9. Applications

9.1. **Large deviations for finite state Markov chains.**

In this section, we study the large deviations of Markov chains. Let Σ be a finite alphabet and $\mathbf{\Pi} = \{\pi(i,j)\}_{i,j=1}^{|\Sigma|}$ be a stochastic matrix, i.e., a nonnegative matrix with sum of each row 1. Consider a Markov chain $\{Y_k\}_{k\geq 0}$ with the transition probability $\mathbf{\Pi}$, for which we denote by $\mathbb{P}_\sigma^\pi$ the Markov probability measure with initial state $\sigma \in \Sigma$, i.e., the probability measure satisfying

$$\mathbb{P}_\sigma^\pi(Y_1 = y_1, \cdots, Y_n = y_n) = \pi(\sigma, y_1) \prod_{i=1}^{n-1} \pi(y_i, y_{i+1}).$$

In the following, we consider the empirical mean

$$Z_n = \sum_{k=1}^{n} X_k,$$

where $X_k = f(Y_k)$ with $f : \Sigma \to \mathbb{R}^d$ a given function. To express the logarithmic moment generating function, we introduce the nonnegative matrix

$$\pi_\lambda(i,j) = \pi(i,j) e^{\langle \lambda, f(j) \rangle} \quad (i, j \in \Sigma).$$

Theorem 9.1 (Perron-Frobenius)**.** *Let $\mathbf{B} = \{B(i,j)\}_{i,j=1}^{|\Sigma|}$ be an irreducible matrix. Then B possesses an eigenvalue ρ (called the Perron–Frobenius eigenvalue) such that:*

(1) $\rho > 0$ is real.
(2) For any eigenvalue λ of B, $|\lambda| \leq \rho$.
(3) There exist left and right eigenvectors corresponding to the eigenvalue ρ that have strictly positive coordinates.
(4) The left and right eigenvectors μ, θ corresponding to the eigenvalue ρ are unique up to a constant multiple.
(5) For every $i \in \Sigma$ and every $\varphi = \left(\varphi_1, \cdots, \varphi_{|\Sigma|}\right)$ such that $\varphi_j > 0$ for all j,

$$\lim_{n\to\infty} \frac{1}{n} \log \left[\sum_{j=1}^{|\Sigma|} B^n(i,j) \varphi_j \right] = \lim_{n\to\infty} \frac{1}{n} \log \left[\sum_{j=1}^{|\Sigma|} \varphi_j B^n(j,i) \right] = \log \rho.$$

Proof. See, for example, Wikipedia. □

Theorem 9.2. *Let $\{Y_k\}$ be a finite state Markov chain possessing an irreducible transition matrix $\mathbf{\Pi}$. For every $z \in \mathbb{R}^d$, define*

$$I(z) = \sup_{\lambda \in \mathbb{R}^d} \{ \langle \lambda, z \rangle - \log \rho(\mathbf{\Pi}_\lambda) \}$$

Then the empirical mean Z_n satisfies the LDP with the convex, good rate function I. Explicitly, for any set $\Gamma \subseteq \mathbb{R}^d$, and any initial state $\sigma \in \Sigma$,

$$-\inf_{z \in \mathring{\Gamma}} I(z) \leq \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_\sigma^\pi(Z_n \in \Gamma) \leq \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_\sigma^\pi(Z_n \in \Gamma) \leq -\inf_{z \in \overline{\Gamma}} I(z).$$

Proof. Consider the logarithmic generating function

$$\Lambda_n(\lambda) := \log \mathbb{E}_\sigma^\pi \left[e^{\langle \lambda, Z_n \rangle} \right].$$

By Gärtner–Ellis theorem (Theorem 8.7), it is enough to check that the limit

$$\Lambda(\lambda) := \lim_{n\to\infty} \frac{1}{n}\Lambda_n(n\lambda)$$

exists, is finite and differentiable everywhere in $\mathbb{R}^d$, and satisfies $\Lambda(\lambda) = \log \rho(\mathbf{\Pi}_\lambda)$. To begin, note that

$$\Lambda_n(\lambda) = \log \sum_{y_1,\cdots,y_n} \mathbb{P}^\pi_\sigma(Y_1 = y_1, \cdots, Y_n = y_n) \prod_{i=1}^n e^{\langle \lambda, f(y_k)\rangle} = \log \sum_{y_n} (\mathbf{\Pi}_\lambda)^n(\sigma, y_n).$$

