Weak convergence and large deviation theory

- Large deviation principle
- Convergence in distribution
- The Bryc-Varadhan theorem
- Tightness and Prohorov’s theorem
- Exponential tightness
- Tightness for processes
- (Exponential) tightness and results for finite dimensional distributions
- Conditions for (exponential) tightness

Joint work with Jin Feng

Second Lecture
Large deviation principle

\((S, d)\) a (complete, separable) metric space.

\(X_n, n = 1, 2, \ldots\) \(S\)-valued random variables

\(\{X_n\}\) satisfies a \textit{large deviation principle} (LDP) if there exists a lower semicontinuous function \(I : S \to [0, \infty]\) such that for each open set \(A\),

\[
\liminf_{n \to \infty} \frac{1}{n} \log P\{X_n \in A\} \geq -\inf_{x \in A} I(x)
\]

and for each closed set \(B\),

\[
\limsup_{n \to \infty} \frac{1}{n} \log P\{X_n \in B\} \leq -\inf_{x \in B} I(x).
\]
The rate function

$I$ is called the rate function for the large deviation principle.

A rate function is good if for each $a \in [0, \infty)$, $\{x : I(x) \leq a\}$ is compact.

If $I$ is a rate function for $\{X_n\}$, then

$$I_*(x) = \lim_{\epsilon \to 0} \inf_{y \in B_\epsilon(x)} I(y)$$

is also a rate function for $\{X_n\}$. $I_*$ is lower semicontinuous.

If the large deviation principle holds with lower semicontinuous rate $I$ function, then

$$I(x) = \lim_{\epsilon \to 0} \lim_{n \to \infty} \frac{1}{n} \log P\{X_n \in B_\epsilon(x)\} = \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log P\{X_n \in \overline{B}_\epsilon(x)\}$$
Convergence in distribution

\( \{X_n\} \) converges in distribution to \( X \) if and only if for each \( f \in C_b(S) \)

\[
\lim_{n \to \infty} E[f(X_n)] = E[f(X)]
\]
Equivalent statements: Large deviation principle

$\{X_n\}$ satisfies an LDP with a good rate function if and only if $\{X_n\}$ is exponentially tight and

$$
\Lambda(f) \equiv \lim_{n \to \infty} \frac{1}{n} \log E[e^{nf(X_n)}]
$$

for each $f \in C_b(S)$. Then

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

and

$$
\Lambda(f) = \sup_{x \in S} \{f(x) - I(x)\}
$$

Bryc, Varadhan
Equivalent statements: Convergence in distribution

\{X_n\} converges in distribution to \(X\) if and only if

\[
\liminf_{n \to \infty} P\{X_n \in A\} \geq P\{X \in A\}, \text{ each open } A,
\]

or equivalently

\[
\limsup_{n \to \infty} P\{X_n \in B\} \leq P\{X \in B\}, \text{ each closed } B
\]

LDP
Tightness

A sequence \( \{X_n\} \) is tight if for each \( \epsilon > 0 \), there exists a compact set \( K_\epsilon \subset S \) such that

\[
\sup_n P\{X_n \notin K_\epsilon\} \leq \epsilon.
\]

Prohorov’s theorem

**Theorem 1** Suppose that \( \{X_n\} \) is tight. Then there exists a subsequence \( \{n(k)\} \) along which the sequence converges in distribution.
Exponential tightness

\( \{X_n\} \) is exponentially tight if for each \( a > 0 \), there exists a compact set \( K_a \subset S \) such that

\[
\limsup_{n \to \infty} \frac{1}{n} \log P\{X_n \notin K_a\} \leq -a.
\]

Analog of Prohorov’s theorem

**Theorem 2** (Puhalskii, O’Brien and Vervaat, de Acosta) Suppose that \( \{X_n\} \) is exponentially tight. Then there exists a subsequence \( \{n(k)\} \) along which the large deviation principle holds with a good rate function.
**Stochastic processes in** $D_E[0, \infty)$

$(E, r)$ complete, separable metric space

$S = D_E[0, \infty)$

Modulus of continuity:

$$w'(x, \delta, T) = \inf \{ t_i \} \max_{i} \sup_{s, t \in [t_i-1, t_i]} r(x(s), x(t))$$

where the infimum is over $\{t_i\}$ satisfying

$$0 = t_0 < t_1 < \cdots < t_{m-1} < T \leq t_m$$

and $\min_{1 \leq i \leq n}(t_i - t_{i-1}) > \delta$

$X_n$ stochastic process with sample paths in $D_E[0, \infty)$

$X_n$ adapted to $\{\mathcal{F}_t^n\}$: For each $t \geq 0$, $X_n(t)$ is $\mathcal{F}_t^n$-measurable.
Tightness in $D_E[0, \infty)$

