k} =: -\frac{\sum_{k=1}^N \nu_k A_k}{2\sum_{k=1}^N \nu_k^2 Z_k}.$$

Proposition (CGH '16)

$$\mathbb{E}(Z_k) \asymp \frac{1}{\mu_k^2\theta^2}\left(u_k^2(0) + \sigma^2 T\lambda_k^{-2\gamma}\right)^2, \quad k \to \infty,$$

$$\mathit{Var}(Z_k) \asymp \frac{\lambda_k^{-2\gamma}}{\nu_k^5\theta^5}\left(u_k^2(0) + \sigma^2 T\lambda_k^{-2\gamma}\right)^3,$$

$$\mathbb{E}(A_k) \asymp \frac{\lambda_k^{-2\gamma}}{\nu_k^2\theta^2}\left(u_k^2(0) + \sigma^2 T\lambda_k^{-2\gamma}\right),$$

$$\mathit{Var}(A_k) \asymp \frac{\lambda_k^{-2\gamma}}{\nu_k^3\theta^3}\left(u_k^2(0) + \sigma^2 T\lambda_k^{-2\gamma}\right)^3.$$

TFE: Consistency

Theorem (CGH '16)

Assume that

$$\sum_{k=}^{\infty} \lambda_k^{-4\gamma} = \infty.$$

Then,

$$\lim_{N\to\infty} \widetilde{\theta}_N = \theta, \quad \mathbb{P} - a.s..$$

TFE: Asymptotic Normality

Theorem (CGH '16)

If in addition

$$\sum_{k=1}^{\infty} \lambda_k^{-8\gamma} \nu_k^{-1} = \infty.$$

Then, as $N \to \infty$,

$$\frac{\widetilde{\theta}_N - \theta + a_N}{b_N} \xrightarrow{d} \mathcal{N}(0,1), \tag{5.2}$$

where

$$a_N := \frac{\sum_{k=1}^N \nu_k \, \mathbb{E}(A_k)}{2\sum_{k=1}^N \nu_k^2 \, \mathbb{E}(Z_k)}, \quad b_N := \frac{\sqrt{\sum_{k=1}^N \nu_k^2 \, \mathit{Var}(A_k)}}{2\sum_{k=1}^N \nu_k^2 \, \mathbb{E}(Z_k)}, \tag{5.3}$$

and where $\xrightarrow{d}$ *denotes the convergence in distribution.*

Example

Fractional stochastic heat equation driven by an additive noise:

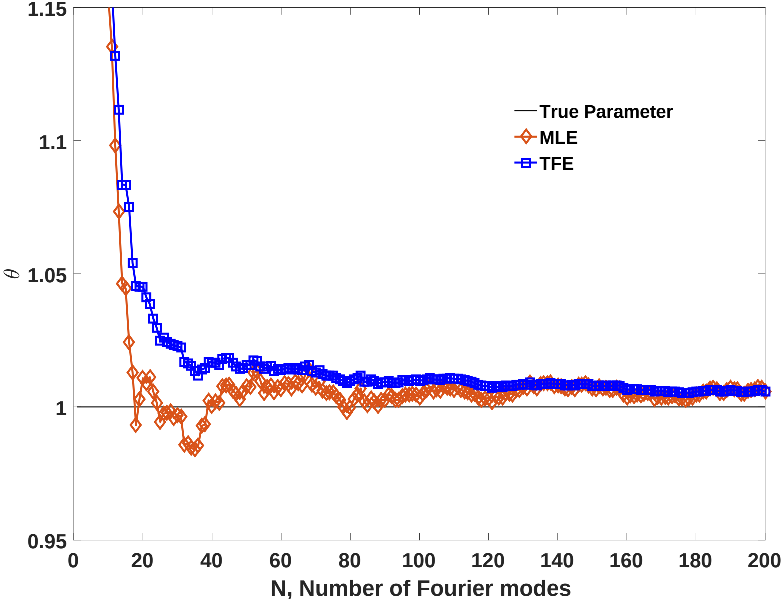
$$du(t,x)+\theta(-\Delta)^{\beta}u(t,x)\,dt=\sigma\sum_{k=1}^{\infty}\lambda_k^{-\gamma}h_k(x)\,dw_k(t),\quad t\in[0,T],\quad x\in G,$$

with initial condition $u(0,x)=u_0(x)\in H$, where $\theta>0$, $\beta>0$, $\gamma\geq 0$ and $\sigma\in\mathbb{R}\setminus\{0\}$ are constants. In this case,

$$\nu_k\sim c_1k^{2\beta/d},\quad \lambda_k\sim\sqrt{c_1}\,k^{1/d},\quad k\to\infty.$$

