Concentration of Measure for Stochastic Heat Equation

Andrey Sarantsev

University of California, Santa Barbara
Department of Statistics and Applied Probability

March 29, 2018

Joint work with Davar Khoshnevisan

Wasserstein Distance

Order $p \geq 1$.

Metric space $(\mathcal{X}, \rho)$.

Two probability measures $\mathbb{P}$ and $\mathbb{Q}$ on $\mathcal{X}$.

Wasserstein distance of order p:

$$\mathcal{W}_p(\mathbb{P}, \mathbb{Q}) = \inf \left[\mathbb{E}\rho^p(X, Y)\right]^{1/p}$$

where the infinum is taken over all couplings $(X, Y) \sim (\mathbb{P}, \mathbb{Q})$.

It is also called Kantorovich distance

Convergence in $\mathcal{W}_p$ = weak conv. + conv. of pth moments

Relative Entropy

Metric space $(\mathcal{X}, d)$.

Two probability measures $\mathbb{P}$ and $\mathbb{Q}$ on $\mathcal{X}$.

Relative Entropy or Kullback-Leibler divergence:

$$\mathcal{H}(\mathbb{Q} \mid \mathbb{P}) = \mathbb{E}^{\mathbb{P}}\left[\varphi \ln(\varphi)\right], \quad \varphi := \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}},$$

if $\mathbb{Q} \ll \mathbb{P}$ and ∞ otherwise.

This is a generalization of entropy of the distribution $(p_1, \ldots, p_n)$:

$$H(p) = -p_1 \ln p_1 - \ldots - p_n \ln p_n.$$

Talagrand Concentration Inequalities

We write $\mathbb{P} \in T_p(C)$ if for every $\mathbb{Q} \ll \mathbb{P}$ we have:

$$\mathcal{W}_p(\mathbb{P}, \mathbb{Q}) \le \sqrt{2C\mathcal{H}(\mathbb{Q} \mid \mathbb{P})}.$$

We say $\mathbb{P}$ satisfies transportation-cost information inequality or Talagrand concentration inequality of order p with constant C.

For $1 \le q < p$, $T_p(C)$ is stronger than $T_q(C)$.

Examples

Gaussian measure $\mathcal{N}(0, I_d)$ satisfies $T_2(C)$ with $C = 1$ on the space $\mathbb{R}^d$ with the Euclidean norm.

Brownian motion $W = (W(t), 0 \le t \le T)$ satisfies $T_2(C)$ with $C = T$ on $C[0, T]$ with the max-norm.

Pinsker inequality: Every $\mathbb{P}$ satisfies $T_1(C)$ with $C = 1/4$ with discrete metric $\rho(x, y) = 1$ for $x \neq y$.

Applications: Gaussian Tail Estimate

1-Lipschitz function $f : \mathcal{X} \to \mathbb{R}$: $|f(x) - f(y)| \le \rho(x, y)$.

Theorem (Marton, 1996)

If $\mathbb{P} \in T_1(C)$, then for any 1-Lipschitz function $f : \mathcal{X} \to \mathbb{R}$ with median $m(f)$ we have a Gaussian tail estimate

$$\mathbb{P}(|f - m(f)| \ge \delta) \le 2\exp\left(-\delta^2/(8C)\right), \quad \delta \ge 2\sqrt{2C\log 2}.$$

In fact, the converse is also true: Gaussian tail implies T_1.

Theorem (Bobkov, Gotze, 1999; Djellout, Guillin, Wu, 2004)

If $\mathbb{P}$ has first moment on $\mathcal{X}$, then $\mathbb{P} \in T_1(C)$ if and only if for all 1-Lipschitz functions $f : \mathcal{X} \to \mathbb{R}$ with $\int f \,\mathrm{d}\mathbb{P} = 0$, and all $a > 0$,

$$\int e^{af}\,\mathrm{d}\mathbb{P} \le e^{a^2C/2}.$$

Importance of Order 2

If $\mathbb{P}, \mathbb{Q} \in T_2(C)$, then $\mathbb{P} \times \mathbb{Q} \in T_2(C)$ on the product space $\mathcal{X} \times \mathcal{X}$ with distance

$$\rho_2((x_1, y_1), (x_2, y_2)) = \left[\rho^2(x_1, x_2) + \rho^2(y_1, y_2)\right]^{1/2}.$$

This property holds only for order 2. (Ledoux, 2001)

Poincare inequality $Var_\mu(f) \le C \int |\nabla f|^2 \, \mathrm{d}\mu$ follows from $T_2(C)$.

Finer Structure

Any probability measure with Gaussian tail satisfies T_1.

A Bernoulli measure on $\{0, 1\}$ does not satisfy T_p for $p > 1$.

Therefore, any measure with disconnected support (where components are at a positive distance from each other) does not satisfy T_p for $p > 1$.