**Theorem 3 (Skorohod)** Suppose that for $t \in \mathcal{T}_0$, a dense subset of $[0, \infty)$, $\{X_n(t)\}$ is tight. Then $\{X_n\}$ is tight if and only if for each $\epsilon > 0$ and $T > 0$

$$\lim_{\delta \to 0} \limsup_{n \to \infty} P\{w'(X_n, \delta, T) > \epsilon\} = 0.$$

**Theorem 4 (Puhalskii)** Suppose that for $t \in \mathcal{T}_0$, a dense subset of $[0, \infty)$, $\{X_n(t)\}$ is exponentially tight. Then $\{X_n\}$ is exponentially tight if and only if for each $\epsilon > 0$ and $T > 0$

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \frac{1}{n} \log P\{w'(X_n, \delta, T) > \epsilon\} = -\infty.$$
Identification of limit distribution

Theorem 5 If \( \{X_n\} \) is tight in \( D_E[0, \infty) \) and
\[
(X_n(t_1), \ldots, X_n(t_k)) \Rightarrow (X(t_1), \ldots, X(t_k))
\]
for \( t_1, \ldots, t_k \in \mathcal{T}_0, \mathcal{T}_0 \) dense in \( [0, \infty) \), then \( X_n \Rightarrow X \).

Identification of rate function

Theorem 6 If \( \{X_n\} \) is exponentially tight in \( D_E[0, \infty) \) and for each \( 0 \leq t_1 < \cdots < t_m \), \( \{X_n(t_1), \ldots, X_n(t_m)\} \) satisfies the large deviation principle in \( E^m \) with rate function \( I_{t_1, \ldots, t_m} \), then \( \{X_n\} \) satisfies the large deviation principle in \( D_E[0, \infty) \) with good rate function
\[
I(x) = \sup_{\{t_i\} \subset \Delta_x} I_{t_1, \ldots, t_m}(x(t_1), \ldots, x(t_m)),
\]
where \( \Delta_x \) is the set of discontinuities of \( x \).
Conditions for tightness

\( S^n_0(T) \) collection of discrete \( \{\mathcal{F}_t^n\} \)-stopping times

\[ q(x, y) = 1 \wedge r(x, y) \]

Suppose that for \( t \in \mathcal{T}_0 \), a dense subset of \([0, \infty)\), \( \{X_n(t)\} \) is tight. Then the following are equivalent.

a) \( \{X_n\} \) is tight in \( D_E[0, \infty) \).
Conditions for tightness

b) For $T > 0$, there exist $\beta > 0$ and random variables $\gamma_n(\delta, T), \delta > 0,$ satisfying

$$
E[q^\beta(X_n(t + u), X_n(t)) \land q^\beta(X_n(t), X_n(t - v))|\mathcal{F}_t^n]
\leq E[\gamma_n(\delta, T)|\mathcal{F}_t^n]
$$

(1)

for $0 \leq t \leq T$, $0 \leq u \leq \delta$, and $0 \leq v \leq t \land \delta$ such that

$$
\lim_{\delta \to 0} \limsup_{n \to \infty} E[\gamma_n(\delta, T)] = 0
$$

and

$$
\lim_{\delta \to 0} \limsup_{n \to \infty} E[q^\beta(X_n(\delta), X_n(0))] = 0.
$$

(2)
Conditions for tightness

c) Condition (2) holds, and for each $T > 0$, there exists $\beta > 0$ such that

$$C_n(\delta, T) \equiv \sup_{\tau \in S^n_0(T)} \sup_{u \leq \delta} \sup_{v \leq \delta \wedge \tau} E\left[ \sup_{v \leq \delta \wedge \tau} q^\beta(X_n(\tau + u), X_n(\tau)) \wedge q^\beta(X_n(\tau), X_n(\tau - v)) \right]$$

satisfies $\lim_{\delta \to 0} \limsup_{n \to \infty} C_n(\delta, T) = 0$.

Aldous
Conditions for exponential tightness

$S_0^n(T)$ collection of discrete $\{F^n_t\}$-stopping times

$q(x, y) = 1 \wedge r(x, y)$

Suppose that for $t \in T_0$, a dense subset of $[0, \infty)$, $\{X_n(t)\}$ is exponentially tight. Then the following are equivalent.

a) $\{X_n\}$ is exponentially tight in $D_E[0, \infty)$. 
Conditions for exponential tightness

b) For $T > 0$, there exist $\beta > 0$ and random variables $\gamma_n(\delta, \lambda, T)$, $\delta, \lambda > 0$, satisfying

$$E[e^{n\lambda q^\beta(X_n(t+u),X_n(t))\wedge q^\beta(X_n(t),X_n(t-v))}|\mathcal{F}_t^n] \leq E[e^{\gamma_n(\delta, \lambda, T)}|\mathcal{F}_t^n]$$

for $0 \leq t \leq T$, $0 \leq u \leq \delta$, and $0 \leq v \leq t \wedge \delta$ such that