The consistency and the asymptotic normality hold for the TFE $\tilde{\theta}_N$, whenever

$$2\beta+8\gamma\leq d.$$



True Parameter
MLE
TFE
θ
N, Number of Fourier modes
1.15
1.1
1.05
1
0.95
0
20
40
60
80
100
120
140
160
180
200

PART II: Hypothesis Testing for SPDEs

I. Cialenco, L. Xu, Hypothesis testing for stochastic PDEs driven by additive noise, Stochastic Processes and their Appl., 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.

Similar Setup

Fractional heat equation driven by additive noise:

$$\mathrm{d}U(t,x) + \theta(-\Delta)^{\beta}U(t,x)\mathrm{d}t = \sigma \sum_{k=1}^{\infty} \lambda_k^{-\gamma} h_k(x)\mathrm{d}w_k(t),$$

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

- 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;
- $\{h_k\}$ are the eigenfunctions of Δ in $L^2(G)$; $\{\rho_k\}$ are the eigenvalues; $\lambda_k = \sqrt{-\rho_k} \sim k^{1/d}$;
- consider solution in $(H^{\beta+s}(G), H^{s}(G), H^{-\beta+s}(G))$;
- $\theta > 0$ (**Unknown**),
  all other parameters $\beta > 0,\ \gamma \geq 0,\ \sigma \in \mathbb{R} \setminus \{0\}$ known.

Simple Hypothesis

$$\mathrm{d}U(t,x) + \theta(-\Delta)^{\beta}U(t,x)\mathrm{d}t = \sigma \sum_{k=1}^{\infty} \lambda_k^{-\gamma} h_k(x)\mathrm{d}w_k(t)$$

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

Consider a simple hypothesis:

$$\begin{aligned} \mathscr{H}_0: &\quad \theta = \theta_0, \\ \mathscr{H}_1: &\quad \theta = \theta_1. \end{aligned}$$

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

Construction of the Test

$$\mathrm{d}U(t,x) + \theta(-\Delta)^{\beta}U(t,x)\mathrm{d}t = \sigma \sum_{k=1}^{\infty} \lambda_k^{-\gamma} h_k(x)\mathrm{d}w_k(t), \quad U(0,x) = 0.$$

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

$$\mathrm{d}u_k = -\theta\lambda_k^{2\beta}u_k\mathrm{d}t + \sigma\lambda_k^{-\gamma}\mathrm{d}w_k(t), \quad u_k(0) = 0,$$
$$u_k(t) = \sigma\lambda_k^{-\gamma}\int_0^t e^{-\theta\lambda_k^{2\beta}(t-s)}dw_k, \quad k \geq 1.$$

- 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 $U_T^N(t) = (u_1,\ldots,u_N)$ up to time T.

Observable: First N Fourier coefficients $u_1(t), \dots, u_N(t)$, for all $t \in [0, T]$.

- 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

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

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

Definition

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

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

Neyman-Pearson Lemma

Theorem (C. and Xu, '14, '15)

Take the Likelihood Ratio

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

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 $\mathcal{K}_\alpha$.

The Difficulty:

The problem is that c_α has no explicit formula for finite T and N.

The Difficulty:

The problem is that c_α has no explicit formula for finite T and N.

We suggest/take **"Asymptotic Method"**

(1) Fix N, let $T \to \infty$;

(2) Fix T, let $N \to \infty$.

In this talk we focus on case (1), large time asymptotics;

For case (2) see [CX '14 and '15].

Asymptotic Method in Time T:

Define a new class

$$\mathcal{K}^*_\alpha := \left\{ (R_T)_{T\in\mathbb{R}_+} : R_T \in \mathcal{B}(C([0,T];\mathbb{R}^N)),\ \limsup_{T\to\infty} \mathbb{P}^{N,T}_{\theta_0}(R_T) \le \alpha \right\},$$

where N is fixed, and α is the **"Asymptotic Significance Level"**.