Since $\mathbf{\Pi}_\lambda$ is irreducible, we have

$$\Lambda(\lambda) = \log \rho(\mathbf{\Pi}_\lambda).$$

To show that it is differentiable, we apply the implicit function theorem. Explicitly, consider the functions $F_y : \mathbb{R} \times \mathbb{R}^{|\Sigma|} \times \mathbb{R} \to \mathbb{R} \times \mathbb{R}^{|\Sigma|}$, parametrized by $y \in \mathbb{R}^{|\Sigma|}$, defined by

$$F_y(\rho, x, \lambda) = (\langle y, x\rangle - 1, \mathbf{\Pi}_\lambda x - \rho x).$$

Clearly, F_y is continuously differentiable. Hence, it suffices to show that for every λ_0 with the associated Perron-Frobenius eigenvalue ρ_0 and left and right eigenvectors y_0 and x_0 of $\mathbf{\Pi}_{\lambda_0}$ satisfying $\langle y_0, x_0\rangle = 1$, we have (1) $F_{y_0}(\rho_0, x_0, \lambda_0) = 0$ and (2) the Jacobian matrix

$$\partial_{\rho,x} F_{y_0}(\rho_0, x_0, \lambda_0) = \begin{pmatrix} 0 & y_0^T \\ -x_0 & \mathbf{\Pi}_{\lambda_0} - \rho_0 I. \end{pmatrix}$$

is invertible. Indeed, (1) is clear and (2) holds for if otherwise, there exists $(u, v) \in \mathbb{R} \times \mathbb{R}^{|\Sigma|} \setminus \{(0, 0)\}$ such that

$$\begin{cases} \langle y_0, v\rangle = 0, \\ \left\langle y_0, -x_0 + \mathbf{\Pi}_{\lambda_0} v - \rho_0 v\right\rangle = 0, \end{cases}$$

which is impossible. This implies, by the implicit function theorem, that there exists a continuously differentiable function $\lambda \mapsto (\rho(\lambda), x(\lambda))$ on a neighborhood of λ_0 such that $F_{y_0}(\rho(\lambda), x(\lambda), \lambda) = 0$. In particular, $\lambda \mapsto \log \rho(\mathbf{\Pi}_\lambda) = \log \rho(\lambda)$ is continuously differentiable on a neighborhood of λ_0. The proof is finished by noting that λ_0 is arbitrary. □

As a corollary, we have the following.

Corollary 9.3. *The empirical averages $L^{\mathbf{Y}}_n(i) = \frac{1}{n}\sum_{k=1}^n \mathbb{1}_i(Y_k)$, with $i \in \Sigma$, satisfies the LDP with the good rate function*

$$I(q) = \sup_{\lambda\in\mathbb{R}^d} \{\langle \lambda, q\rangle - \log \rho(\mathbf{\Pi}_\lambda)\} = \begin{cases} \sup_{u>0} \sum_{j\in\Sigma} q_j \log\left[\frac{u_j}{(\mathbf{u\Pi})_j}\right] & q \in M_1(\Sigma), \\ \infty & q \notin M_1(\Sigma) \end{cases}$$

where $\pi_\lambda(i, j) = \pi(i, j)e^{\lambda_j}$ and the inequality between vectors is compared entrywise.