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \frac{1}{n} \log E[e^{\gamma_n(\delta, \lambda, T)}] = 0,$$

and

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \frac{1}{n} \log E[e^{n\lambda q^\beta(X_n(\delta),X_n(0))}] = 0. \quad (3)$$
Conditions for exponential tightness

c) Condition (3) holds, and for each $T > 0$, there exists $\beta > 0$ such that for each $\lambda > 0$

$$C_n(\delta, \lambda, T) \equiv \sup_{\tau \in S_0^n(T)} \sup_{u \leq \delta} \sup_{v \leq \delta \land \tau} E \left[ e^{n\lambda \beta (X_n(\tau + u), X_n(\tau)) \land q^\beta (X_n(\tau), X_n(\tau - v))} \right]$$

satisfies $\lim_{\delta \to 0} \limsup_{n \to \infty} \frac{1}{n} \log C_n(\delta, \lambda, T) = 0$. 
Example

$W$ standard Brownian motion

$X_n = \frac{1}{\sqrt{n}} W$

$$E[e^{n\lambda |X_n(t+u)-X_n(t)|} | \mathcal{F}_t^W] = E[e^{\lambda \sqrt{n} |W(t+u)-W(t)|} | \mathcal{F}_t^W] \leq 2e^{\frac{1}{2} n\lambda^2 u}$$

so

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \frac{1}{n} \log E[e^{\gamma_n(\delta,\lambda,T)}] = \lim_{\delta \to 0} \frac{1}{2} \lambda^2 \delta = 0.$$
Equivalence to tightness for functions

Theorem 7 \{X_n\} is tight in \(D_E[0, \infty)\) if and only if

\begin{enumerate}
\item (Compact containment condition) For each \(T > 0\) and \(\epsilon > 0\), there exists a compact \(K_{\epsilon,T} \subset E\) such that
\[
\limsup_{n \to \infty} P(\exists t \leq T \ni X_n(t) \notin K_{\epsilon,T}) \leq \epsilon
\]

\item There exists a family of functions \(F \subset C(E)\) that is closed under addition and separates points in \(E\) such that for each \(f \in F\), \(\{f(X_n)\}\) is tight in \(D_R[0, \infty)\).
\end{enumerate}

Kurtz, Jakubowski
Equivalence to exponential tightness for functions

Theorem 8 \( \{X_n\} \) is exponentially tight in \( D_E[0, \infty) \) if and only if

a) For each \( T > 0 \) and \( a > 0 \), there exists a compact \( K_{a,T} \subset E \) such that
\[
\limsup_{n \to \infty} \frac{1}{n} \log P(\exists t \leq T \ni X_n(t) \notin K_{a,T}) \leq -a
\]  

b) There exists a family of functions \( F \subset C(E) \) that is closed under addition and separates points in \( E \) such that for each \( f \in F \), \( \{f(X_n)\} \) is exponentially tight in \( D_R[0, \infty) \).

Schied
Large Deviations for Markov Processes

- Martingale problems and semigroups
- Semigroup convergence and the LDP
- Control representation of the rate function
- Viscosity solutions and semigroup convergence
- Summary of method
Markov processes

\( X_n = \{X_n(t), t \geq 0\} \) is a Markov process if

\[
E[g(X_n(t + s))|\mathcal{F}_t^n] = E[g(X_n(t + s))|X_n(t)]
\]

The generator of a Markov process determines its short time behavior

\[
E[g(X_n(t + \Delta t)) - g(X_n(t))|\mathcal{F}_t] \approx A_n g(X_n(t)) \Delta t
\]
Martingale problems

$X_n$ is a solution of the martingale problem for $A_n$ if and only if

$$g(X_n(t)) - g(X_n(0)) - \int_0^t A_ng(X_n(s))ds$$

(5)

is an $\{\mathcal{F}_t^n\}$-martingale for each $g \in \mathcal{D}(A_n)$.

If $g$ is bounded away from zero, (5) is a martingale if and only if

$$g(X_n(t)) \exp\left\{-\int_0^t \frac{A_ng(X_n(s))}{g(X_n(s))} ds\right\}$$

is a martingale. (You can always add a constant to $g$.)
Nonlinear generator

Define $\mathcal{D}(H_n) = \{ f \in B(E) : e^{nf} \in \mathcal{D}(A_n) \}$ and set

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

Then

$$\exp\{nf(X_n(t)) - nf(X(0)) - \int_0^t nH_n f(X(s)) ds\}$$

is a $\{\mathcal{F}_t^n\}$-martingale.
Tightness for solutions of MGPs

\[ E[f(X_n(t + u)) - f(X_n(t))|F^n_t] \]
\[ = E[\int_t^{t+u} A_n f(X_n(s))ds|F^n_t] \leq u\|A_n f\| \]

