Large Deviations and Variational Representations

Freidlin-Wentzell Asymptotics in Infinite Dimensions

Amarjit Budhiraja

Department of Statistics & Operations Research

University of North Carolina at Chapel Hill

with P. Dupuis and V. Maroulas

Outline.

1. Background and basic definitions.
2. Classical approaches to small noise LDP.
3. Advantages of the current approach.
4. Main Result.
5. Some applications.
6. Key ingredient in the proof: Variational repn. for inf. dim. BM.
7. Proof sketch of the variational repn. .
8. Extensions.

Background and Definitions.

- Concerned with decay rate of probabilities of rare events.
- E.g. consider a k-dimensional SDE:

$$dX^\epsilon(t) = b\big(X^\epsilon(t)\big)dt + \sqrt{\epsilon}a\big(X^\epsilon(t)\big)dW(t), \; X^\varepsilon(0) = x_0, \; t \in [0,T],$$

As $\epsilon \to 0, \; X^\epsilon \xrightarrow{\mathbb{P}} X^0$ in $\mathcal{C}([0,T] : \mathbb{R}^k) \equiv \mathcal{C}$, where X^0 solves the ODE

$$\dot{X}^0 = b(X^0).$$

- Freidlin-Wentzell theory describes precise asymptotics of probabilities such as

$$\mathbb{P}\big\{\sup_{0\le t\le T} |X^\epsilon(t) - X^0(t)| > a\big\}$$

through a Large Deviation Principle.

Large Deviation Principle

Definition. Consider a sequence $\{X^\varepsilon\}_{\varepsilon>0}$ of $\mathcal{E}$ valued r.vs. $\mathcal{E}$ - Polish.

(1) A function I from $\mathcal{E}$ to $[0,\infty]$ is called a rate function on $\mathcal{E}$ if for each $M<\infty$ $\{x\in\mathcal{E} : I(x)\le M\}$ is compact.

(2) $\{X^\epsilon\}$ is said to satisfy the large deviation principle on $\mathcal{E}$ (as $\varepsilon\to 0$) with rate function I if:
(a) For each closed subset F of $\mathcal{E}$

$$\limsup_{\epsilon\to 0}\epsilon\log\mathbb{P}(X^\epsilon\in F)\le -\inf_{x\in F} I(x).$$

(b) For each open subset G of $\mathcal{E}$

$$\liminf_{\epsilon\to 0}\epsilon\log\mathbb{P}(X^\epsilon\in G)\ge -\inf_{x\in G} I(x).$$

Large Deviation Principle

Formally, for small ε:

$$\mathbb{P}(X^\epsilon \in A) \approx \exp\left\{-\frac{\inf_{x\in A} I(x)}{\varepsilon}\right\}, \; A \in \mathcal{B}(\mathcal{E}).$$

Small Noise SDE

- For the f.d. SDEs above, F-W show that the family, $\{X^\varepsilon\}$ of $\mathcal{C}$ valued random elements satisfies LDP with rate function I given as

$$I(f) = \inf_{u \in \mathcal{A}_f} \frac{1}{2} \int_0^T |u_s|^2 ds,$$

$$\mathcal{A}_f = \{u \in L^2([0,T] : \mathbb{R}^m) : \dot{f}_t = b(f_t) + a(f_t)u_t, a.e.\, t,\ f_0 = x_0\}.$$

  In particular:
  $\limsup_{\varepsilon \to 0} \varepsilon \log \mathbb{P}\{||X^\epsilon - X^0|| \geq a\} \leq -\inf_{x \in F_a} I(x)$, where

$$F_a = \{x \in \mathcal{C} : ||x - X^0|| \geq a\}.$$

- Used in study of exit time and invariant measure asymptotics.

Infinite Dimensional Noise

- Here we consider SDEs with an infinite dim. driving noise.
- Well Studied: Faris & Jonas-Lasino (1982), Freidlin(1988), Imauikin & Komech (1988), Zabczyk (1988), Sowers(1992), Chow(1992), Peszat(1994), Kallianpur & Xiong(1996), Cardon-Webber(1999), Cerrai & Roeckner(2004), Feng & Kurtz(2006)...