Proof. The first equality follows immediately from Theorem 9.2 by taking

$$f = \left(\mathbb{1}_1, \mathbb{1}_2, \cdots, \mathbb{1}_{|\Sigma|}\right).$$

To prove the second inequality, we first note the one inequality is more obvious than the other: By taking $\mathbf{u}$ to be the left probability eigenvector of $\mathbf{\Pi}_\lambda$, we see the inequality "$\le$". To prove the other inequality, assume $q \in M_1(\Sigma)$ and choose $\lambda_j = \log\big[u_j/(\mathbf{u\Pi})_j\big]$ so that $\mathbf{u\Pi}_\lambda = \mathbf{u}$ and that $\rho(\mathbf{\Pi}_\lambda) = 1$. Therefore, by definition,

$$I(q) \ge \sum_{j=1}^{|\Sigma|} q_j \log\left[\frac{u_j}{(\mathbf{u\Pi})_j}\right],$$

finishing the proof. □

We also consider a derived process $\{(Y_k Y_{k+1})\}_{k\ge 0}$ of consecutive pairs to obtain a Sanov's theorem. Such a process has the transition matrix $\mathbf{\Pi}^{(2)}$ defined by

$$\pi^{(2)}((k\times\ell, i\times j)) = \mathbb{1}_\ell(i)\,\pi(i,j).$$

As is discussed in the following, one can determine the large deviations of the pair empirical measure

$$L^{\mathbf{Y}}_{n,2}(y) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_y(Y_{i-1}Y_i)$$

For $q \in M_1(\Sigma^2)$, we write

$$q_1(i) = \sum_{j=1}^{|\Sigma|} q(i,j) \text{ and } q_2(i) = \sum_{j=1}^{|\Sigma|} q(j,i),$$

and call q shift-invariant if $q_1 = q_2$.

Theorem 9.4. *Assume that $\mathbf{\Pi}$ is irreducible. Then for every probability measure $q \in M_1(\{(i,j) : \pi(i,j) > 0\})$,*

$$I_2(q) = \begin{cases} \sum_i q_1(i) H(q(\cdot\,|i) \mid \pi(i,\cdot)) & \text{if } q \text{ is shift-invariant,}\\ \infty & \text{otherwise.}\end{cases}$$

Proof. By Corollary 9.3,

$$I_2(q) = \sup_{u>0} \sum_{i,j\in\Sigma} q(i,j)\log\left[\frac{u_{i,j}}{(\mathbf{u\Pi}^{(2)})_{i,j}}\right] = \sup_{u>0}\sum_{i,j\in\Sigma} q_{i,j}\log\left[\frac{u_{i,j}}{\sum_k u_{k,i}\pi_{i,j}}\right].$$

If q is not invariant, then $q_1(j_0) < q_2(j_0)$ for some j_0. For $\mathbf{u}$ such that $u(\cdot, j) = 1$ when $j \ne j_0$ and $u(\cdot, j_0) = e^\alpha$, we have $I_2(q) = \infty$ if we let $\alpha \to \infty$.

If q is invariant,

$$\sum_{i,j\in\Sigma} q(i,j)\log\left[\frac{\sum_k u_{k,i} q_2(j)}{\sum_k u_{k,j} q_1(i)}\right] = 0.$$

Hence,

$$I_2(q) - \sum_i q_1(i) H(q(\cdot\,|i) \mid \pi(i|\,\cdot)) = \sup_{u>0}\sum_{i,j\in\Sigma} q(i,j)\log\left[\frac{u_{i,j}q_1(i)}{\sum_k u_{k,i} q(i,j)}\right]$$

$$= \sup_{u>0}\left\{-\sum_j q_2(j) H(q'(\cdot\,|j) \mid u'(\cdot\,|j))\right\},$$

where $u'(\cdot\,|j) = \frac{u(\cdot,j)}{\sum_i u(i,j)}$ and $q'(\cdot\,|j) = \frac{q(\cdot,j)}{\sum_i q(i,j)}$. This implies

$$I_2(q) \leq \sum_{i\in\Sigma} q_1(i) H(q(\cdot\,|i) \mid \pi(i,\cdot)).$$

Taking $\mathbf{u} > 0$ approaching q proves the theorem. □

9.2. Random walks in $\mathbb{R}^d$.

In this subsection, we assume that $\mathcal{D}_\Lambda = \mathbb{R}^d$. Define

$$Z_n(t) = \frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor} X_i, \quad 0 \leq t < 1,$$

and let μ_n denote the law of $Z_n(\cdot)$ (as a measure on $L^\infty([0,1])$).