For \( \gamma_n(\delta, T) = \delta(\|A_n f^2\| + 2\|f\||A_n f\|) \) (see (1))

\[ E[(f(X_n(t + u)) - f(X_n(t)))^2|F^n_t] \]
\[ = E[\int_t^{t+u} A_n f^2(X_n(s))ds|F^n_t] \]
\[ - 2f(X_n(t))E[\int_t^{t+u} A_n f(X_n(s))ds|F^n_t] \]
\[ \leq u(\|A_n f^2\| + 2\|f\||A_n f\|) \leq \gamma_n(\delta, T) \]
Exponential tightness

\[ E\left[ e^{n(\lambda f(X_n(t+u)) - \lambda f(X_n(t)) - \int_t^{t+u} H_n[\lambda f](X_n(s))ds} \right| \mathcal{F}_t^n \right] = 1 \]

so

\[ E\left[ e^{n\lambda (f(X_n(t+u)) - f(X_n(t))) \right| \mathcal{F}_t^n \right] \leq e^{nu\|H_n\lambda f\|} \]

and

\[ \gamma_n(\delta, \lambda, T) = \delta n(\|H_n[\lambda f]\| + \|H_n[-\lambda f]\|) \]

ET Conditions
The Markov process semigroup

Assume that the martingale problem for $A_n$ is well-posed.

Define

$$T_n(t)f(x) = E[f(X_n(t))|X_n(0) = x]$$

By the Markov property

$$T_n(s)T_n(t)f(x) = T_n(t + s)f(x)$$

$$\lim_{t \to 0} \frac{T_n(t)f(x) - f(x)}{t} = A_n f(x)$$
Iterating the semigroup

For $0 \leq t_1 \leq t_2$,

$$E[f_1(X_n(t_1))f_2(X_n(t_2))|X_n(0) = x] = T_n(t_1)(f_1T_n(t_2 - t_1)f_2)(x)$$

and in general

$$E[f_1(X_n(t_1))\cdots f_k(X_n(t_k))|X_n(0) = x] = E[f_1(X_n(t_1))\cdots f_{k-1}(X_n(t_{k-1}))$$

$$T_n(t_k - t_{k-1})f_k(X_n(t_{k-1}))[X_n(0) = x]$$

Convergence of the semigroups implies convergence of the finite dimensional distributions.
A nonlinear semigroup (Fleming)

Assume that the martingale problem for $A_n$ is well-posed.

Define

$$V_n(t)f(x) = \frac{1}{n} \log E_x[e^{nf(X_n(t))}]$$

By the Markov property

$$V_n(s)V_n(t)f(x) = V_n(t+s)f(x)$$

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

Exponential generator
Iterating the semigroup

For $0 \leq t_1 \leq t_2$, define

$$V_n(t_1, t_2, f_1, f_2)(x) = V_n(t_1)(f_1 + V_n(t_2 - t_1)f_2)(x)$$

and inductively

$$V_n(t_1, \ldots, t_k, f_1, \ldots, f_k)(x)$$

$$= V_n(t_1)(f_1 + V_n(t_2, \ldots, t_k, f_2, \ldots, f_k))(x).$$

Then

$$E[e^n(f_1(X_n(t_1)) + \cdots + f_k(X_n(t_k)))]$$

$$= E[e^{nV_n(t_1, \ldots, t_k, f_1, \ldots, f_k)(X_n(0))}]$$

By the Bryc-Varadhan result, convergence of semigroup should imply the finite dimensional LDP
Weaker conditions for the LDP

A collection of functions $D \subset C_b(S)$ isolates points in $S$, if for each $x \in S$, each $\epsilon > 0$, and each compact $K \subset S$, there exists $f \in D$ satisfying $|f(x)| < \epsilon$, $\sup_{y \in K} f(y) \leq 0$, and

$$\sup_{y \in K \cap B_\epsilon(x)} f(y) < -\frac{1}{\epsilon}.$$ 

A collection of functions $D \subset C_b(S)$ is bounded above if $\sup_{f \in D} \sup_y f(y) < \infty$. 
A rate determining class

Proposition 9 Suppose $\{X_n\}$ is exponentially tight, and let

$$\Gamma = \{f \in C_b(S) : \Lambda(f) = \lim_{n \to \infty} \frac{1}{n} \log E[e^{nf(X_n)}] \text{ exists}\}.$$

If $D \subset \Gamma$ is bounded above and isolates points, then $\Gamma = C_b(S)$ and

$$I(x) = \sup_{f \in D} \{f(x) - \Lambda(f)\}.$$
Semigroup convergence and the LDP

Suppose $D \subset C_b(E)$ contains a set that is bounded above and isolates points.

