A Hamilton-Jacobi equation approach to Large Deviation of Markov Processes

Jin Feng (joint work with Tom Kurtz, U. Wisconsin)

Department of Mathematics, University of Kansas, Lawrence, KS 66045, USA

June 5th, Kac Lecture at Utrech University, the Netherlands

Essence of my talk

- $$a \vee b = \lim_{n \to \infty} \frac{1}{n} \log(e^{na} + e^{nb})$$

Essence of my talk

- $$a \vee b = \lim_{n\to\infty} \frac{1}{n} \log(e^{na} + e^{nb})$$
- Left hand: calculus of variation

Essence of my talk

- $$a \vee b = \lim_{n\to\infty} \frac{1}{n} \log(e^{na} + e^{nb})$$
- Left hand: calculus of variation
- Right hand: probability

Essence of my talk

- $$a \vee b = \lim_{n\to\infty} \frac{1}{n} \log(e^{na} + e^{nb})$$
- Left hand: calculus of variation
- Right hand: probability
- A poor man's version of "everything about Large Deviation"

Principles of Laplace v.s. Large Deviation

- The Laplace principle

$$\sup_{z \in S} f(z) = \lim_{n \to \infty} \frac{1}{n} \log \int_S e^{nf(z)} P(dz).$$

Principles of Laplace v.s. Large Deviation

- The Laplace principle

$$\sup_{z\in S} f(z) = \lim_{n\to\infty} \frac{1}{n} \log \int_S e^{nf(z)} P(dz).$$

- What if $P(dx) := P_n(dx) \sim e^{-nI(x)}\pi(dx)$?

Principles of Laplace v.s. Large Deviation

- The Laplace principle

$$\sup_{z\in S} f(z) = \lim_{n\to\infty} \frac{1}{n} \log \int_S e^{nf(z)} P(dz).$$

- What if $P(dx) := P_n(dx) \sim e^{-nI(x)}\pi(dx)$?
- LDP : $P(X_n \in dx) \sim Z_n^{-1} e^{-nI(x)}\pi(dx)$.

Principles of Laplace v.s. Large Deviation

- The Laplace principle

$$\sup_{z\in S} f(z) = \lim_{n\to\infty} \frac{1}{n} \log \int_S e^{nf(z)} P(dz).$$

- What if $P(dx) := P_n(dx) \sim e^{-nI(x)}\pi(dx)$?
- LDP : $P(X_n \in dx) \sim Z_n^{-1} e^{-nI(x)}\pi(dx)$.
- LDP, a functional level re-formulation,

$$\begin{aligned}\lim_n \frac{1}{n} \log E[e^{nf(X_n)}] &= \lim_n \frac{1}{n} \log \int e^{n(f-I)(x)}\pi(dx) \\ &= \sup_{x\in S}\{f(x) - I(x)\}, \quad \forall f.\end{aligned}$$

Pressure-Entropy duality, the nature of large deviation

- Entropy-Pressure duality à la Gibbs:

$$\log \int_S e^{g(z)} P(dz) = \sup_Q \{\langle g, Q \rangle - R(Q \| P)\}$$

Pressure-Entropy duality, the nature of large deviation

- Entropy-Pressure duality à la Gibbs:

$$\log \int_S e^{g(z)} P(dz) = \sup_Q \{\langle g, Q\rangle - R(Q\|P)\}$$

- Renormalize: let $dP^g := \frac{e^g}{Z_g} dP$, then

$$\sup_Q \{\langle g, Q\rangle - R(Q\|P)\} = -\inf_Q R(Q\|P^g) + \log Z_g.$$

Large Deviation - Rigorous definition

Let $\{X_n : n = 1, 2, \ldots\}$ be metric space S-valued r.v.s.

- There exists lower semicontinuous $I : S \mapsto [0, +\infty]$ satisfying

$$
\begin{aligned}
-\inf_{x \in A^\circ} I(x) \;\; &\leq \;\; \liminf_n \frac{1}{n} \log P(X_n \in A^\circ) \\
&\leq \limsup_n \frac{1}{n} P(X_n \in \bar{A}) \leq -\inf_{x \in \bar{A}} I(x).
\end{aligned}
$$

If rate function(or action functional) I has compact level set, I is said to be good.

Large Deviation - Rigorous definition

Let $\{X_n : n = 1, 2, \ldots\}$ be metric space S-valued r.v.s.

- There exists lower semicontinuous $I : S \mapsto [0, +\infty]$ satisfying

$$
\begin{aligned}
-\inf_{x \in A^\circ} I(x) \quad &\leq \quad \liminf_n \frac{1}{n} \log P(X_n \in A^\circ) \\
&\leq \limsup_n \frac{1}{n} P(X_n \in \bar{A}) \leq -\inf_{x \in \bar{A}} I(x).
\end{aligned}
$$

If rate function(or action functional) I has compact level set, I is said to be good.

