Integration by Parts Formula and Shift Harnack Inequality for Stochastic Equations

Feng-Yu Wang

(Beijing Normal University)

Workshop on Markov Processes and Related Topics

16-21 July 2012

Outline

- ♣ Driver's integration by parts formula
- ♣ Shift-Harnack inequality
- ♣ Backward coupling method
- ♣ Stochastic Hamiltonian systems
- ♣ Stochastic Functional differential equations
- ♣ Semi-linear SPDEs

Bismut's formula

- Bismut's formula (1984): Let P_t be the heat semigroup on a Riemannian manifold with curvature bounded below. For fixed $t > 0$ and vector v, one has

$$\nabla_v P_t f = \mathbb{E}\big[f(X_t)M_t\big], \quad f \in \mathscr{B}_b,$$

  - X_t: the Brownian motion on the manifold;
  - M_t is a random variable explicitly given by v and the curvature.

Let $p_t(x,y)$ be the heat kernel w.r.t. the volume measure. This formula implies

$$\nabla_v \log p_t(\cdot, y) = \mathbb{E}\big(M_t \big| X_t = y\big).$$

- Application: regularity of heat kernel in the first variable.

Driver's formula

- Driver's formula (1997): Let P_t be the heat semigroup on a Riemannian manifold with curvature bounded up to first order derivatives. For fixed $t > 0$ and smooth vector field V with compact support, one has

$$P_t(\nabla_V f)(x) = \mathbb{E}\big[f(X_t)N_t\big], \quad f \in C^1,$$

  - X_t: the Brownian motion on the manifold starting at x;
  - N_t is a random variable given by V, the curvature and its derivatives.

Then

$$\nabla_V \log p_t(x,\cdot)(y) = \mathbb{E}\Big(N_t \Big| X_t = y\Big)$$

provided $\mathrm{div} V(y) = 0$.

- Application: regularity of heat kernel in the second variable.

Bismut's formula has been well studied for SDEs and SPDEs, but much less is known on Driver's formula.

Shift-Harnack inequality

For simplicity, we only consider a diffusion semigroup P_t on $\mathbb{R}^d$. By Young inequality, Driver's formula

$$P_t(\nabla_e f) = \mathbb{E}[f(X_t)N_t]$$

for a vector $e \in \mathbb{R}^d$ implies

$$|P_t(\nabla_e f)| \le \delta\big\{P_t f \log f - (P_t f)\log P_t f\big\} + \delta \log \mathbb{E}\mathrm{e}^{N_t/\delta}, \quad \delta > 0$$

for any positive $f \in C_b^1$. Combining this with the following result one derives the shift-Harnack inequality from Driver's formula.

Shift-Harnack inequality

Proposition

Let P be a Markov operator on $\mathscr{B}_b(E)$ for some Banach space E. Let $e \in E$ and $\beta_e \in C((0,\infty)\times E;[0,\infty))$. Then

$$|P(\nabla_e f)| \le \delta\big\{P(f\log f) - (Pf)\log Pf\big\} + \beta_e(\delta,\cdot)Pf, \quad \delta > 0$$

holds for any positive $f \in C_b^1(E)$ if and only if

$$(Pf)^p(\cdot) \le \big(P\{f^p(re+\cdot)\}\big) \times \exp\left[\int_0^1 \frac{pr}{1+(p-1)s}\beta_e\Big(\frac{p-1}{r+r(p-1)s},\ \cdot + sre\Big)\mathrm{d}s\right]$$

holds for any positive $f \in \mathscr{B}_b(E), r \in (0,\infty)$ and $p > 1$.

Backward coupling

Differently from known Harnack inequalities, in this the shift-Harnack inequality the reference function rather than the variable is shifted. There are a number applications of the shift-Harnack inequality to heat kernel estimates and ultracontractivity property w.r.t. the Lebesgue measure.

♣ Classical coupling: Construct two processes starting at different points such that they move together as soon as possible. In particular, to establish Bismut's formula and Harnack inequalities, we need to ensure that they move together before a fixed time.