Stochastic Differential Equations

The process $X = (X(t),\ t \ge 0)$ in $\mathbb{R}^1$ satisfies

$$\mathrm{d}X(t) = g(t, X(t))\,\mathrm{d}t + \sigma(t, X(t))\,\mathrm{d}W(t),\ \ X(0) = x.$$

Bounded σ: $|\sigma(t,x)| \le K_\sigma$. Lipschitz g and σ:

$$|g(t,x) - g(t,y)| \le L_g|x - y|, \quad |\sigma(t,x) - \sigma(t,y)| \le L_\sigma|x - y|,$$

Then X in $C[0, T]$ satisfies $T_2(C_T)$ with

$$C_T := 3K_\sigma^2 T \exp\left[3T^2L_g^2 + 12L_\sigma^2 T\right].$$

Similarly in $\mathbb{R}^d$, with Euclidean norm and Frobenius matrix norm. (Pal, 2012)

Proof Sketch

For every $\mathbb{Q} \ll \mathbb{P}$, there exist a process Z such that, under $\mathbb{Q}$,

$$\tilde{W}(t) = W(t) - \int_0^t Z(s)\,\mathrm{d}s, \quad \text{is a Brownian motion;}$$

$$\mathcal{H}(\mathbb{Q} \mid \mathbb{P}) = \frac{1}{2}\mathbb{E}^{\mathbb{Q}} \int_0^T Z^2(t)\,\mathrm{d}t.$$

Couple $(\mathbb{P}, \mathbb{Q})$ as follows under $\mathbb{Q}$:

$$\mathrm{d}X(t) = g(t, X(t))\,\mathrm{d}t + \sigma(t, X(t))Z(t)\,\mathrm{d}t + \sigma(t, X(t))\,\mathrm{d}\tilde{W}(t),$$
$$\mathrm{d}Y(t) = g(t, Y(t))\,\mathrm{d}t + \sigma(t, Y(t))\,\mathrm{d}\tilde{W}(t),\ X(0) = x.$$

Apply martingale inequalities and Gronwall's lemma to prove

$$\mathbb{E}^{\mathbb{Q}} \max_{0 \le t \le T} |X(t) - Y(t)|^2 \le C_T \cdot \mathbb{E}^{\mathbb{Q}} \int_0^T Z^2(t)\,\mathrm{d}t.$$

Stochastic Heat Equation

Unknown function: $u(t, x)$, $t \geq 0$, $0 \leq x \leq 1$.

$$\frac{\partial u}{\partial t} = \mathcal{L}u(t, x) + g(x, u(t, x)) + \sigma(x, u(t, x))\,\dot{W}(t, x).$$

Operator: $\mathcal{L} = \frac{1}{2}\frac{\partial^2}{\partial x^2}$ (Laplace in 1D)

Initial condition: $u\big|_{t=0} = u_0(x)$, deterministic.

Boundary condition: $u\big|_{x=0} = u\big|_{x=1} = 0$ (Dirichlet).

Space-time white noise: $W(t, x)$, "flickers" of independent noise at every point (t, x).

Mild Solution

Defined as a function $u(t,x)$, satisfying

$$u(t,x) = \int_{\mathbb{R}} u_0(y)\, G(t,x,y)\,\mathrm{d}y + \int_{\mathbb{R}} \int_0^t g(y,u(s,y))\, G(t-s,x,y)\,\mathrm{d}s\,\mathrm{d}y + \int_{\mathbb{R}} \int_0^t \sigma(y,u(s,y))\, G(t-s,x,y)\, W(\mathrm{d}s,\mathrm{d}y).$$

$G(t,x,y)$: Fundamental solution (heat kernel) of operator $\mathcal{L}$ with given boundary conditions; transition density of the corresponding stochastic process (absorbed Brownian motion on $[0,1]$)

Assumptions

Drift coefficient g: $|g(x, u) - g(x, v)| \le L|u - v|$.

Diffusion coefficient $\sigma \equiv 1$.

Solution exists and is unique, is a.s. continuous.

Works only in dimension 1: For spatial dimension 2 or more, the solution to the stochastic heat equation as a function does not even exist!

Main Result

Consider the max-norm on the space $C([0,T]\times[0,1])$ of continuous functions $u:[0,T]\times[0,1]\to\mathbb{R}$.

Theorem (Khoshnevisan, S, 2017)

The distribution of u satisfies $T_2(C)$ in the space $C([0,T]\times[0,1])$, with

$$C = 2G_T\exp(2L^2T^2),\quad G_T := \pi^{-1/2}\sqrt{T}.$$

Proof Sketch

Similarly to SDE, we represent $\mathbb{Q} \ll \mathbb{P}$ by Girsanov transformation: There exists a field $Z(t, x)$ such that, under $\mathbb{Q}$,

$$\tilde{W}(\mathrm{d}t, \mathrm{d}x) = W(\mathrm{d}t, \mathrm{d}x) - Z(t, x)\,\mathrm{d}t\,\mathrm{d}x,$$

is a space-time white noise. Moreover,

$$\mathcal{H}(\mathbb{Q} \mid \mathbb{P}) = \frac{1}{2}\mathbb{E}^{\mathbb{Q}} \int_0^T \int_{\mathbb{R}} Z^2(t, x)\,\mathrm{d}t\,\mathrm{d}x.$$

Couple $(\mathbb{P}, \mathbb{Q})$ via solutions of SPDE.

Other Operators

Similar results hold for other operators $\mathcal{L}$ instead of Laplacian:

- fractional Laplacian: α-stable Lévy process
- general second-order differential operator: stochastic differential equation

and different boundary conditions:

- Neumann: $u_x\big|_{x=0} = u_x\big|_{x=1} = 0$: reflected process
- periodic: $u\big|_{x=0} = u\big|_{x=1}$, $u_x\big|_{x=0} = u_x\big|_{x=1}$: process on the circle

$$\text{Need} \quad G_T := \sup_{0\le x\le 1}\int_0^T\int_0^1 G^2(t,x,y)\,\mathrm{d}y\,\mathrm{d}t.$$

Another Result

Instead of $C([0,T] \times [0,1])$, take $L^2([0,T] \times [0,1])$, with L^2-norm.

Diffusion σ is not necessarily 1, needs to be Lipschitz and bounded.

Another result, with a complicated constant C_T.