Exponential Tightness

- For each $a > 0$, there exists a compact $K_a \subset\subset S$ such that

$$\limsup_n \frac{1}{n} \log P(X_n \notin K_a) \leq -a.$$

Exponential Tightness

- For each $a > 0$, there exists a compact $K_a \subset\subset S$ such that

$$\limsup_n \frac{1}{n} \log P(X_n \notin K_a) \leq -a.$$

- Equivalently (S complete separable metric space), for each $a > 0$ and $\epsilon > 0$, there exists $K_{a,\epsilon} \subset\subset S$,

$$\limsup_n \frac{1}{n} \log P(X_n \notin K_{a,\epsilon}^{\epsilon}) \leq -a$$

The Laplace Principle

$\{X_n : n = 1, 2, \ldots\}$ is S-valued (S is Polish)

- (Varadhan) $\{X_n : n = 1, 2, \ldots\}$ satisfies LDP with good rate function I, then

$$\lim_n \frac{1}{n} \log E[e^{nf(X_n)}] = \sup_{x \in S}\{f(x) - I(x)\}.$$

The Laplace Principle

$\{X_n : n = 1, 2, \ldots\}$ is S-valued (S is Polish)

- (Varadhan) $\{X_n : n = 1, 2, \ldots\}$ satisfies LDP with good rate function I, then

$$\lim_n \frac{1}{n} \log E[e^{nf(X_n)}] = \sup_{x \in S}\{f(x) - I(x)\}.$$

- (Bryc) Let $\{X_n : n = 1, 2, \ldots\}$ be exponentially tight,

$$\Lambda(f) = \lim_n \frac{1}{n} \log E[e^{nf(X_n)}], \quad f \in D \subset C_b(S).$$

  Then $\{X_n : n = 1, 2, \ldots\}$ satisfies LDP with good rate function $I(x) = \sup_{f \in C_b(S)}\{f(x) - \Lambda(f)\}$.

Smaller class of test functions $D \subset C_b(S)$

Definition: $D \subset C_b(S)$ is **rate function determining** if

$$\sup_{f \in D}\{f(x) - \Lambda(f)\} = \sup_{f \in C_b(S)}\{f(x) - \Lambda(f)\}.$$

- Useful test functions are of the form $-md(x, \cdot)$ ...

In the following talk,

- $X_n(\cdot)$ is a sequence of Markov processes, hence

$$S := D_E[0, \infty).$$

Outline of main result:

- Basic setup

Outline of main result:

- Basic setup

1. E_n, E are complete separable metric spaces;

Outline of main result:

- Basic setup
  1. E_n, E are complete separable metric spaces;
  2. for each n, $Y_n(\cdot)$ is E_n-valued Markov Process with (MGP) generator A_n;

Outline of main result:

- Basic setup
  1. E_n, E are complete separable metric spaces;
  2. for each n, $Y_n(\cdot)$ is E_n-valued Markov Process with (MGP) generator A_n;
  3. $\eta_n : E_n \mapsto E$ and $X_n(t) := \eta_n(Y_n(t)) = \eta_n \circ Y_n(t)$ is E-valued process;

Outline of main result:

- Basic setup
  1. E_n, E are complete separable metric spaces;
  2. for each n, $Y_n(\cdot)$ is E_n-valued Markov Process with (MGP) generator A_n;
  3. $\eta_n : E_n \mapsto E$ and $X_n(t) := \eta_n(Y_n(t)) = \eta_n \circ Y_n(t)$ is E-valued process;
- Main questions:

Outline of main result:

- Basic setup
  1. E_n, E are complete separable metric spaces;
  2. for each n, $Y_n(\cdot)$ is E_n-valued Markov Process with (MGP) generator A_n;
  3. $\eta_n : E_n \mapsto E$ and $X_n(t) := \eta_n(Y_n(t)) = \eta_n \circ Y_n(t)$ is E-valued process;
- Main questions:
  1. Can we derive large deviation behavior of $\{X_n(\cdot) : n = 1, 2, \ldots\}$ based on information in sequence of operators $\{A_n : n = 1, 2, \ldots\}$?