Standard Approaches

- Build on the ideas of Azencott(1980).
- One approximates by Gaussian systems by freezing the diffusion coefficient and suitable time discretizations.
- ...or one considers approximations by SDEs with f.d. noise.
- LDP for approximations follow from classical results.
- Finally one obtains suitable exponential continuity estimates in order to obtain the LDP for the original non-Gaussian infinite dim. system.
- Feng-Kurtz: proofs based on exponential tightness estimates and uniqueness theory for infinite dimensional HJ equations.
- Exponential continuity and tightness estimates are hard.
- ...sometimes obtained under “sub-optimal" conditions.

Advantages of Current Approach

- No approximations or discretizations.
- Exponential prob. estimates are completely bypassed.
- Proofs of LDP reduce to demonstrating basic qualitative properties of certain perturbations of the original system.

- For f.d. SDEs this amounts to showing:
  (i) For any $\theta \in [0,1)$, $x \in \mathbb{R}^k$ and any L^2–bounded control u the SDE below has a unique solution.

$$dX^{\theta,u}(t)=b\Big(X^{\theta,u}(t)\Big)dt+\theta a\Big(X^{\theta,u}(t)\Big)dW(t)+a\Big(X^{\theta,u}(t)\Big)u(t)dt,\; X^{\theta,u}(0)=x$$

  (ii) If $\theta_n \to 0$ and $u_n \Rightarrow u$, where $\{u_n\}$ are uniformly L^2–bounded controls, then $X^{\theta_n,u_n} \to X^{0,u}$ in distribution.

Brownian Sheet

- Let $\mathcal{O}$ be a bounded open set in $\mathbb{R}^d$ and $\{B(t,x),(t,x) \in [0,T] \times \mathcal{O}\}$ be a Brownian sheet.

  I.e. it is a mean zero, continuous, Gaussian random field such that
  - $Cov\big(B(t,x), B(s,y)\big) = \text{Leb}(A_{t,x} \cap A_{s,y})$, where

$$A_{t,x} \doteq \big\{(s,y) : s \in [0,t],\ y \in \mathcal{O} \cap [0,x]\big\}.$$

- B is a $(\mathbb{C}, \mathcal{B}(\mathbb{C}))$ valued r.v., where $\mathbb{C} = C([0,T] \times \overline{\mathcal{O}} : \mathbb{R})$ and $\mathcal{B}(\mathbb{C})$ the Borel sigma-field.

- Denote by μ the induced Wiener measure.

- Henceforth B is the canonical process on $(\mathbb{C}, \mathcal{B}(\mathbb{C}), \mu)$.

A General LDP

For $\varepsilon > 0$ and Polish space $\mathcal{E}$, let $\mathcal{G}^\varepsilon : \mathbb{C} \to \mathcal{E}$ be a measurable map.

Interested in LDP (as $\varepsilon \to 0$) for

$$X^\varepsilon \doteq \mathcal{G}^\varepsilon(\sqrt{\epsilon}B).$$

Typical example of X^ε: Solution of a small noise SPDE.

LDP for $X^\varepsilon \doteq \mathcal{G}^\varepsilon(\sqrt{\epsilon}B)$

- Some Notation: Let

$$S^N \doteq \left\{\phi \in H \equiv L^2([0,T]\times\mathcal{O}) : ||\phi||_H^2 \leq N\right\},$$

where

$$||\phi||_H^2 = \int_{[0,T]\times\mathcal{O}} \phi^2(s,x)dsdx.$$

- S^N is a compact Polish space with the weak topology.
- With $\{\mathcal{F}_t\}$ the canonical filtration on $(\mathbb{C},\mathcal{B}(\mathbb{C}))$, define

$$\mathcal{P}_2^N \doteq \left\{u : u \text{ is } \mathcal{P}\otimes\mathcal{B}(\mathcal{O}) \text{ measurable and } u(\omega) \in S^N, \mu - a.s.\right\}.$$

- For $\phi \in H$, define $\text{Int}(\phi) \in \mathbb{C}$ by

$$\text{Int}(\phi)(t,x) \doteq \int_{A_{t,x}} \phi(s,y)dsdy,$$

LDP for $X^\varepsilon \doteq \mathcal{G}^\varepsilon(\sqrt{\epsilon}B)$ (contd.)