Asymptotic Method in Time T:

Define a new class

$$\mathcal{K}_\alpha^* := \left\{ (R_T)_{T\in\mathbb{R}_+} : R_T \in \mathcal{B}(C([0,T];\mathbb{R}^N)),\ \limsup_{T\to\infty} \mathbb{P}_{\theta_0}^{N,T}(R_T) \le \alpha \right\},$$

where N is fixed, and α is the **"Asymptotic Significance Level"**.

Goal:

We want to find a rejection region $(R_T^*)_{T\in\mathbb{R}_+}$ such that

$$\lim_{T\to\infty} \mathbb{P}_{\theta_0}^{N,T}(R_T^*) = \alpha.$$

Attempt:

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

To find c_α, we make the following heuristic argument: by Itō's Formula,

$$\begin{aligned}&\mathbb{P}_{\theta_0}^{N,T}(L(\theta_0,\theta_1,U_T^N)\geq c_\alpha^*)\\=&\mathbb{P}_{\theta_0}^{N,T}\left(X_T-\frac{2(\theta_1+\theta_0)}{(\theta_1-\theta_0)\sigma\sqrt{T}}Y_T\geq\frac{4\theta_0\ln c_\alpha^*}{(\theta_1-\theta_0)^2T}+M\right),\end{aligned}$$

where

$$\begin{aligned}M:=\sum_{k=1}^N\lambda_k^{2\beta},\quad X_T:=\sum_{k=1}^N\frac{\lambda_k^{2\beta+2\gamma}u_k^2(T)}{\sigma^2T},\\ Y_T:=\frac{1}{\sqrt{T}}\sum_{k=1}^N\lambda_k^{2\beta+\gamma}\int_0^Tu_kdw_k.\end{aligned}$$

We can prove:

- And we have the split:

$$\mathbb{P}_{\theta_0}^{N,T}(L(\theta_0,\theta_1,U_T^N) \geq c_\alpha^*) \leq \mathbb{P}_{\theta_0}^{N,T}(X_T \geq \delta)$$
$$+ \mathbb{P}_{\theta_0}^{N,T}\left(-\frac{2(\theta_1+\theta_0)}{(\theta_1-\theta_0)\sigma\sqrt{T}} Y_T \geq \frac{4\theta_0 \ln c_\alpha^*}{(\theta_1-\theta_0)^2 T} + M - \delta\right).$$

- For any fixed $\delta > 0$, $\mathbb{P}_{\theta_0}^{N,T}(X_T \geq \delta) \to 0$ as $T \to \infty$.
- $Y_T \xrightarrow{d} \mathcal{N}(0, \sigma^2 M/(2\theta_0))$ as $T \to \infty$.

We can prove:

- And we have the split:

$$\mathbb{P}_{\theta_0}^{N,T}(L(\theta_0,\theta_1,U_T^N) \geq c_\alpha^*) \leq \mathbb{P}_{\theta_0}^{N,T}(X_T \geq \delta)$$
$$+ \mathbb{P}_{\theta_0}^{N,T}\left(-\frac{2(\theta_1+\theta_0)}{(\theta_1-\theta_0)\sigma\sqrt{T}} Y_T \geq \frac{4\theta_0 \ln c_\alpha^*}{(\theta_1-\theta_0)^2 T} + M - \delta\right).$$

- For any fixed $\delta > 0$, $\mathbb{P}_{\theta_0}^{N,T}(X_T \geq \delta) \to 0$ as $T \to \infty$.
- $Y_T \xrightarrow{d} \mathcal{N}(0, \sigma^2 M/(2\theta_0))$ as $T \to \infty$.

It Is Reasonable To Take:

$$-\sqrt{\frac{2\theta_0}{M}} \frac{(\theta_1-\theta_0)\sqrt{T}}{2(\theta_1+\theta_0)} \left[\frac{4\theta_0 \ln c_\alpha^*}{(\theta_1-\theta_0)^2 T} + M\right] = q_\alpha. \qquad (6.1)$$

Solve (6.1) to get

$$c_\alpha^\sharp(T) = \exp\left(-\frac{(\theta_1-\theta_0)^2}{4\theta_0}MT - \frac{\theta_1^2-\theta_0^2}{2\theta_0}\sqrt{\frac{MT}{2\theta_0}}q_\alpha\right). \qquad (6.2)$$

Solve (6.1) to get

$$c_\alpha^\sharp(T) = \exp\left(-\frac{(\theta_1-\theta_0)^2}{4\theta_0}MT - \frac{\theta_1^2-\theta_0^2}{2\theta_0}\sqrt{\frac{MT}{2\theta_0}}q_\alpha\right). \tag{6.2}$$