Outline of main result:

- Basic setup
  1. E_n, E are complete separable metric spaces;
  2. for each n, $Y_n(\cdot)$ is E_n-valued Markov Process with (MGP) generator A_n;
  3. $\eta_n : E_n \mapsto E$ and $X_n(t) := \eta_n(Y_n(t)) = \eta_n \circ Y_n(t)$ is E-valued process;
- Main questions:
  1. Can we derive large deviation behavior of $\{X_n(\cdot) : n = 1, 2, \ldots\}$ based on information in sequence of operators $\{A_n : n = 1, 2, \ldots\}$?
  2. What is its connection with variational problems?

Large deviation for Markov processes $\{Y_n(\cdot) : n = 1, 2, \ldots\}$

Take the case $E = E_n$ and $\eta_n = I$, so
$X_n(t) := \eta_n(Y_n(t)) = Y_n(t)$:

- One step transition determines everything

$$(V_n(t)f)(x) := \frac{1}{n} \log E[e^{nf(X_n(t))} | X_n(0) = x].$$

Large deviation for Markov processes $\{Y_n(\cdot) : n = 1, 2, \ldots\}$

Take the case $E = E_n$ and $\eta_n = I$, so
$X_n(t) := \eta_n(Y_n(t)) = Y_n(t)$:

- One step transition determines everything

$$(V_n(t)f)(x) := \frac{1}{n} \log E[e^{nf(X_n(t))}|X_n(0) = x].$$

- V_n is a semigroup

$$V_n(t+s) = V_n(t)V_n(s).$$

Large deviation for Markov processes $\{Y_n(\cdot) : n = 1, 2, \ldots\}$

Take the case $E = E_n$ and $\eta_n = I$, so
$X_n(t) := \eta_n(Y_n(t)) = Y_n(t)$:

- One step transition determines everything

$$(V_n(t)f)(x) := \frac{1}{n} \log E[e^{nf(X_n(t))} | X_n(0) = x].$$

- V_n is a semigroup

$$V_n(t+s) = V_n(t)V_n(s).$$

- V_n is the Nisio semigroup by entropy-pressure dual

$$V_n(t)f(x) = \sup_{Q_n^x} E^{Q_n^x}[f(X_n(t)) - \log \frac{dQ_n^x}{dP_n^x}|_{\mathcal{F}_t^n}(X_n(\cdot))]$$

Large deviation for Markov processes $\{Y_n(\cdot) : n = 1, 2, \ldots\}$

- Therefore convergence of V_n would imply LDP (needs exponential tightness, why?)

Large deviation for Markov processes $\{Y_n(\cdot) : n = 1, 2, \ldots\}$

- Therefore convergence of V_n would imply LDP (needs exponential tightness, why?)
- What happened infinitesimally?

$$\lim_{t \to 0+} \frac{1}{t}(V_n(t)f(x) - f(x)) = \frac{1}{n} e^{-nf} A_n e^{nf}(x) := H_n f(x)$$

Large deviation for Markov processes $\{Y_n(\cdot) : n = 1, 2, \ldots\}$

- Therefore convergence of V_n would imply LDP (needs exponential tightness, why?)
- What happened infinitesimally?

$$\lim_{t \to 0+} \frac{1}{t}(V_n(t)f(x) - f(x)) = \frac{1}{n} e^{-nf} A_n e^{nf}(x) := H_n f(x)$$

- Here is a program for LDP:

$$1.(H_n \to H) + 2.(\text{ exp. tight.}) + 3.(\text{HJ equations}) \Rightarrow \text{LDP.}$$

Large deviation for Markov processes $\{Y_n(\cdot) : n = 1, 2, \dots\}$

- Rate function? $V_n \to V$ is Nisio semigroup convergence and

$$R(Q_n^x \| P_n^x) \to \int_0^t L(x(s), u(s)) ds := I(x(\cdot))$$

and

$$V(t) f(x) = \sup_u \{ f(x(t)) - \int_0^t L(x(s), u(s)) ds : x(0) = x \}$$

Large deviation for Markov processes $\{Y_n(\cdot) : n = 1, 2, \ldots\}$

- Rate function? $V_n \to V$ is Nisio semigroup convergence and

$$R(Q_n^x \| P_n^x) \to \int_0^t L(x(s), u(s))ds := I(x(\cdot))$$

and

$$V(t)f(x) = \sup_u \{f(x(t)) - \int_0^t L(x(s), u(s))ds : x(0) = x\}$$

- In general, relaxed control (Young measure) replacing u...
One can completely formulate a program on co-tangent space (Hamiltonian equations) until the limit for representation ..