Suppose $X_n(0) = x$ and $\{X_n(t)\}$ is exponentially tight. If $V_n(t)f(x) \to V(t)f(x)$ for each $f \in D$, then $\{X_n(t)\}$ satisfies a LDP with rate function

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

and hence

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

Think of $I_t(y|x)$ as the large deviation analog of a transition density.
Iterating the semigroup

Suppose $D$ is closed under addition, $V(t) : D \rightarrow D$, $t \geq 0$, and $0 \leq t_1 \leq t_2$. Define

$$V(t_1, t_2, f_1, f_2)(x) = V(t_1)(f_1 + V(t_2 - t_1)f_2)(x)$$

and inductively

$$V(t_1, \ldots, t_k, f_1, \ldots, f_k)(x) = V(t_1)(f_1 + V(t_2, \ldots, t_k, f_2, \ldots, f_k))(x)$$
Semigroup convergence and the LDP

**Theorem 10** For each $n$, let $A_n \subset C_b(E) \times B(E)$, and suppose that existence and uniqueness holds for the $D_E[0,\infty)$-martingale problem for $(A_n, \mu)$ for each initial distribution $\mu \in \mathcal{P}(E)$.

Let $D \subset C_b(E)$ be closed under addition and contain a set that is bounded above and isolates points, and suppose that there exists an operator semigroup $\{V(t)\}$ on $D$ such that for each compact $K \subset E$

$$\sup_{x \in K} |V(t)f(x) - V_n(t)f(x)| \to 0, \quad f \in D.$$
Suppose that \( \{X_n\} \) is exponentially tight, and that \( \{X_n(0)\} \) satisfies a large deviation principle with good rate function \( I_0 \). Define

\[
\Lambda_0(f) = \lim_{n \to \infty} \frac{1}{n} \log E[e^{nf(X_n(0))}], \quad f \in C_b(E).
\]

a) For each \( 0 \leq t_1 < \cdots < t_k \) and \( f_1, \ldots, f_k \in D \),

\[
\lim_{n \to \infty} \frac{1}{n} \log E[e^{nf_1(X_n(t_1))} + \cdots + nf_k(X_n(t_k))]
= \Lambda_0(V(t_1, \ldots, t_k, f_1, \ldots, f_k)).
\]

Recall

\[
E[e^{n(f_1(X_n(t_1)) + \cdots + f_k(X_n(t_k)))}]
= E[e^{nV_n(t_1, \ldots, t_k, f_1, \ldots, f_k)(X_n(0))}]
\]
b) For $0 \leq t_1 < \ldots < t_k \{(X_n(t_1), \ldots, X_n(t_k))\}$ satisfies the large
deviation principle with rate function

$$I_{t_1, \ldots, t_k}(x_1, \ldots, x_k) = \sup_{f_1, \ldots, f_k \in D \cap C_b(E)} \{f_1(x_1) + \ldots + f_k(x_k)$$

$$- \Lambda_0(V(t_1, \ldots, t_k, f_1, \ldots, f_k))\}$$

$$= \inf_{x_0 \in E} (I_0(x_0) + \sum_{i=1}^k I_{t_i-t_{i-1}}(x_i|x_{i-1}))$$

c) $\{X_n\}$ satisfies the large deviation principle in $D_E[0, \infty)$ with rate
function

$$I(x) = \sup_{\{t_i\} \subset \Delta^c_x} I_{t_1, \ldots, t_k}(x(t_1), \ldots, x(t_k))$$

$$= \sup_{\{t_i\} \subset \Delta^c_x} (I_0(x(0)) + \sum_{i=1}^k I_{t_i-t_{i-1}}(x(t_i)|x(t_{i-1})))$$
Example: Freidlin and Wentzell small diffusion

Let $X_n$ satisfying the Itô equation

$$X_n(t) = x + \frac{1}{\sqrt{n}} \int_0^t \sigma(X_n(s-))dW(s) + \int_0^t b(X_n(s))ds,$$

and define $a(x) = \sigma^T(x) \cdot \sigma(x)$. Then

$$A_n g(x) = \frac{1}{2n} \sum_{ij} a_{ij}(x) \partial_i \partial_j g(x) + \sum_i b_i(x) \partial_i g(x),$$

Take $\mathcal{D}(A_n)$ to be the collection of functions of the form $c + f$, $c \in \mathbb{R}$ and $f \in C^2_c(\mathbb{R}^d)$. 
Convergence of the nonlinear generator

\[ H_n f(x) = \frac{1}{2n} \sum_{ij} a_{ij}(x) \partial_{ij} f(x) + \frac{1}{2} \sum_{ij} a_{ij}(x) \partial_i f(x) \partial_j f(x) \]

\[ + \sum_i b_i(x) \partial_i f(x). \]

and \( H f = \lim_{n \to \infty} H_n f \) is

\[ H f(x) = \frac{1}{2} (\nabla f(x))^T \cdot a(x) \cdot \nabla f(x) + b(x) \cdot \nabla f(x). \]
A control problem

Let \((E, r)\) and \((U, q)\) be complete, separable metric spaces, and let \(A : \mathcal{D}(A) \subset C_b(E) \to C(E \times U)\)

Let \(H\) be as above, and suppose that there is a nonnegative, lower semicontinuous function \(L\) on \(E \times U\) such that

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

\(\{V(t)\}\) should be the Nisio semigroup corresponding to an optimal control problem with “reward” function \(-L\).