Assumption. There exists a measurable map $\mathcal{G}^0 : \mathbb{C} \to \mathcal{E}$ such that: For every $N < \infty$:

- Whenever $\{u_n\} \subset \mathcal{P}_2^N$ is such that $u_n \Rightarrow u$ (as S^N–valued random elements), and $\varepsilon_n \in [0,1)$ is such that $\varepsilon_n \to 0$, we have

$$\mathcal{G}^{\epsilon_n}\left(\sqrt{\epsilon_n}B + \text{Int}(u_n)\right) \Rightarrow \mathcal{G}^0\left(\text{Int}(u)\right).$$

Theorem. Suppose Assumption holds. Then, the family $\{X^\epsilon\}$ satisfies LDP on $\mathcal{E}$, with rate function

$$I(f) \doteq \inf_{\{u \in H : f = \mathcal{G}^0(\text{Int}(u))\}} \left\{\frac{1}{2}||u||_H^2\right\}.$$

Extensions.

- A slight strengthening of Assumption 1 gives a uniform LDP.
- Analogous results can be obtained for an infinite sequence of real i.i.d. BMs, a cylindrical BM, and a Hilbert space valued BM.

An Application: A Toy Example.

Nonlinear stochastic cable equation: Let $L = \alpha I - \beta \frac{d^2}{dx^2}$.

$$dX^\varepsilon(t,r) = -LX^\varepsilon(t,r)dr\,dt + \sqrt{\epsilon}F(X^\varepsilon(t,r))B(drdt),\ \ x \in (0,b),\ t \in [0,T].$$

with initial and boundary condition $X^\varepsilon(0,r) = f(r)$, $r \in [0,b]$, $\frac{\partial}{\partial x}X^\varepsilon(t,0) = \frac{\partial}{\partial x}X^\varepsilon(t,b) = 0$, $t \in [0,T]$.

There is a unique continuous mild solution:

$$\begin{aligned} X^\varepsilon(t,r) &= \int_{[0,b]} G(t,r,q)f(q)dq \\ &+ \sqrt{\epsilon}\int_{[0,t]\times[0,b]} G(t-s,r,q)F(X^\varepsilon(s,q))B(dqds). \end{aligned}$$

Thus there is a measurable map $\mathcal{G}^\varepsilon : \mathbb{C} \to \mathbb{C}$ such that

$$X^\varepsilon = \mathcal{G}^\varepsilon(\sqrt{\varepsilon}B).$$

Verifying Assumption for $\mathcal{G}^\varepsilon$

Defining $\mathcal{G}^0$: For $\phi \in H$, let ξ_ϕ be the unique soln. of

$$\begin{aligned} \xi_\phi(t,r) &= \int_{[0,b]} G(t,r,q) f(q) dq \\ &+ \int_{[0,t]\times[0,b]} G(t-s,r,q) F(\xi_\phi(s,q))\phi(s,q) dq\, ds. \end{aligned}$$

Define $\mathcal{G}^0 : \mathbb{C} \to \mathbb{C}$ as

$$\mathcal{G}^0(v) = \xi_\phi, \quad \text{if } v = \text{Int}(\phi), \quad \text{for some } \phi \in H.$$

Set $\mathcal{G}^0(v) = 0$ otherwise.

Verifying Assumption (Contd.)

Let $\{u^\varepsilon\}$ be a sequence in $\mathcal{P}_2^N$ such that $u^\varepsilon \Rightarrow u$.

Let $X^{u^\varepsilon} = \mathcal{G}^\varepsilon(\sqrt{\varepsilon}B + \text{Int}(u^\varepsilon))$ and $X^0 = \mathcal{G}^0(\text{Int}(u))$.