♣ Backward coupling: Construct two processes starting at a same point such that at a given time the difference of these processes reaches a given vector.

For the shift-Harnack inequality

Let P_t be a Markov semigroup. For fixed $T > 0$ and $x, e \in \mathbb{R}^d$, let X_t and Y_t be two process such that

(i) $X_0 = Y_0 = x$ and $Y_T = X_T + e$;

(ii) under probability $\mathbb{P}$ the process X_t is associated to P_t, i.e. $P_t f(x) = \mathbb{E}_{\mathbb{P}} f(X_t)$ for $f \in \mathscr{B}_b$;

(iii) under a weighted probability $\mathbb{Q} := R\mathbb{P}$, the process Y_t is associated to P_t.

Then for any $p > 1$ and positive $f \in \mathscr{B}_b$,

$$\begin{aligned}(P_T f)^p(x) &= \big(\mathbb{E}_{\mathbb{Q}} f(Y_T)\big)^p = \big(\mathbb{E}_{\mathbb{P}}[f(X_T+e)R]\big)^p \\ &\le \big(\mathbb{E}_{\mathbb{P}} f^p(X_T+e)\big)\big(\mathbb{E}_{\mathbb{P}} R^{p/(p-1)}\big)^{p-1} \\ &= \big(\mathbb{E}_{\mathbb{P}} R^{p/(p-1)}\big)^{p-1} P_T\{f^p(e+\cdot)\}(x).\end{aligned}$$

This gives a shift-Harnack inequality for P_T.

For Driver's formula

Suppose that we have constructed a family of couplings (X_t, Y_t^ε) and a family weighted probability measures $Q_\varepsilon := R_\varepsilon \mathbb{P},\ \varepsilon > 0$ such that

(i) $X_0 = Y_0^\varepsilon = x$ and $Y_T^\varepsilon = X_T + \varepsilon e$;

(ii) under probability $\mathbb{P}$ the process X_t is associated to P_t, i.e. $P_t f(x) = \mathbb{E}_\mathbb{P} f(X_t)$ for $f \in \mathscr{B}_b$;

(iii) under $\mathbb{Q}_\varepsilon$ the process Y_t^ε is associated to P_t;

(iv) $N_T := \lim_{\varepsilon\downarrow 0} \frac{1-R_\varepsilon}{\varepsilon}$ exists in $L^1(\mathbb{P})$.

Then for $f \in C_b^1$,

$$
\begin{aligned}
P_T(\nabla_e f)(x) &= P_T\Big(\lim_{\varepsilon\downarrow 0} \frac{f(\cdot+\varepsilon e) - f}{\varepsilon}\Big)(x)\\
&= \lim_{\varepsilon\downarrow 0} \frac{\mathbb{E}_\mathbb{P} f(X_T+\varepsilon e) - \mathbb{E}_{\mathbb{Q}_\varepsilon} f(Y_T^\varepsilon)}{\varepsilon}\\
&= \lim_{\varepsilon\downarrow 0} \mathbb{E}_\mathbb{P}\Big[\frac{f(X_T+\varepsilon e) - f(X_T+\varepsilon e)R_\varepsilon}{\varepsilon}\Big] = \mathbb{E}_\mathbb{P}[f(X_T)N_T].
\end{aligned}
$$

Stochastic Hamiltonian system

Consider the following degenerate stochastic differential equation for $(X(t), Y(t)) \in \mathbb{R}^{m+d} = \mathbb{R}^m \times \mathbb{R}^d (m, d \geq 1)$:

$$\begin{cases} \mathrm{d}X(t) = \big\{AX(t) + BY(t)\big\}\mathrm{d}t, \\ \mathrm{d}Y(t) = Z(X(t), Y(t))\mathrm{d}t + \sigma \mathrm{d}W(t), \end{cases}$$

- A: $m \times m$-matrix;
- B: $m \times d$-matrix;
- $Z \in C^1(\mathbb{R}^{m+d}; \mathbb{R}^d)$;
- σ: invertible $d \times d$-matrix;
- $W(t)$: d-dimensional Brownian motion.