(cf. Book by Dupuis and Ellis)
Dynamics of control problem

Require

\[ f(x(t)) - f(x(0)) - \int_{U \times [0,t]} Af(x(s),u) \lambda_s(du \times ds) = 0, \]

for each \( f \in \mathcal{D}(A) \) and \( t \geq 0 \), where \( x \in D_E[0, \infty) \) and \( \lambda \in \mathcal{M}_m(U) \), the space of measures on \( U \times [0, \infty) \) satisfying \( \lambda(U \times [0, t]) = t \).

For each \( x_0 \in E \), we should have

\[
V(t)g(x_0) = \sup_{(x, \lambda) \in \mathcal{J}_{x_0}^t} \left\{ g(x(t)) - \int_{[0,t] \times U} L(x(s),u) \lambda(du \times ds) \right\}
\]
Representation theorem

**Theorem 11** Suppose \((E, r)\) and \((U, q)\) are complete, separable, metric spaces. Let \(A : \mathcal{D}(A) \subset C_b(E) \to C(E \times U)\) and lower semicontinuous \(L(x, u) \geq 0\) satisfy

1. \(\mathcal{D}(A)\) is convergence determining.
2. For each \(x_0 \in E\), there exists \((x, \lambda) \in \mathcal{J}\) such that \(x(0) = x_0\) and
   \[
   \int_{U \times [0, t]} L(x(s), u) \lambda(du \times ds) = 0, \quad t \geq 0.
   \]
3. For each \(f \in \mathcal{D}(A)\), there exists a nondecreasing function \(\psi_f : [0, \infty) \to [0, \infty)\) such that
   \[
   |Af(x, u)| \leq \psi_f(L(x, u)), \quad (x, u) \in E \times U,
   \]
   and \(\lim_{r \to \infty} r^{-1} \psi_f(r) = 0\).
4. There exists a tightness function $\Phi$ on $E \times U$, such that $\Phi(x, u) \leq L(x, u)$ for $(x, u) \in E \times U$.

Let $\{V(t)\}$ be an LDP limit semigroup and satisfy the control identity. Then

$$I(x) = I_0(x(0)) + \inf_{\lambda: (x, \lambda) \in \mathcal{J}} \left\{ \int_{U \times [0, \infty)} L(x(s), u) \lambda(du \times ds) \right\}.$$
Small diffusion

\[
H f(x) = \frac{1}{2} (\nabla f(x))^T \cdot a(x) \cdot \nabla f(x) + b(x) \cdot \nabla f(x)
\]

For

\[
A f(x, u) = u \cdot \nabla f(x)
\]

and

\[
L(x, u) = \frac{1}{2} (u - b(x) a(x)^{-1}(u - b(x)),
\]

\[
H f(x) = \sup_{u \in \mathbb{R}^d} (A f(x, u) - L(x, u))
\]

\[
I(x) = \int_0^\infty \frac{1}{2} (\dot{x}(s) - b(x(s)) a(x(s))^{-1}(\dot{x}(s) - b(x(s)))) ds
\]
Alternative representation

For

\[ Af(x, u) = (u^T \sigma(x) + b(x)) \nabla f(x) \]

and

\[ L(x, u) = \frac{1}{2} |u|^2, \]

again

\[ Hf(x) = \sup_{u \in \mathbb{R}^d} (Af(x, u) - L(x, u)) \]

\[ I(x) = \inf \left\{ \int_0^\infty \frac{1}{2} |u(s)|^2 ds : \dot{x}(t) = u^T(t)\sigma(x(t)) + b(x(t)) \right\} \]
Legendre transform approach

If $H_f(x) = H(x, \nabla f(x))$, where $H(x, p)$ is convex and continuous in $p$, then

$$L(x, u) = \sup_{p \in \mathbb{R}^d} \{ p \cdot u - H(x, p) \}$$

and

$$H(x, p) = \sup_{u \in \mathbb{R}^d} \{ p \cdot u - L(x, u) \},$$

so taking $A f(x, u) = u \cdot \nabla f(x)$,

$$H_f(x) = \sup_{u \in \mathbb{R}^d} \{ u \cdot \nabla f(x) - L(x, u) \}$$
Viscosity solutions

Let $E$ be compact, $H \subset C(E) \times B(E)$, and $(f, g) \in H$ imply $(f + c, g) \in H$. Fix $h \in C(E)$ and $\alpha > 0$.