Need to show: $X^{u^\varepsilon} \Rightarrow X^0$. Can Check X^{u^ε} solves:

$$
\begin{aligned}
X^{u^\varepsilon}(t,r) &= \int_{[0,b]} G(t,r,q) f(q) dq \\
&+ \sqrt{\epsilon} \int G(...) F(X^{u^\varepsilon}) B(dqds) + \int G(...) F(X^{u^\varepsilon}) u^\varepsilon(s,q) dqds.
\end{aligned}
$$

Also X^0 solves

$$
\begin{aligned}
X^0(t,r) &= \int_{[0,b]} G(t,r,q) f(q) dq \\
&+ \int G(...) F(X^0) u(s,q) dq\, ds.
\end{aligned}
$$

Weak convergence follows by standard estimates...

Applications

- Hilbert Space Valued Diffusions. Unique solvability studied in Leha and Ritter (1984). Small noise LDP in B.-Dupuis(2000).
- Stochastic reaction diffusion equations. Prior works on LDP: Freidlin(1988), Zabczyk(1988), Sowers(1992), Kallianpur and Xiong(1995). These papers assume diffusion coefficient is bounded, “cone condition" on domain... conditions needed for tail probability estimates on certain stochastic convolutions in Holder norms.

  Conditions relaxed in B.-Dupuis-Maroulas(2008).
- Stochastic flows of diffeomorphisms. B.-Dupuis-Maroulas(2009). Prior works include Millet, Nualart and Sanz-Sole(1992), Ben Arous and Castell(1995)–these concern finite dimensional flows.

Other Applications.

- Ren and Zhang (J. Func. Anal.(2004), Bull. Sci. Math.(2005)) - SDE driven with infinitely many BMs and non-Lipschitz diffusion coefficients...Brownian motion on the homeomorphism group of the circle...homeomorphism flows of non-Lipschitz multi-dimensional SDEs.
- Sritharan and Sundar (Stoc. Proc. App.(2006)) - 2D Navier-Stokes equation with multiplicative noise.
- Wang and Duan (Preprint(2007).) Stochastic parabolic PDEs with rapidly varying (random) boundary conditions.
- Liu (App. Math. Opt.(2009)) Stochastic evolution equations– general monotone drift and multiplicative noise.
- Bo and Jiang (preprint) Stochastic variational inequalities, reflected SPDEs.

Proof of the general LDP: $X^\varepsilon \doteq \mathcal{G}^\varepsilon(\sqrt{\epsilon}B)$

Suffices to show that Laplace principle holds: For all $h \in C_b(\mathcal{E})$

$$\lim_{\epsilon\to 0} -\epsilon \log\ \mathbb{E}\left\{\exp\left[-\frac{1}{\epsilon}h(X^\epsilon)\right]\right\} = \inf_{x\in\mathcal{E}}\{h(x)+I(x)\}.$$

A Variational Representation (B.-Dupuis(2000)): Let $f : \mathbb{C} \to \mathbb{R}$ be a bounded measurable map. Let B be a Brownian sheet. Then

$$-\log \mathbb{E}(\exp\{-f(B)\}) = \inf_{u\in\mathcal{P}_2} \mathbb{E}\left(\frac{1}{2}||u||_H^2 + f(B + \mathsf{Int}(u))\right).$$

Recall $X^\varepsilon \doteq \mathcal{G}^\varepsilon(\sqrt{\epsilon}B)$. Applying repn. with $f = \frac{1}{\epsilon} h o \mathcal{G}^\varepsilon(\sqrt{\varepsilon}\cdot)$ we have

$$-\epsilon \log \mathbb{E}\left\{\exp\left[-\frac{1}{\epsilon}h(X^\epsilon)\right]\right\} = \inf_u \mathbb{E}\left(\frac{1}{2}||u||_H^2 + h(X^{\epsilon,u})\right),$$

where $X^{\epsilon,u} = \mathcal{G}^\varepsilon(\sqrt{\epsilon}B + \mathsf{Int}(u))$

Proof: $\inf_u \mathbb{E}\Big(\frac{1}{2}||u||_H^2 + h(X^{\epsilon,u})\Big) \to \inf_{x\in\mathcal{E}}\{h(x)+I(x)\}.$

Proof of Upper Bound. Fix $\delta \in (0,1)$ and choose for each ε, u^ϵ such that

$$\text{LHS} \geq \mathbb{E}\left(\frac{1}{2}||u^\varepsilon||_H^2 + h(X^{\epsilon,u^\varepsilon})\right) - \delta.$$

- WLOG, for some $N < \infty$, $\sup_{\epsilon>0} ||u^\varepsilon||_H^2 \leq N$.