Theorem (C. and Xu)

Suppose

$$R_T^\sharp := \{U_T^N : L(\theta_0,\theta_1,U_T^N) \geq c_\alpha^\sharp(T)\}, \qquad \textit{for all } T,$$

where $c_\alpha^\sharp$ *is given by* (6.2). *Then, the rejection region* $(R_T^\sharp)_{T\in\mathbb{R}_+} \in \mathcal{K}_\alpha^*$, *and moreover*

$$\lim_{T\to\infty} \mathbb{P}_{\theta_0}^{N,T}(R_T^\sharp) = \alpha.$$

The Next Question:

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

The Next Question:

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

Theorem (C. and Xu)

$$1 - \mathbb{P}_{\theta_1}^{N,T}(R_T^\sharp) \sim \exp(-I(\theta_0, \theta_1, N)T + o(T)), \quad as\ T \to \infty,$$

where $I(\theta_0, \theta_1, N) = (\theta_1 - \theta_0)^2 M / 4\theta_0$.

The Next Question:

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

Theorem (C. and Xu)

$$1 - \mathbb{P}_{\theta_1}^{N,T}(R_T^\sharp) \sim \exp(-I(\theta_0, \theta_1, N)T + o(T)), \quad as\ T \to \infty,$$

where $I(\theta_0, \theta_1, N) = (\theta_1 - \theta_0)^2 M / 4\theta_0$.

Sketch of the Proof:

- Calculate the Moment Generating Function of the Log-Likelihood ratio (Gapeev and Küchler [2008])
  - Use Feynman-Kac Formula to derive a PDE
  - Make some transforms and guess the solution
- Apply a theorem for Large Deviation in Lin'kov [1999]
- Use some technics in limit theory to get the final result.

Questions to be answered:

- Except for $(R_T^\sharp)$, how do other rejection regions work for the testing? Is $(R_T^\sharp)$ the best one?
- Is the class $\mathcal{K}_\alpha^*$ the best to take for the testing?
- How large T shall we take to insure the accuracy?

Asymptotically The Most Powerful Test

Definition

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

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

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

Theorem (C. and Xu)

There exists rejection region $(\hat{R}_T) \in \mathcal{K}_\alpha^*$ *which is* **Asymptotically More Powerful** *than* $(R_T^\sharp)$, *that is*

$$\limsup_{T\to\infty} \frac{1 - \mathbb{P}_{\theta_1}^{N,T}(\hat{R}_T)}{1 - \mathbb{P}_{\theta_1}^{N,T}(R_T^\sharp)} < 1.$$

Theorem (C. and Xu)

There exists rejection region $(\hat{R}_T) \in \mathcal{K}_\alpha^*$ *which is* **Asymptotically More Powerful** *than* $(R_T^\sharp)$*, that is*

$$\limsup_{T\to\infty} \frac{1 - \mathbb{P}_{\theta_1}^{N,T}(\hat{R}_T)}{1 - \mathbb{P}_{\theta_1}^{N,T}(R_T^\sharp)} < 1.$$

Theorem (C. and Xu)

The rejection region of the form

$$R_T := \left\{ U_T^N : L(\theta_0, \theta_1, U_T^N) \geq c_\alpha(T) \right\},$$

with $c_\alpha(T) > 0$*, can not be asymptotically the most powerful in the class* $\mathcal{K}_\alpha^*$.

Refined Asymptotic Class

Consider the class of the form:

$$\mathcal{K}_{\alpha}^{\sharp} := \left\{ (R_T) : \limsup_{T\to\infty} \left(\mathbb{P}_{\theta_0}^{N,T}(R_T) - \alpha \right) \sqrt{T} \leq \alpha_1 \right\}.$$

where α_1 is some explicitly computable quantity.

Theorem (C. and Xu)

The rejection region $(R_T^{\sharp})$ is **asymptotically the most powerful** *in the class $\mathcal{K}_{\alpha}^{\sharp}$.*

Thank You !

