Parameter Estimation for SPDEs

Sergey Lototsky
Department of Mathematics
University of Southern California

Statistical Inference

- From noisy observations to (useful) information.

- Random information: filtering.

- Deterministic information (θ): estimation ($\widehat{\theta}$).

- This talk:
  – Only Gaussian noise;
  – Source of problems is model rather than method: any consistent estimator will do.

Regular Model: LAN

Small Noise $y(t) = y_0 + \theta \int_0^t y(s)ds + \varepsilon W(t)$,
$\boxed{y_0 \neq 0}$.

Large Time $y(t) = -\theta \int_0^t y(s)ds + W(t),\ \theta > 0$.

$$\mathbf{MLE}: \widehat{\theta} = \int_0^T y(t)dy(t) \left(\int_0^T y^2(t)dt \right)^{-1}.$$

- *Inherent asymptotic*:

Must pass to a limit (ε or T).

- More: Ibragimov–Khasminskii (1981),

Yu. Kutoyants (1994, 2004).

Singular Models: no LAN

- All are different.
- Might admit a "magic device".
- *Computational* asymptotic.

Example: Noise as information.

$$y(t) = -\int_0^t y(t)dt + \theta W(t),\ \widehat{\theta} = \sqrt{\langle y \rangle_T / T}.$$

SPDEs

$(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$, $u = u(t,x)$, $0 \leq t \leq T$, $x \in G$,
G is d-dimensional;
$\{h_k,\, k \geq 1\}$ — CONS in $L_2(G)$,
W_k — independent Wiener processes.

$$du = (\mathcal{A}u + f)dt + \sum_{k\geq 1} (\mathcal{M}_k u + g_k) dW_k(t), \; u(0,x) = u_0(x).$$

- **Additive noise:** $\mathcal{M}_k = 0$ for all k.
- **Multiplicative noise:** otherwise.

- **Space-time white noise:**

$$dW(t,x) = \sum_{k\geq 1} h_k(x) dW_k(t).$$

$\mathcal{M}_k u = h_k \mathcal{M} u$,
$g_k(t,x) = h_k(x) g(t,x)$:

$$du = (\mathcal{A}u + f)dt + (\mathcal{M}u + g)dW(t,x).$$

Most difficult: multiplicative space-time white noise.

Diagonalizable model

$$du = (\theta\mathcal{A}_1 + \mathcal{A}_0)u dt + \sum_{k\geq 1}(\mathcal{M}_k u + g_k)dW_k(t)$$

$u(t,x) = \sum_{k=1}^{\infty} u_k(t)h_k(x)$, u_k **uncoupled**.

At the very least, need:

- $\theta = \theta(t)$;
- $\mathcal{A}_0 h_k = \rho_k h_k,\ \mathcal{A}_1 = \nu_k h_k$.

Then, two options:

(1) **Additive** *space-time* white noise:
$\mathcal{M}_k = 0,\ g_k = h_k$.

(2) **Multiplicative** *time* white noise:
$g_k = g_k(t) = 0$ (for simplicity), $\mathcal{M}_n h_k = \mu_k^{(n)} h_k$.

Idea: estimate θ from observations $(u_1(t), \ldots, u_N(t))$, $t \in [0, T]$. Then let $N \to \infty$.

General result: additive noise

$du = (\theta\mathcal{A}_1 + \mathcal{A}_0)u\,dt + dW(t,x),\ \ u(0,x) = 0.$
order$(\theta\mathcal{A}_1 + \mathcal{A}_0) = 2m$, order$(\mathcal{A}_1) = m_1$.

Key parameter: $\boxed{q = \frac{2(m_1 - m)}{d}.}$

$du_k(t) = (\theta\nu_k + \rho_k)u_k(t)dt + dW_k(t)$
(independent processes).

Different θ — absolutely continuous measures in $\mathrm{C}((0,T);\mathbb{R}^N)$.

Maximum Likelihood Estimator for θ:

$$\widehat{\theta}_N = \frac{\int_0^T (\Pi^N A_1 u, d\Pi^N u - \Pi^N A_0 u dt)_{L_2(G)}}{\int_0^T \|\Pi^N u(t)\|^2_{L_2(G)} dt}.$$

Π^N is the orthogonal projection on the span of $h_1, \ldots, h_N$.

Theorem (Huebner–Khasminskii–Rozovskii (1992), Huebner–Rozovskii (1995))

If $q \geq -1$, then
- SPDE measures are singular
- $\widehat{\theta}_N,\ N \to \infty$ is nice.

If $q < -1$, then SPDE measures are abs. continuous and $\widehat{\theta}_N,\ N \to \infty$ is biased.