Pick a subsequence along which u^ϵ converges in distribution to some u. From Assumption:

$$\begin{aligned}
&\liminf_{\epsilon\to 0} \mathbb{E}\left[\frac{1}{2}||u^\varepsilon||_H^2 + h \circ \mathcal{G}^\epsilon\left(\sqrt{\epsilon}B + \text{Int}(u^\epsilon)\right)\right] \\
\geq\ & \mathbb{E}\left[\frac{1}{2}||u||_H^2 + h \circ \mathcal{G}^0\left(\text{Int}(u)\right)\right] \\
\geq\ & \inf_{\{(x,u)\in\mathcal{E}\times H: x=\mathcal{G}^0(\text{Int}(u))\}} \left\{\frac{1}{2}||u||_H^2 + h(x)\right\} \\
\geq\ & \inf_{x\in\mathcal{E}}\{I(x)+h(x)\}.
\end{aligned}$$

Proof: $\lim_{\epsilon\to 0}\inf_u \mathbb{E}\Big(\frac{1}{2}||u||_H^2+h(X^{\epsilon,u})\Big)=\inf_{x\in\mathcal{E}}\{h(x)+I(x)\}.$

Proof of the lower bound. Fix $\delta > 0$ and let $x_0 \in \mathcal{E}$ be such that $I(x_0) + h(x_0) \leq \inf_{x\in\mathcal{E}}\{I(x) + h(x)\} + \frac{\delta}{2}$.

Choose $\widetilde{u} \in H$ such that:

$$\frac{1}{2}||\tilde{u}||_H^2 \leq I(x_0) + \frac{\delta}{2} \quad \text{and} \quad x_0 = \mathcal{G}^0\left(\int_0^{\cdot} \tilde{u}(s)ds\right).$$

Then

LHS

$$\begin{aligned}
&= \limsup_{\epsilon\to 0}\inf_u \mathbb{E}\left[\frac{1}{2}||u||_H^2 + h\circ\mathcal{G}^\epsilon\left(\sqrt{\epsilon}B + \mathsf{Int}(u)\right)\right] \\
&\leq \limsup_{\epsilon\to 0}\mathbb{E}\left[\frac{1}{2}||\tilde{u}||_H^2 + h\circ\mathcal{G}^\epsilon\left(\sqrt{\epsilon}B + \mathsf{Int}(\tilde{u})\right)\right] \\
&= \frac{1}{2}||\tilde{u}||_H^2 + \mathbb{E}\left[h\circ\mathcal{G}^0\left(\mathsf{Int}(\tilde{u})\right)\right] = \frac{1}{2}||\tilde{u}||_H^2 + h(x_0) \\
&\leq I(x_0) + h(x_0) + \frac{\delta}{2} \leq \inf_{x\in\mathcal{E}}\{I(x) + h(x)\} + \delta.
\end{aligned}$$

Variational Repn.: Sketch of Proof.

$$-\log \mathbb{E}^{\mu}(\exp\{-f(B)\}) = \inf_{u\in\mathcal{P}_2} \mathbb{E}^{\mu}\left(\tfrac{1}{2}||u||_H^2 + f(B + \mathsf{Int}(u))\right). \tag{1}$$

Donsker-Varadhan variational formula:

$$-\log \mathbb{E}^{\mu}(\exp\{-f(B)\}) = \inf_{\gamma\in\mathcal{P}(\mathbb{C})} (R(\gamma \parallel \mu) + \mathbb{E}^{\gamma}(f(B))).$$

Upper Bound (In (1), LHS $\leq$ RHS): For $u \in \mathcal{P}_{\mathsf{SIM}}$, let $\gamma^u \in \mathcal{P}(\mathbb{C})$ be defined as

$$d\gamma^u = \exp\left(\int_{[0,T]\times\mathcal{O}} u(s,x)W(dsdx) - \frac{1}{2}||u||_H^2\right) d\mu.$$

Then $R(\gamma^u \parallel \mu) = \mathbb{E}^{\gamma^u}(\frac{1}{2}||u||_H^2)$.