Entropy-Pressure Dual at infinitesimal level

Fleming- Sheu's logarithmic transformation :

- $Hf := e^{-f}Ae^{f}$;

Entropy-Pressure Dual at infinitesimal level

Fleming- Sheu's logarithmic transformation :

- $Hf := e^{-f} A e^{f}$;
- $A^g f := e^{-g} A(f e^{g}) - (e^{-g} f) A e^{g}$;

Entropy-Pressure Dual at infinitesimal level

Fleming- Sheu's logarithmic transformation :

- $Hf := e^{-f} A e^{f}$;
- $A^g f := e^{-g} A(f e^g) - (e^{-g} f) A e^g$;
- $Lg := A^g g - Hg = e^{-g} A(g e^g) - (e^{-g} g) A e^g - e^{-g} A e^g$;

Entropy-Pressure Dual at infinitesimal level

Fleming- Sheu's logarithmic transformation :

- $Hf := e^{-f}Ae^{f}$;
- $A^g f := e^{-g}A(fe^g) - (e^{-g}f)Ae^g$;
- $Lg := A^g g - Hg = e^{-g}A(ge^g) - (e^{-g}g)Ae^g - e^{-g}Ae^g$;
- Then, by maximum principle,

$$
\begin{aligned}
Hf(x) &= \sup_g \left(A^g f(x) - Lg(x)\right) \\
Lg(x) &= \sup_f \left(A^g f(x) - Hf(x)\right).
\end{aligned}
$$

What's going on?

- Normalizing transition measure (g as a control)

$$P^g(t;x,dy) := \frac{e^{g(y)}P(t;x,dy)}{\int e^{g(z)}P(t;x,dz)}.$$

What's going on?

- Normalizing transition measure (g as a control) $P^g(t;x,dy) := \frac{e^{g(y)}P(t;x,dy)}{\int e^{g(z)}P(t;x,dz)}$.
-

$$
\begin{aligned}
A^g f(x) &:= \lim_{t\to 0+} E^{P^g_x}[f(X(t)) - f(X(0))] \\
&= \lim_{t\to 0+} \frac{E^x[e^{g(X(t))}(f(X(t)) - f(X(0)))]}{E^x[e^{g(X(t))}]} \\
&= e^{-g}A(fe^g)(x) - (e^{-g}f)Ae^g(x); \\
Lg(x) &:= \lim_{t\to 0+} R(P^g(t;x,\cdot)\|P(t;x,\cdot)) = \ldots \\
&= A^g g(x) - Hg(x).
\end{aligned}
$$

Diffusion Example

- $A_n = \frac{1}{\sqrt{n}}\Delta + b \cdot \nabla,$

Diffusion Example

- $A_n = \frac{1}{\sqrt{n}}\Delta + b \cdot \nabla,$
- $H_n f = \frac{1}{\sqrt{n}}\Delta f + b\nabla f + \frac{1}{2}|\nabla f|^2,$

Diffusion Example

- $A_n = \frac{1}{\sqrt{n}}\Delta + b \cdot \nabla,$
- $H_n f = \frac{1}{\sqrt{n}}\Delta f + b\nabla f + \frac{1}{2}|\nabla f|^2,$
- $L_n g = \frac{1}{2}|\nabla g|^2,$

Diffusion Example

- $A_n = \frac{1}{\sqrt{n}}\Delta + b \cdot \nabla,$
- $H_n f = \frac{1}{\sqrt{n}}\Delta f + b\nabla f + \frac{1}{2}|\nabla f|^2,$
- $L_n g = \frac{1}{2}|\nabla g|^2,$
- $A_n^g f(x) = \frac{1}{\sqrt{n}}\Delta f(x) + (b + \nabla g) \cdot \nabla f(x)$

Diffusion Example

- $A_n = \frac{1}{\sqrt{n}}\Delta + b \cdot \nabla,$
- $H_n f = \frac{1}{\sqrt{n}}\Delta f + b\nabla f + \frac{1}{2}|\nabla f|^2,$
- $L_n g = \frac{1}{2}|\nabla g|^2,$
- $A_n^g f(x) = \frac{1}{\sqrt{n}}\Delta f(x) + (b + \nabla g) \cdot \nabla f(x)$
- $H_n f(x) = \sup_g \{A_n^g f(x) - L_n g(x)\}$

Diffusion Example

- $A_n = \frac{1}{\sqrt{n}}\Delta + b \cdot \nabla,$
- $H_n f = \frac{1}{\sqrt{n}}\Delta f + b\nabla f + \frac{1}{2}|\nabla f|^2,$
- $L_n g = \frac{1}{2}|\nabla g|^2,$
- $A_n^g f(x) = \frac{1}{\sqrt{n}}\Delta f(x) + (b + \nabla g) \cdot \nabla f(x)$
- $H_n f(x) = \sup_g \{A_n^g f(x) - L_n g(x)\}$
- Embedding (reduction) of control space to R^d:

$$H_n f(x) = \sup_{u \in R^d} \{\frac{1}{\sqrt{n}}\Delta f(x) + (b(x) + u) \cdot \nabla f(x) - \frac{1}{2}|u|^2\}.$$

A pure jump example

- Take $Af(x) := \lambda(x) \int (f(y) - f(x))\eta(x, dy)$.