The Hörmander condition holds if and only if

(H) There exists $0 \leq k \leq m-1$ such that $\mathrm{Rank}[B, AB, \cdots, A^k B] = m$.

Stochastic Hamiltonian system

We only consider Driver's integration by parts formula, since it is easier to establish the shift-Harnack. To this end, let $T > 0$ and $e = (e_1, e_2) \in \mathbb{R}^{m+d}$ be fixed. Assume that

$$\sup_{t\in[0,T]} \mathbb{E}\Big\{ \sup_{B(X(t),Y(t);r)} |\nabla Z|^2 \Big\} < \infty$$

holds for some $r > 0$. For non-negative $\phi \in C([0,T])$ with $\phi > 0$ in $(0,T)$, define

$$Q_\phi = \int_0^T \phi(t) \mathrm{e}^{(T-t)A} B B^* \mathrm{e}^{(T-t)A^*} \mathrm{d}t.$$

Then (H) implies that Q_ϕ is invertible.

Stochastic Hamiltonian system

Theorem

Let $\phi, \psi \in C^1([0,T])$ *such that* $\phi(0) = \phi(T) = 0, \phi > 0$ *in* $(0,T)$, *and* $\psi(T) = 1$, $\psi(0) = 0$, $\int_0^T \psi(t) \mathrm{e}^{(T-t)A} B \mathrm{d}t = 0$. *Moreover, let*

$$h(t) = \phi(t) B^* \mathrm{e}^{(T-t)A^*} Q_\phi^{-1} e_1 + \psi(t) e_2 \in \mathbb{R}^d,$$

$$\Theta(t) = \left(\int_0^t \mathrm{e}^{(t-s)A} B h(s) \mathrm{d}s,\ h(t) \right) \in \mathbb{R}^{m+d}, \quad t \in [0,T].$$

Then for any $f \in C_b^1(\mathbb{R}^{m+d})$,

$$P_T(\nabla_e f) = \mathbb{E}\big\{ f(X(T), Y(T)) N_T \big\} \text{ holds for}$$

$$N_T = \int_0^T \Big\langle \sigma^{-1}\big\{ h'(t) - \nabla_{\Theta(t)} Z(X(t), Y(t)) \big\},\ \mathrm{d}W(t) \Big\rangle.$$

Sketch of proof

For $\varepsilon \in (0,1]$, let $(X^\varepsilon(t), Y^\varepsilon(t))$ solve the equation

$$\begin{cases} \mathrm{d}X^\varepsilon(t) = \big\{AX^\varepsilon(t) + BY^\varepsilon(t)\big\}\mathrm{d}t, \\ \mathrm{d}Y^\varepsilon(t) = \sigma \mathrm{d}W(t) + \big\{Z(X(t), Y(t)) + \varepsilon h'(t)\big\}\mathrm{d}t \end{cases}$$

with $(X^\varepsilon(0), Y^\varepsilon(0)) = (X(0), Y(0))$. Then

$$\begin{cases} Y^\varepsilon(t) = Y(t) + \varepsilon h(t), \\ X^\varepsilon(t) = X(t) + \varepsilon \int_0^t \mathrm{e}^{(t-s)A} Bh(s)\mathrm{d}s. \end{cases}$$

In particular,

$$(X^\varepsilon(T), Y^\varepsilon(T)) = (X(T), Y(T)) + \varepsilon e$$

and

$$(*) \quad \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(X^\varepsilon(t), Y^\varepsilon(t))\Big|_{\varepsilon=0} = \Big(h(t), \int_0^t \mathrm{e}^{(t-s)A} Bh(s)\mathrm{d}s\Big).$$