$\bar{f} \in B(E)$ is a viscosity subsolution of

$$f - \alpha Hf = h$$

if and only if $\bar{f}$ is upper semicontinuous and for each $(f_0, g_0) \in H$ there exists $x_0 \in E$ satisfying $(\bar{f} - f_0)(x_0) = \sup_x (\bar{f}(x) - f_0(x))$ and

$$\frac{\bar{f}(x_0) - h(x_0)}{\alpha} \leq (g_0)^*(x_0)$$

or equivalently

$$\bar{f}(x_0) \leq \alpha (g_0)^*(x_0) + h(x_0)$$
\( f \in B(E) \) is a *viscosity supersolution* of (7) if and only if \( f \) is lower semicontinuous and for each \((f_0, g_0) \in H\) there exists \(x_0 \in E\) satisfying 
\[
(f_0 - f)(x_0) = \sup_x (f_0(x) - f(x)) 
\]
and
\[
\frac{f(x_0) - h(x_0)}{\alpha} \geq (g_0)_*(x_0) 
\]
or
\[
f(x_0) \geq \alpha(g_0)_*(x_0) + h(x_0) 
\]

A function \( f \in C(E) \) is a *viscosity solution* of \( f - \alpha H f = h \) if it is both a subsolution and a supersolution.
Comparison principle

The equation $f - \alpha H f = h$ satisfies a comparison principle, if $\bar{f}$ a viscosity subsolution and $\underline{f}$ a viscosity supersolution implies $\bar{f} \leq \underline{f}$ on $E$. 
Viscosity approach to semigroup convergence

Theorem 12  Let $(E, r)$ be a compact metric space, and for $n = 1, 2, \ldots$, assume that the martingale problem for $A_n \subset B(E) \times B(E)$ is well-posed.

Let

$$H_n f = \frac{1}{n} e^{-nf} A_n e^{nf}, \quad e^{nf} \in \mathcal{D}(A_n),$$

and let $H \subset C(E) \times B(E)$ with $\mathcal{D}(H)$ dense in $C(E)$. Suppose that for each $(f, g) \in H$, there exists $(f_n, g_n) \in H_n$ such that $\|f - f_n\| \to 0$ and $\|g - g_n\| \to 0$. 
Fix $\alpha_0 > 0$. Suppose that for each $0 < \alpha < \alpha_0$, there exists a dense subset $D_\alpha \subset C(E)$ such that for each $h \in D_\alpha$, the comparison principle holds for

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

Then there exists $\{V(t)\}$ on $C(E)$ such that

$$\sup_x |V(t)f(x) - V_n(t)f(x)| \to 0, \quad f \in C(E).$$

If $\{X_n(0)\}$ satisfies a large deviation principle with a good rate function. Then $\{X_n\}$ is exponentially tight and satisfies a large deviation principle with rate function $I$ given above (6).
Proof of a large deviation principle

1. Verify convergence of the sequence of operators $H_n$ and derive the limit operator $H$. In general, convergence will be in the extended limit or graph sense.

2. Verify exponential tightness. Given the convergence of $H_n$, exponential tightness typically follows provided one can verify the exponential compact containment condition.

3. Verify the range condition or the comparison principle for the limiting operator $H$. The rate function is characterized by the limiting semigroup.

4. Construct a variational representation for $H$. This representation typically gives a more explicit representation of the rate function.
$\mathbb{R}^d$-valued processes

Let $a = \sigma \sigma^T$, and define

$$A_n f(x) = n \int_{\mathbb{R}^d} \left( f(x + \frac{1}{n} z) - f(x) - \frac{1}{n} z \cdot \nabla f(x) \right) \eta(x, dz)$$

$$+ b(x) \cdot \nabla f(x) + \frac{1}{2n} \sum_{ij} a_{ij}(x) \partial_i \partial_j f(x)$$
Nonlinear generator

The operator $H_n f = \frac{1}{n} e^{-nf} A_n e^{nf}$ is given by

$$H_n f(x) = \int_{\mathbb{R}^d} \left( e^n(f(x + \frac{1}{n} z) - f(x)) - 1 - z \cdot \nabla f(x) \right) \eta(x, dz)$$

$$+ \frac{1}{2n} \sum_{ij} a_{ij}(x) \partial_i \partial_j f(x)$$

$$+ \frac{1}{2} \sum_{ij} a_{ij}(x) \partial_i f(x) \partial_j f(x) + b(x) \cdot \nabla f(x)$$
Limiting operator

\[ H f(x) = \int_{\mathbb{R}^d} (e^{\nabla f(x) \cdot z} - 1 - z \cdot \nabla f(x)) \eta(x, dz) \]

\[ + \frac{1}{2} \sum_{i,j} a_{ij}(x) \partial_i f(x) \partial_j f(x) + b(x) \cdot \nabla f(x) \]

Note that \( H \) has the form

\[ H f(x) = H(x, \nabla f(x)) \]

for

\[ H(x, p) = \frac{1}{2} |\sigma^T(x)p|^2 + b(x) \cdot p + \int_{\mathbb{R}^d} (e^{p \cdot z} - 1 - p \cdot z) \eta(x, dz) \]
Gradient limit operators