A pure jump example

- Take $Af(x) := \lambda(x) \int (f(y) - f(x))\eta(x, dy)$.
- w.l.o.g., all Markov processes can be approximated by such form by Yosida appro. $A_\epsilon = A(I - \epsilon A)^{-1}$.

A pure jump example

- Take $Af(x) := \lambda(x) \int (f(y) - f(x))\eta(x, dy)$.
- w.l.o.g., all Markov processes can be approximated by such form by Yosida appro. $A_\epsilon = A(I - \epsilon A)^{-1}$.
- Infinitesimal change of measure $\eta^g(x, dy) := \frac{e^{g(y)}}{e^{g(x)}}\eta(x, dy)$.

A pure jump example

- Take $Af(x) := \lambda(x)\int(f(y) - f(x))\eta(x, dy)$.
- w.l.o.g., all Markov processes can be approximated by such form by Yosida appro. $A_\epsilon = A(I - \epsilon A)^{-1}$.
- Infinitesimal change of measure $\eta^g(x, dy) := \frac{e^{g(y)}}{e^{g(x)}}\eta(x, dy)$.
- $Hf(x) = \lambda(x)\int(e^{f(y)-f(x)} - 1)\eta(x, dy)$.

A pure jump example

- Take $Af(x) := \lambda(x) \int (f(y) - f(x))\eta(x, dy)$.
- w.l.o.g., all Markov processes can be approximated by such form by Yosida appro. $A_\epsilon = A(I - \epsilon A)^{-1}$.
- Infinitesimal change of measure
  $\eta^g(x, dy) := \frac{e^{g(y)}}{e^{g(x)}}\eta(x, dy)$.
- $Hf(x) = \lambda(x) \int (e^{f(y)-f(x)} - 1)\eta(x, dy)$.
- $A^g f(x) = \lambda(x) \int (f(y) - f(x))\eta^g(x, dy)$.

A pure jump example

- Take $Af(x) := \lambda(x)\int(f(y)-f(x))\eta(x,dy)$.
- w.l.o.g., all Markov processes can be approximated by such form by Yosida appro. $A_\epsilon = A(I-\epsilon A)^{-1}$.
- Infinitesimal change of measure
  $\eta^g(x,dy) := \frac{e^{g(y)}}{e^{g(x)}}\eta(x,dy)$.
- $Hf(x) = \lambda(x)\int(e^{f(y)-f(x)}-1)\eta(x,dy)$.
- $A^g f(x) = \lambda(x)\int(f(y)-f(x))\eta^g(x,dy)$.
- $Lg(x) = \lambda(x)\int\left((g(y)-g(x))\frac{e^{g(y)}}{e^{g(x)}} - \frac{e^{g(y)}}{e^{g(x)}} + 1\right)\eta(x,dy)$
  $= \lambda(x)R(\eta^g(x,\cdot)\|\eta(x,\cdot))$.

Nonlinear semigroup via viscosity solution language

Let

$$(I - \alpha H_n) f_n = h_n, \quad (I - \alpha H) f = h.$$

- Suppose we know that, if $h_n \to h$ and $H \subset \lim_n H_n$ (what sense?), then $f_n = (I - \alpha H_n)^{-1} h_n \to f = (I - \alpha H)^{-1} h$.

Nonlinear semigroup via viscosity solution language

Let

$$(I-\alpha H_n)f_n = h_n, \quad (I-\alpha H)f = h.$$

- Suppose we know that, if $h_n \to h$ and $H \subset \lim_n H_n$ (what sense?), then $f_n = (I-\alpha H_n)^{-1}h_n \to f = (I-\alpha H)^{-1}h$.
- Then

$$\lim_{n\to\infty} V_n(t)h_n = \lim_n \lim_{m\to\infty} (I - \frac{t}{m}H_n)^{-m}h_n$$

$$= \lim_m \lim_n (I - \frac{t}{m}H_n)^{-m}h_n = \lim_m (I - \frac{t}{m}H_n)^{-m}h$$

$$= V(t)h$$

What is viscosity solution in this context?

Problem: E is just a Polish space.
Solution: Use (nonlinear) maximum principle.

- $f, g \in M(E)$,

$$\begin{aligned} [f,g]_+ &:= \inf_{\epsilon>0} \epsilon^{-1}(\|f+\epsilon g\| - \|f\|), \\ [f,g]_- &:= \sup_{\epsilon<0} \epsilon^{-1}(\|f+\epsilon g\| - \|f\|). \end{aligned}$$

What is viscosity solution in this context?