Sketch of proof

Moreover, by the Girsanov theorem, $(X^\varepsilon(T), Y^\varepsilon(T))$ is associated to P_T under the weighted probability $R_\varepsilon \mathbb{P}$, where

$$R_\varepsilon = \exp\left[-\int_0^T \langle \sigma^{-1}\xi_\varepsilon(s), \mathrm{d}W(s)\rangle - \frac{1}{2}\int_0^T |\sigma^{-1}\xi_\varepsilon(s)|^2 \mathrm{d}s\right],$$

$$\xi_\varepsilon(s) := \varepsilon h'(s) + Z(s, X(s), Y(s)) - Z(s, X^\varepsilon(s), Y^\varepsilon(s))$$

Then the proof is completed since (*) and the assumption on ∇Z imply that

$$\lim_{\varepsilon \downarrow 0} \frac{1 - R_\varepsilon}{\varepsilon} = N_T$$

holds in $L^1(\mathbb{P})$.

Stochastic Functional differential equations

Let $\tau > 0$ be a fixed number, let $\mathcal{C} = C([-\tau, 0]; \mathbb{R}^d)$ be equipped with uniform norm $\|\cdot\|_\infty$. For a path $\gamma : [-\tau, \infty) \to \mathbb{R}^d$ and $t \geq 0$, $\gamma_t \in \mathcal{C}$ is given by

$$\gamma_t(s) = \gamma(t+s), \quad s \in [-\tau, 0].$$

Consider the stochastic functional equation

$$\mathrm{d}X(t) = b(X_t)\mathrm{d}t + \sigma \mathrm{d}W(t), \quad t \geq 0,$$

- $W(t)$: Brownaian motion on $\mathbb{R}^d$;
- $b \in C_b^1(\mathcal{C}; \mathbb{R}^d)$;
- σ: invertible $d \times d$-matrix.

The segment solution X_t is Markovian with semigroup P_t given by

$$P_t f(\xi) := \mathbb{E}\big(f(X_t)\big|X_0 = \xi\big), \quad \xi \in \mathcal{C}, t \geq 0, f \in \mathscr{B}_b(\mathcal{C}).$$

Stochastic Functional differential equations

Introduce the Carmeron-Martin space
$\mathbb{H} := \big\{h \in \mathcal{C} :\ \|h\|_{\mathbb{H}}^2 := \int_{-\tau}^0 |h'(t)|^2 \mathrm{d}t < \infty\big\}.$

Theorem

Let $T > \tau$ *and* $\eta \in \mathbb{H}$ *be fixed. For any* $\phi \in \mathscr{B}_b([0, T-\tau])$ *such that* $\int_0^{T-\tau} \phi(t)\mathrm{d}t = 1$, *let*

$$\Gamma(t) = \phi(t)\eta(-\tau)1_{[0,T-\tau]}(t) + \eta'(t-T)1_{(T-\tau,T]}(t),$$

$$\Theta(t) = \int_0^{t\vee 0} \Gamma(s)\mathrm{d}s, \quad t \in [-\tau, T].$$

Then for any $f \in C_b^1(\mathcal{C})$,

$$P_T(\nabla_\eta F) = \mathbb{E}\bigg(F(X_T)\int_0^T \Big\langle \sigma^{-1}\big(\Gamma(t) - \nabla_{\Theta_t} b(X_t)\big), \mathrm{d}W(t)\Big\rangle\bigg).$$

Sketch of proof

For fixed $\xi \in \mathcal{C}$, let $X(t)$ solve the equation for $X_0 = \xi$. For any $\varepsilon \in [0,1]$, let $X^\varepsilon(t)$ solve the equation

$$\mathrm{d}X^\varepsilon(t) = \{b(t, X_t) + \varepsilon\Gamma(t)\}\mathrm{d}t + \sigma\mathrm{d}W(t), \quad t \geq 0, X_0^\varepsilon = \xi.$$