What else for additive noise

Diagonalizable (bench-mark results):
- Bayesian: Bishwal (2002).
- Several parameters: Huebner (1997).
- $\theta = \theta(t)$: Hubner–Lototsky (2000; sieves and kernel).
- Discrete-time observations:

Piterbarg–Rozovskii (1997); $\boxed{q = \frac{2(m_1 - 2m)}{d} \geq -1.}$
- $\theta(t)$—random: Lototsky (2004).

Other:
- "Weakly" non-diagonalizable:

Lototsky–Rozovskii (1997), Lototsky (2001).
- Small noise: Huebner (1997),

Ibragimov–Khasminskii (1998,1999,...)

General result: multiplicative noise

(Ig. Cialenco (2006))

$du = (\theta \mathcal{A}_1 + \mathcal{A}_0)u\, dt + \mathcal{M}u dW(t)$,
$u(0, x) \neq 0$, non-random.

$du_k(t) = (\theta\nu_k + \rho_k)u_k(t)dt + \mu_k u_k dW(t)$:
Each u_k is a geometric Brownian motion and
$\sigma\Big(u_k(t), 0 \leq t \leq T\Big)$ is the same for all k.

Different θ — absolutely continuous measures in $\mathrm{C}((0,T);\mathbb{R})$ for each k, but not in $\mathrm{C}((0,T);\mathbb{R}^N)$.

Maximum Likelihood Estimator for θ:

$$\widehat{\theta}_N = \frac{1}{\nu_N T}\int_0^T \frac{du_N(t)}{u_N(t)} - \frac{\rho_N}{\nu_N}.$$

Asymptotic properties

Theorem

If $\lim\limits_{N\to\infty} \dfrac{\mu_N}{\nu_N} = 0$ then

$$\lim_{N\to\infty} \widehat{\theta}_N = \theta \ (\mathbb{P}\text{−(a.s.)}),$$

$$\lim_{N\to\infty} \frac{\sqrt{T}\nu_N}{\mu_N}\left(\widehat{\theta}_N - \theta\right) = \mathcal{N}(0,1)$$

in distribution and all moments.

Note: No d explicitly.

Proof

$$\widehat{\theta}_N = \theta + \frac{\mu_N}{\nu_N}\frac{W_T}{T}.$$

Special Effects, I

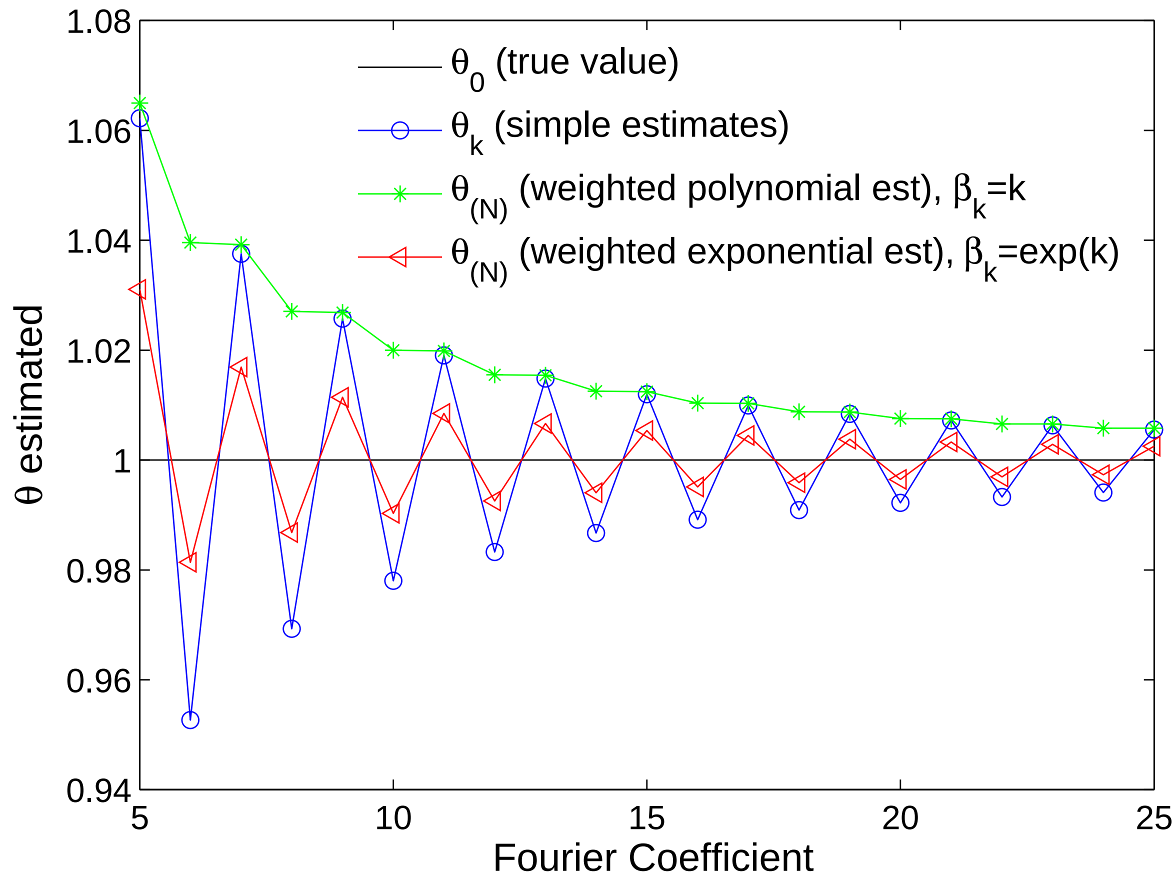
Weights can help:

$$\widetilde{\theta}_N = \frac{\sum_{k=1}^{N} \beta_k \widehat{\theta}_k}{\sum_{k=1}^{N} \beta_k}, \quad 0 \le \beta_k \uparrow +\infty \text{ fast enough.}$$

Example.

$\nu_k = -k^2,\ \rho_k = 0,\ \mu_k = (-1)^k\sqrt{k},\ \theta = 1$:

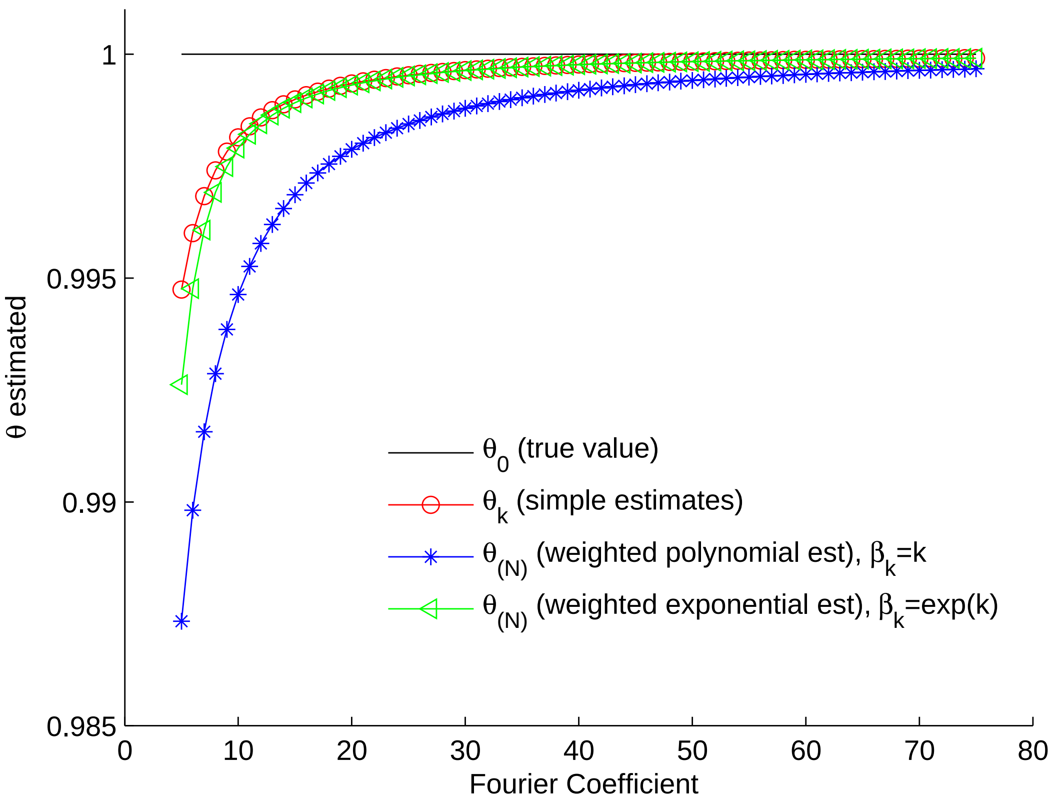
$du = \theta\, u_{xx} dt + \mathcal{M} u\, dW(t)$.



Special Effects, II

Equation with "no solution"

Example. $\nu_k = -k^4,\ \rho_k = 1,\ \mu_k = k^{2.5},\ \theta = 1$:
$du = (-\theta\, u_{xxxx} + u)dt + (-\Delta)^{5/4} u\, dW(t)$.



Note: Solution exists in a weighted Wiener chaos space (Lototsky–Rozovskii (2006)).

Summary

Statistical Inference for SPDEs

is fun!

Try it yourself.