Problem: E is just a Polish space.
Solution: Use (nonlinear) maximum principle.

- $f, g \in M(E)$,

$$[f,g]_+ := \inf_{\epsilon>0} \epsilon^{-1}(\|f+\epsilon g\| - \|f\|),$$
$$[f,g]_- := \sup_{\epsilon<0} \epsilon^{-1}(\|f+\epsilon g\| - \|f\|).$$

- $H \subset M(E) \times M(E)$ has maximum principle:

$$[f_1 - f_2, g_1 - g_2]_\pm \leq 0, \qquad \forall (f_i, g_i) \in H;$$

Viscosity solution

For $(I - \alpha H)f = h$.

- $[f - f_0, Hf - Hf_0]_- \leq 0$.

Viscosity solution

For $(I - \alpha H)f = h$.

- $[f - f_0, Hf - Hf_0]_- \leq 0$.
- If $\sup_x (f - f_0) \geq 0$, $\Rightarrow$ sub-solution ...

Viscosity solution

For $(I - \alpha H)f = h$.

- $[f - f_0, Hf - Hf_0]_- \leq 0$.
- If $\sup_x (f - f_0) \geq 0$, $\Rightarrow$ sub-solution ...
- If $\sup_x (f_0(x) - f(x)) \geq 0$, $\Rightarrow$ super-solution...

Viscosity solution

For $(I - \alpha H)f = h$.

- $[f - f_0, Hf - Hf_0]_- \leq 0$.
- If $\sup_x (f - f_0) \geq 0$, $\Rightarrow$ sub-solution ...
- If $\sup_x (f_0(x) - f(x)) \geq 0$, $\Rightarrow$ super-solution...
- An exercise: Write down the precise expression ...

Barles-Perthame half relaxed limit, an abstract new proof

- Consider, for COMPACT E,

$$(I - \alpha H_n) f_n = h_n, \quad (I - \alpha H) f = h.$$

Q: Suppose $h_n \to h$ and $H_n \to H$, does $f_n \to f$?

Barles-Perthame half relaxed limit, an abstract new proof

- Consider, for COMPACT E,

$$(I - \alpha H_n) f_n = h_n, \quad (I - \alpha H) f = h.$$

Q: Suppose $h_n \to h$ and $H_n \to H$, does $f_n \to f$?

- Minty's device method by L.C.Evans:

$$[f_n - f_0, H_n f_n - H_n f_0]_- \le 0 \Rightarrow [f - f_0, Hf - Hf_0]_- \le 0.$$

Barles-Perthame half relaxed limit, an abstract new proof

- Consider, for COMPACT E,

$$(I - \alpha H_n) f_n = h_n, \quad (I - \alpha H) f = h.$$

Q: Suppose $h_n \to h$ and $H_n \to H$, does $f_n \to f$?

- Minty's device method by L.C.Evans:

$$[f_n - f_0, H_n f_n - H_n f_0]_- \le 0 \Rightarrow [f - f_0, Hf - Hf_0]_- \le 0.$$

- Barles and Perthame:

$$\bar{f}(x) := \limsup_{n,y \to x} f_n(y), \quad \underline{f} := \liminf_{n,y \to x} f_n(y)$$

$$(I - \alpha H)\bar{f} \le h, \quad (I - \alpha H)\underline{f} \ge h.$$

$f = \overline{f} = \underline{f}$? Answer: Comparison Principle.

The noncompact E case

- Idea: Use concentration on compact set property of probability measures:

The noncompact E case

- Idea: Use concentration on compact set property of probability measures:
- Instead of

$$\|f_n - f_0\| \leq \|f_n - \alpha H_n f_n - (f_0 - \alpha H_n f_0)\|,$$

The noncompact E case

- Idea: Use concentration on compact set property of probability measures:
- Instead of

$$\|f_n - f_0\| \leq \|f_n - \alpha H_n f_n - (f_0 - \alpha H_n f_0)\|,$$

- we want $\forall\ K \subset\subset E$ and $\epsilon > 0$, $\exists\ K_\epsilon \subset\subset E$, s.t.

$$\begin{aligned} & \sup_{x \in K_0} |f_n(x) - f_0(x)| \\ \leq\ & \epsilon + \sup_{y \in K_\epsilon} |f_n(y) - \alpha H_n f_n(y) - (f_0(y) - \alpha H_n f_0(y))|. \end{aligned}$$

The noncompact E case

- Idea: Use concentration on compact set property of probability measures:
- Instead of

$$\|f_n - f_0\| \le \|f_n - \alpha H_n f_n - (f_0 - \alpha H_n f_0)\|,$$

- we want $\forall\ K \subset\subset E$ and $\epsilon > 0$, $\exists\ K_\epsilon \subset\subset E$, s.t.

$$\begin{aligned} &\sup_{x \in K_0} |f_n(x) - f_0(x)| \\ \le\ & \epsilon + \sup_{y \in K_\epsilon} |f_n(y) - \alpha H_n f_n(y) - (f_0(y) - \alpha H_n f_0(y))|. \end{aligned}$$

- This a priori estimate comes from exponential tightness.
Why?