Condition 13

1. For each compact $\Gamma \subset \mathbb{R}^d$, there exist $\mu_m \to +\infty$ and $\omega : (0, \infty) \to [0, \infty]$ such that $\{(x_m, y_m)\} \subset \Gamma \times \Gamma$, $\mu_m|x_m - y_m|^2 \to 0$, and

$$\sup_m H_*(y_m, \mu_m(x_m - y_m)) < \infty$$

imply

$$\liminf_{m \to \infty} [\lambda H_*(x_m, \frac{\mu_m(x_m - y_m)}{\lambda}) - H_*(y_m, \mu_m(x_m - y_m))] \leq \omega(\lambda)$$

and

$$\lim_{\epsilon \to 0} \inf_{|\lambda - 1| \leq \epsilon} \omega(\lambda) \leq 0.$$ 

2. If $x_m \to \infty$ and $p_m \to 0$, then $\lim_{m \to \infty} H(x_m, p_m) = 0$. 
Conditions for comparison principle

Lemma 14 If Condition 13 is satisfied, then for \( h \in C(E) \) and \( \alpha > 0 \), the comparison principle holds for

\[
(I - \alpha H)f = h.
\]
Sufficient conditions

Lemma 15 Suppose $\sigma$ and $b$ are bounded and Lipschitz and $\eta = 0$. Then Condition 13 holds with

$$\omega(\lambda) = \begin{cases} 
0 & \lambda > 1 \\
\infty & \lambda \leq 1.
\end{cases}$$

If $H$ is continuous and for each $x, p \in \mathbb{R}^d \lim_{r \to \infty} H(x, rp) = \infty$, then Condition 13.1 holds with

$$\omega(\lambda) = \begin{cases} 
0 & \lambda = 1 \\
\infty & \lambda \neq 1.
\end{cases}$$

If $\sigma$ and $b$ are bounded and

$$\lim_{|p| \to 0} \sup_x \int_{\mathbb{R}^d} (e^{p \cdot z} - 1 - p \cdot z) \eta(x, dz) = 0,$$

then Condition 13.2 holds.
Diffusions with periodic coefficients (Baldi)

Let $\sigma$ be periodic (for each $1 \leq i \leq d$, there is a period $p_i > 0$ such that $\sigma(y) = \sigma(y + p_i e_i)$ for all $y \in \mathbb{R}^d$), and let $X_n$ satisfy the Itô equation

$$dX_n(t) = \frac{1}{\sqrt{n}}\sigma(\alpha_n X_n(t))dW(t),$$

where $\alpha_n > 0$ and $\lim_{n \to \infty} n^{-1} \alpha_n = \infty$. Let $a = \sigma \sigma^T$. Then

$$A_nf(x) = \frac{1}{n} \sum_{ij} a_{ij}(\alpha_n x) \frac{\partial^2}{\partial x_i \partial x_j} f(x),$$

and

$$H_nf(x) = \frac{1}{2n} \sum_{ij} a_{ij}(\alpha_n x) \partial_{ij} f(x) + \frac{1}{2} \sum_{ij} a_{ij}(\alpha_n x) \partial_i f(x) \partial_j f(x).$$
Limit operator

Let $f_n(x) = f(x) + \epsilon_n h(x, \alpha_n x)$, where $\epsilon_n = n\alpha_n^{-2}$.

$\epsilon_n \alpha_n = n\alpha_n^{-1} \to 0$

If $h$ has the same periods in $y$ as the $a_{ij}$ and

$$
\frac{1}{2} \sum_{ij} a_{ij}(y) \left( \frac{\partial^2}{\partial y_i \partial y_j} h(x, y) + \partial_i f(x) \partial_j f(x) \right) = g(x)
$$

for some $g$ independent of $y$, then

$$
\lim_{n \to \infty} H_n f_n(x, y) = g(x).
$$
It follows that
\[ g(x) = \frac{1}{2} \sum_{i,j} \bar{a}_{ij} \partial_i f(x) \partial_j f(x), \]
where \( \bar{a}_{ij} \) is the average of \( a_{ij} \) with respect to the stationary distribution for the diffusion on \([0, p_1] \times \cdots \times [0, p_d]\) whose generator is

\[ A_0 f(y) = \frac{1}{2} \sum_{i,j} a_{ij}(y) \frac{\partial^2}{\partial y_i \partial y_j} f(y) \]

with periodic boundary conditions. In particular,

\[ h(x, y) = \frac{1}{2} \sum_{i,j} h_{ij}(y) \partial_i f(x) \partial_j f(x), \]

where \( h_{ij} \) satisfies

\[ A_0 h_{ij}(y) = \bar{a}_{ij} - a_{ij}(y). \]