The end of the talk ...

but not of the story

Asymptotic Method in Fourier Modes $N \to \infty$

Define class

$$\tilde{\mathcal{K}}_\alpha(\delta) := \left\{ (R_N) : \limsup_{N\to\infty} \left(\mathbb{P}_{\theta_0}^{N,T}(R_N) - \alpha \right) \sqrt{M} \le \tilde{\alpha}_1(\delta) \right\}. \tag{7.1}$$

Definition

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

$$\liminf_{N\to\infty} \frac{1 - \mathbb{P}_{\theta_1}^{N,T}(R_N)}{1 - \mathbb{P}_{\theta_1}^{N,T}(\tilde{R}_N)} \ge 1, \qquad \text{for all } (R_N) \in \tilde{\mathcal{K}}_\alpha. \tag{7.2}$$

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

Following similar argument as in "T part", one can find

$$\tilde{R}_N^\delta = \left\{ U_T^N : L(\theta_0, \theta_1, U_T^N) \geq \tilde{c}_\alpha^\delta(N) \right\},$$

such that

Theorem (Main Result II)

Assume $\beta/d \geq 1/2$. The rejection region $(\tilde{R}_N^\delta)$ is asymptotically the most powerful in $\tilde{\mathcal{K}}_\alpha(\delta)$.

New Tests for Error Control

Theorem (Error Control for $T \to \infty$)

Consider the test statistics of the form

$$R_T^0 = \left\{ U_T^N : \ln L(\theta_0, \theta_1, U_T^N) \geq \eta_0 T \right\},$$

where η_0 *is given by an explicit formula of the form* $-\frac{(\theta_1 - \theta_0)^2}{4\theta_0} M + O(T^{-1/2})$. *If* $T \geq T_0$ *(*T_0 *has explicit formula), then the Type I and Type II errors have the following bound estimates*

$$\mathbb{P}_{\theta_0}^{N,T}\left(R_T^0\right) \leq (1 + \varrho)\alpha,$$

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

where ϱ *denotes a given threshold of error tolerance.*

Theorem (Error Control for $N \to \infty$)

Consider the test statistics of the form

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

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

$$\mathbb{P}_{\theta_0}^{N,T}\left(R_N^0\right) \leq (1+\varrho)\alpha,$$

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

where ϱ denotes a given threshold of error tolerance.

Simulation Results

Table: *Type I error for various α.*

α	0.1	0.05	0.01	0.005
T_0	629	818	1258	1447
$\mathbb{P}_{\theta_0}^{N,T}(R_{T_0}^0)$	0.020	0.006	0.002	0.001

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

Table: *Type I error for various* $T \geq T_0$

T	T_0	$T_0 + T_1$	$T_0 + 2T_1$	$T_0 + 3T_1$	$T_0 + 4T_1$
$\mathbb{P}_{\theta_0}^{N,T}\left(R_T^0\right)$	0.006	0.014	0.010	0.006	0.010
$\mathbb{P}_{\theta_0}^{N,T}\left(R_T^\sharp\right)$	0.054	0.064	0.050	0.028	0.056

Other parameters: $T_1 = 2000$, $\alpha = 0.05$, $\theta_0 = 0.1$, $\theta_1 = 0.2$, $N = 3$, $\rho = 0.1$, $d = \beta = \sigma = 1$, $\gamma = 0$.

Appendix

Theorem (Criterion for Most Powerful Test)

Consider the rejection region of the form

$$R_T^* = \left\{ U_T^N : L(\theta_0, \theta_1, U_T^N) \geq c_\alpha^*(T) \right\}, \tag{7.3}$$

where $c_\alpha^*(T)$ *is a function of* T *such that,* $c_\alpha^*(T) > 0$ *for all* $T > 0$ *and*

$$\lim_{T\to\infty} \mathbb{P}_{\theta_0}^{N,T}(R_T^*) = \alpha, \tag{7.4}$$

$$\lim_{T\to\infty} \frac{c_\alpha^*(T)}{1 - \mathbb{P}_{\theta_1}^{N,T}(R_T^*)} < \infty. \tag{7.5}$$

Then (R_T^*) *is asymptotically the most powerful in* $\mathcal{K}_\alpha^*$.