Exponential Tightness

Main ideas: stochastic generalization of Ascoli-Arzela:

- Modulus of continuity estimate follows from $H_n \to H$;

Exponential Tightness

Main ideas: stochastic generalization of Ascoli-Arzela:

- Modulus of continuity estimate follows from $H_n \to H$;
- Exponential compact containment (construct Lyapunov funct.);

Exponential Tightness

Main ideas: stochastic generalization of Ascoli-Arzela:

- Modulus of continuity estimate follows from $H_n \to H$;
- Exponential compact containment (construct Lyapunov funct.);
- Usually all follow from $\sup_n \sup_x H_n f_n < \infty$ for "wise" choices of f_ns.

First version (there are 2nd, 3rd, ...) of the LDP theorem

$\{X_n(\cdot) : n = 1, 2, \ldots\}$ are E-valued processes that solves martingale problem with generator A_n.

- Compute $H_n f = \frac{1}{n} e^{-nf} A_n e^{nf}$;

First version (there are 2nd, 3rd, ...) of the LDP theorem

$\{X_n(\cdot) : n = 1, 2, \ldots\}$ are E-valued processes that solves martingale problem with generator A_n.

- Compute $H_n f = \frac{1}{n} e^{-nf} A_n e^{nf}$;
- Verify $H \subset \lim_n H_n$ (in some appropriate sense);

First version (there are 2nd, 3rd, ...) of the LDP theorem

$\{X_n(\cdot) : n = 1, 2, \ldots\}$ are E-valued processes that solves martingale problem with generator A_n.

- Compute $H_n f = \frac{1}{n} e^{-nf} A_n e^{nf}$;
- Verify $H \subset \lim_n H_n$ (in some appropriate sense);
- Verify exponential tightness (usually follows from previous step);

First version (there are 2nd, 3rd, ...) of the LDP theorem

$\{X_n(\cdot) : n = 1, 2, \ldots\}$ are E-valued processes that solves martingale problem with generator A_n.

- Compute $H_n f = \frac{1}{n} e^{-nf} A_n e^{nf}$;
- Verify $H \subset \lim_n H_n$ (in some appropriate sense);
- Verify exponential tightness (usually follows from previous step);
- Verify comparison principle: i.e.

$$(I - \alpha H)\overline{f} \leq h, \text{ and} (I - \alpha H)\underline{f} \geq h \text{ implies } \overline{f} \leq \underline{f}.$$

First version (there are 2nd, 3rd, ...) of the LDP theorem

$\{X_n(\cdot) : n = 1, 2, \ldots\}$ are E-valued processes that solves martingale problem with generator A_n.

- Compute $H_n f = \frac{1}{n} e^{-nf} A_n e^{nf}$;
- Verify $H \subset \lim_n H_n$ (in some appropriate sense);
- Verify exponential tightness (usually follows from previous step);
- Verify comparison principle: i.e.

  $(I - \alpha H)\overline{f} \leq h,$ and$(I - \alpha H)\underline{f} \geq h$ implies $\overline{f} \leq \underline{f}$.

- LDP holds for $\{X_n(\cdot) : n = 1, 2, \ldots\}$.

The rate function

- First,

$$I(x(\cdot)) = \sup_{t_0, t_1, \ldots, t_k \in \Delta_x^c} \{I_0(x(0)) + \sum_{i=1}^{k} I_{t_i - t_{i-1}}(x(t_i)|x(t_{i-1}))\}$$

where Δ_x^c is continuity time points of path $x(\cdot)$ and

$$I_t(y|x) = \sup_{f \in D}\{f(y) - V(t)f(x)\}.$$

The rate function

- First,

$$I(x(\cdot)) = \sup_{t_0,t_1,\ldots,t_k \in \Delta_x^c} \{I_0(x(0)) + \sum_{i=1}^{k} I_{t_i - t_{i-1}}(x(t_i)|x(t_{i-1}))\}$$

  where Δ_x^c is continuity time points of path $x(\cdot)$ and

$$I_t(y|x) = \sup_{f \in D} \{f(y) - V(t)f(x)\}.$$

- Second, $\{V(t) : t \geq 0\}$ is a variational semigroup (the Nisio semigroup), so I should have a variational action integral representation.