Upper Bound Proof (ctd.)

$$
\begin{aligned}
-\log \mathbb{E}^{\mu}(\exp\{-f(B)\}) &\leq \inf_{u\in\mathcal{P}_{\mathsf{SIM}}} \mathbb{E}^{\gamma^u}\left(f(B) + \tfrac{1}{2}||u||_H^2\right) \\
&= \inf_{u\in\mathcal{P}_{\mathsf{SIM}}} \mathbb{E}^{\gamma^u}\left(f(\tilde{B}^u + \mathsf{Int}(u)) + \tfrac{1}{2}||u||_H^2\right) \\
&= \inf_{u\in\mathcal{P}_{\mathsf{SIM}}} \mathbb{E}^{\mu}\left(f(B + \mathsf{Int}(u)) + \tfrac{1}{2}||u||_H^2\right) \\
&= \inf_{u\in\mathcal{P}_2} \mathbb{E}^{\mu}\left(f(B + \mathsf{Int}(u)) + \tfrac{1}{2}||u||_H^2\right).
\end{aligned}
$$

Lower Bound Proof.

Infimum in the D-V formula attained at γ^0, with $d\gamma^0 = c\exp(-f(B))d\mu$. I.e.

$$-\log \mathbb{E}^{\mu}(\exp\{-f(B)\}) = \big(R(\gamma^0 \parallel \mu) + \mathbb{E}^{\gamma^0}(f(B))\big).$$

Define a martingale $L(t) = E(\frac{d\gamma_0}{d\mu} \mid \mathcal{F}_t)$. Martingale repn. theorem gives, for some $v \in \mathcal{P}_2$

$$\begin{aligned} L(t) &= 1 + \int_{[0,t]\times\mathcal{O}} v(s,x)dW(s,x) \\ &= 1 + \int_{[0,t]\times\mathcal{O}} L(s)u(s,x)dW(s,x). \end{aligned}$$

So $L(t) = \exp\{\int_{[0,t]\times\mathcal{O}} u(s,x)dW(s,x) - \frac{1}{2}\int_{[0,t]\times\mathcal{O}} |u(s,x)|^2 dsdx.\}$

Thus $\gamma^0 = \gamma^u$. Also: $\widetilde{B}^u = B - \text{Int}(u)$ is a Brownian sheet under γ^0.

Lower Bound Proof (ctd.)

$$
\begin{aligned}
-\log \mathbb{E}^{\mu}(\exp\{-f(B)\}) &= \mathbb{E}^{\gamma^u}\left(f(\widetilde{B}^u + \mathsf{Int}(u)) + \tfrac{1}{2}||u||_H^2\right) \\
&\geq \inf_{u \in \mathcal{P}_2} \mathbb{E}^{\gamma^u}\left(f(\tilde{B}^u + \mathsf{Int}(u)) + \tfrac{1}{2}||u||_H^2\right) \\
&\geq \inf_{u \in \mathcal{P}_{\mathsf{SIM}}} \mathbb{E}^{\gamma^u}\left(f(\tilde{B}^u + \mathsf{Int}(u)) + \tfrac{1}{2}||u||_H^2\right) \\
&= \inf_{u \in \mathcal{P}_{\mathsf{SIM}}} \mathbb{E}^{\mu}\left(f(B + \mathsf{Int}(u)) + \tfrac{1}{2}||u||_H^2\right) \\
&\geq \inf_{u \in \mathcal{P}_2} \mathbb{E}^{\mu}\left(f(B + \mathsf{Int}(u)) + \tfrac{1}{2}||u||_H^2\right).
\end{aligned}
$$

Extensions

- Representation for positive functionals of Poisson Random measures. Zhang(2009) gives a representation - not suitable for large deviation applications.
- more generally: Repn. for functionals of (inf. dim BM, PRM).
- large deviation applications - infinite dimensional jump-diffusions.
- Asymptotics of a large number of interacting diffusions.