Proof for $c_\alpha^\sharp(T)/\left(1-\mathbb{P}_{\theta_1}^{N,T}(R_T^\sharp)\right)\sim\sqrt{T}$:

- Split the probability:

$$1-\mathbb{P}_{\theta_1}^{N,T}(R_T^\sharp)=A_T B_T$$

- After some substitutions and calculations we get

$$A_T \asymp \exp[-I(\theta_0,\theta_1,N)T]$$

- By a series of technical lemmas we proved

$$B_T\sim\exp[o(T)]/\sqrt{T}$$

- Referring to the form of $c_\alpha^\sharp$ in (6.2) we finally have

$$c_\alpha^\sharp(T)/\left(1-\mathbb{P}_{\theta_1}^{N,T}(R_T^\sharp)\right)\sim\sqrt{T}$$

Proof:

By the same reasoning as in **"Neyman-Pearson"**, for a fixed T and any $(R_T) \in \mathcal{K}_\alpha^*$, we have that

$$\mathbb{P}_{\theta_1}^{N,T}(R_T^*) - \mathbb{P}_{\theta_1}^{N,T}(R_T) \geq c_\alpha^*(T)\left(\mathbb{P}_{\theta_0}^{N,T}(R_T^*) - \mathbb{P}_{\theta_0}^{N,T}(R_T)\right),$$

which can be written as

$$\frac{1 - \mathbb{P}_{\theta_1}^{N,T}(R_T)}{1 - \mathbb{P}_{\theta_1}^{N,T}(R_T^*)} \geq 1 + \frac{c_\alpha^*(T)}{1 - \mathbb{P}_{\theta_1}^{N,T}(R_T^*)}\left(\mathbb{P}_{\theta_0}^{N,T}(R_T^*) - \mathbb{P}_{\theta_0}^{N,T}(R_T)\right).$$

From here, using (7.4) and (7.5), we deduce

$$
\begin{aligned}
\liminf_{T\to\infty}\frac{1-\mathbb{P}_{\theta_1}^{N,T}(R_T)}{1-\mathbb{P}_{\theta_1}^{N,T}(R_T^*)} \geq & 1+\lim_{T\to\infty}\frac{c_\alpha^*(T)}{1-\mathbb{P}_{\theta_1}^{N,T}(R_T^*)}\lim_{T\to\infty}\mathbb{P}_{\theta_0}^{N,T}(R_T^*) \\
& -\lim_{T\to\infty}\frac{c_\alpha^*(T)}{1-\mathbb{P}_{\theta_1}^{N,T}(R_T^*)}\limsup_{T\to\infty}\mathbb{P}_{\theta_0}^{N,T}(R_T) \\
= & 1+\lim_{T\to\infty}\frac{c_\alpha^*(T)}{1-\mathbb{P}_{\theta_1}^{N,T}(R_T^*)}\left(\alpha-\limsup_{T\to\infty}\mathbb{P}_{\theta_0}^{N,T}(R_T)\right) \\
\geq & 1.
\end{aligned}
$$

This completes the proof.

Sketch of the proof for main theorem:

By the same reasoning as in **"Neyman-Pearson"**, for a fixed T and any $(R_T) \in \mathcal{K}_\alpha^*$, we have that

$$\mathbb{P}_{\theta_1}^{N,T}(R_T^\sharp) - \mathbb{P}_{\theta_1}^{N,T}(R_T) \geq c_\alpha^\sharp(T)\left(\mathbb{P}_{\theta_0}^{N,T}(R_T^\sharp) - \mathbb{P}_{\theta_0}^{N,T}(R_T)\right),$$

which can be written as

$$\frac{1 - \mathbb{P}_{\theta_1}^{N,T}(R_T)}{1 - \mathbb{P}_{\theta_1}^{N,T}(R_T^\sharp)} \geq 1 + \frac{c_\alpha^\sharp(T)}{1 - \mathbb{P}_{\theta_1}^{N,T}(R_T^\sharp)}\left(\mathbb{P}_{\theta_0}^{N,T}(R_T^\sharp) - \mathbb{P}_{\theta_0}^{N,T}(R_T)\right).$$

Taking the 'liminf', we deduce

$$\begin{aligned}\liminf_{T\to\infty}\frac{1 - \mathbb{P}_{\theta_1}^{N,T}(R_T)}{1 - \mathbb{P}_{\theta_1}^{N,T}(R_T^\sharp)} \geq & 1 + \liminf_{T\to\infty}\frac{c_\alpha^\sharp(T)}{1 - \mathbb{P}_{\theta_1}^{N,T}(R_T^\sharp)}\left(\mathbb{P}_{\theta_0}^{N,T}(R_T^\sharp) - \alpha\right) \\ & - \limsup_{T\to\infty}\frac{c_\alpha^\sharp(T)}{1 - \mathbb{P}_{\theta_1}^{N,T}(R_T^\sharp)}\left(\mathbb{P}_{\theta_0}^{N,T}(R_T) - \alpha\right).\end{aligned}$$