The rate function and Nisio semigroup

Recall the infinitesimal entropy-pressure dual (always)

$$H_n f(x) = \sup_{g \in D(H)} \{A_n^g f(x) - L_n g(x)\}.$$

Idea: embed function space $D(H)$ into a nicer one.

- $A_n = \frac{1}{2n}\Delta$,then H_n is associated with ($U = D(H_n)$)

$$dX_n = \nabla g(t, X_n(t))dt + \frac{1}{\sqrt{n}} dW_n,$$

$$L(X_n(t), g(t, X_n(t))) = \frac{1}{2}|\nabla g(t, X_n(t))|^2.$$

The rate function and Nisio semigroup

Recall the infinitesimal entropy-pressure dual (always)

$$H_n f(x) = \sup_{g \in D(H)} \{A_n^g f(x) - L_n g(x)\}.$$

Idea: embed function space $D(H)$ into a nicer one.

- $A_n = \frac{1}{2n}\Delta$,then H_n is associated with ($U = D(H_n)$)

$$dX_n = \nabla g(t, X_n(t))dt + \frac{1}{\sqrt{n}} dW_n,$$

$$L(X_n(t), g(t, X_n(t))) = \frac{1}{2}|\nabla g(t, X_n(t))|^2.$$

- We could take $U = R^d$, $dX_n = udt + \frac{1}{\sqrt{n}} dW_n$ and $L(x, u) = \frac{1}{2}|u|^2$.

The rate function and Nisio semigroup - II

- Assume $\exists$ metric space U, $Af : E \times U \mapsto R$ and $L : E \times U \mapsto R \cup \{+\infty\}$ s.t.

$$Hf(x) = \sup_{u \in U} \{Af(x,u) - L(x,u)\}$$

The rate function and Nisio semigroup - II

- Assume $\exists$ metric space U, $Af : E \times U \mapsto R$ and $L : E \times U \mapsto R \cup \{+\infty\}$ s.t.

$$Hf(x) = \sup_{u \in U}\{Af(x,u) - L(x,u)\}$$

- Then

$$V(t)f(x) = \sup_{x(\cdot),\mu(\cdot)\in\Gamma_x}\{f(x(t)) - \int_{[0,t]\times U} L(x,u)\mu(du|s)ds\}$$

The rate function and Nisio semigroup - II

- Assume $\exists$ metric space U, $Af : E \times U \mapsto R$ and $L : E \times U \mapsto R \cup \{+\infty\}$ s.t.

$$Hf(x) = \sup_{u \in U} \{Af(x,u) - L(x,u)\}$$

- Then

$$V(t)f(x) = \sup_{x(\cdot),\mu(\cdot) \in \Gamma_x} \{f(x(t)) - \int_{[0,t]\times U} L(x,u)\mu(du|s)ds\}$$

- Γ_x means solving a controlled functional eqn

$$f(x(t)) - f(x(s)) = \int_s^t \int_U Af(x(r),u)\mu(du|r)dr, \quad \forall f \in D($$

The rate function and Nisio semigroup - III

- $$\begin{aligned} I(x(\cdot)) &= \inf\{\int_0^T \int_U L(x(s),u)\mu(du|s)ds : \\ &\quad (x(\cdot),\mu(\cdot)) \in \Gamma_{x(0)}\}. \end{aligned}$$

The rate function and Nisio semigroup - III

- $$\begin{aligned} I(x(\cdot)) &= \inf\{\int_0^T \int_U L(x(s), u)\mu(du|s)ds : \\ &\qquad (x(\cdot), \mu(\cdot)) \in \Gamma_{x(0)}\}. \end{aligned}$$

- To rigororize above: comparison principle for $(I - \alpha H)f = h$.

What is the catch

- Classical LDP approach – analysis in path space (tangent space analysis)

What is the catch

- Classical LDP approach – analysis in path space (tangent space analysis)
- This approach : More emphasize on analysis on function space (co-tangent space analysis).

What is the catch

- Classical LDP approach – analysis in path space (tangent space analysis)
- This approach : More emphasize on analysis on function space (co-tangent space analysis).
- variations on path v.s. variations on test functions

What is the catch

- Classical LDP approach – analysis in path space (tangent space analysis)
- This approach : More emphasize on analysis on function space (co-tangent space analysis).
- variations on path v.s. variations on test functions
- Technical subtlety: replace approximation for paths by proof of comparison principle for PDEs.