Then

$$X_t^\varepsilon = X_t + \varepsilon\Theta_t, \quad t \in [0, T].$$

So,

$$X_T^\varepsilon = X_T + \varepsilon\eta, \quad \frac{\mathrm{d}}{\mathrm{d}\varepsilon}X_t^\varepsilon\Big|_{\varepsilon=0} = \Theta_t.$$

Let

$$R_\varepsilon = \exp\bigg[-\int_0^T \Big\langle \sigma^{-1}\big\{\varepsilon\Gamma(t) + b(X_t) - b(X_t^\varepsilon)\big\}, \mathrm{d}W(t)\Big\rangle - \frac{1}{2}\int_0^T \Big|\sigma^{-1}\big\{\varepsilon\Gamma(t) + b(X_t) - b(X_t^\varepsilon)\big\}\Big|^2\mathrm{d}t\bigg].$$

Sketch of the proof

By the Girsanov theorem, under the weighted probability $R_\varepsilon \mathbb{P}$, X_t^ε is associated to P_t. Then the proof is completed by noting that

$$\lim_{\varepsilon \downarrow 0} \frac{1 - R_\varepsilon}{\varepsilon} = \int_0^T \left\langle \sigma^{-1}\big(\Gamma(t) - \nabla_{\Theta_t} b(X_t)\big), \mathrm{d}W(t) \right\rangle$$

holds in $L^1(\mathbb{P})$.

Semi-linear SPDEs

Let $(H, \langle\cdot,\cdot\rangle)$ be a real separable Hilbert space, $(W(t))_{t\geq 0}$ a cylindrical Wiener process on H. Consider the semi-linear equation on H:

$$\mathrm{d}X(t) = \{AX(t) + b(X(t))\}\mathrm{d}t + \sigma \mathrm{d}W(t),$$

- $(A, \mathscr{D}(A))$: linear operator on H generating a contractive C_0-semigroup such that $\int_0^1 \|\mathrm{e}^{sA}\|_{HS}^2 \mathrm{d}s < \infty$;
- $b \in C_b^1(H; H)$;
- σ: bounded and invertible operator on H.

Semi-linear SPDEs

The equation has a unique mild solution, and let P_t be the associated semigroup.

Theorem

Let $T > 0$ and $e \in H$ be fixed. For any $f \in C_b^1(H)$,

$$P_T(\nabla_e f) = \mathbb{E}\bigg(f(X(T)) \int_0^T \Big\langle \sigma^{-1}\Big(\frac{e}{T} - \nabla_{\frac{te}{T}} b(X(t))\Big), \mathrm{d}W(t)\Big\rangle\bigg).$$

Sketch of the proof

For $\varepsilon > 0$, let $X^\varepsilon(t)$ solve the equation

$$\mathrm{d}X^\varepsilon(t) = \Big\{b(t, X(t)) + \frac{\varepsilon}{T}e\Big\}\mathrm{d}t + \sigma \mathrm{d}W(t),\ \ t \ge 0, X^\varepsilon(0) = X(0).$$

Then

$$X^\varepsilon(t) = X(t) + \frac{t\varepsilon}{T}e,\ \ t \in [0, T].$$

In particular, $X^\varepsilon(T) = X(T) + \varepsilon e$. Moreover, let

$$R_\varepsilon = \exp\Big[-\int_0^T \Big\langle \sigma^{-1}\Big\{\frac{\varepsilon e}{T} + b(X(t)) - b(X^\varepsilon(t))\Big\}, \mathrm{d}W(t)\Big\rangle - \frac{1}{2}\int_0^T \Big|\sigma^{-1}\Big\{\frac{\varepsilon e}{T} + b(X(t)) - b(X^\varepsilon(t))\Big\}\Big|^2 \mathrm{d}t\Big].$$

Then under weighted probability $R_\varepsilon \mathbb{P}$ the process $X^\varepsilon(t)$ is associated to P_t. The proof is done.

Final remark

The argument may work also for many other models, for instance

- SPDEs;
- Multiplicative noises;
- Degenerate SFDEs;
- SDEs driven by Lévy processes, fractional BM;

....

Thank You