Theorem 9.5 (Mogulskii). *The measures μ_n satisfy in $(L^\infty([0,1]), \|\cdot\|_\infty)$ the LDP with the good rate function*

$$I(\varphi) = \begin{cases} \int_0^1 \Lambda^*(\dot\varphi(t))dt \text{ if } \varphi \in \mathcal{AC},\ \varphi(0)=0, \\ \infty \qquad\qquad\qquad\quad \text{otherwise,} \end{cases}$$

where

$$\mathcal{AC} = \{f \in C([0,1]) : f \text{ is absolutely continuous}\}.$$

Remark 9.6. *The derivative in Theorem 9.5 is well-defined and satisfies the fundamental theorem of calculus $\varphi(s) = \int_0^s \dot\varphi(t)dt$ since φ is absolutely continuous.*

Definition 9.7. Let $(\mathcal{Y}, d)$ be a metric space. The probability measures $\{\mu_\varepsilon\}$ and $\{\tilde\mu_\varepsilon\}$ on $\mathcal{Y}$ are said to be exponentially equivalent if there exists a probability space $(\Omega, \mathcal{B}_\varepsilon, \mathbb{P}_\varepsilon)$ and two family of $\mathcal{Y}$-valued random variables $\{Z_\varepsilon\}$ and $\{\tilde Z_\varepsilon\}$ with joint laws $\mathbb{P}_\varepsilon$ and marginals $\{\mu_\varepsilon\}$ and $\{\tilde\mu_\varepsilon\}$, respectively, such that $\{\omega : (Z_\varepsilon, \tilde Z_\varepsilon) \in \Gamma_\delta\} \in \mathcal{B}_\varepsilon$ and

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}_\varepsilon(\Gamma_\delta) = -\infty,$$

where

$$\Gamma_\delta = \{(y, \tilde y) : d(y, \tilde y) > \delta\}.$$

Proposition 9.8. *Suppose μ_ε and $\tilde\mu_\varepsilon$ are exponentially equivalent. If $\tilde\mu_\varepsilon$ satisfies the LDP with the good rate function I, then so does μ_ε with rate I.*

Proof. Since I is good, it suffices to show the μ_ε satisfies the weak LDP with rate I. For the upper bound, we observe that for every compact set K and every $x \in K$, there exists a finite open cover $\{A_{x_i}\}_{i=1}^N$ such that for every $\delta > 0$,

$$x_i \in K \text{ and } \inf_{x\in A_{x_i}} I(x) > I^\delta(x_i).$$

Hence,

$$\begin{aligned} \limsup_{\varepsilon\to 0} \varepsilon\log\mu_\varepsilon(K) &\leq \limsup_{\varepsilon\to 0} \varepsilon \log\left[\tilde\mu_\varepsilon\left(\cup_{i=1}^N A_i\right) + \mathbb{P}_\varepsilon(\Gamma_\delta)\right] \\ \leq -\min_{1\leq i\leq N} I^\delta(x_i) &\leq -\inf_{x\in K} I^\delta(x), \end{aligned}$$

which proves the upper bound when letting $\delta \to 0$. As for the lower bound, for every open set G, there exists $G^\delta := \{y \in G : d(y, G^c) \geq \delta\}$ such that

$$\mu_\varepsilon(G) \geq \tilde{\mu}_\varepsilon\big(G^\delta\big) + \mathbb{P}_\varepsilon\big(G^\delta\big),$$

amounting to

$$\liminf_{\varepsilon\to 0} \varepsilon \log \mu_\varepsilon(G) \geq -\inf_{x\in G^\delta} I(x).$$

The proof is hence complete by taking $\delta \to 0$. $\square$

Lemma 9.9. *Let $\tilde{\mu}_n$ denote the law of $\tilde{Z}_n(\cdot)$ in $L^\infty([0,1])$, where*

$$\tilde{Z}_n(t) := Z_n(t) + \left(t - \frac{\lfloor nt \rfloor}{n}\right) X_{\lfloor nt\rfloor+1} \tag{9.1}$$

is the polygonal approximation of $Z_n(t)$. Then the probability measures μ_n and $\tilde{\mu}_n$ are exponentially equivalent in $L^\infty([0,1])$.

Proof. The sets $\big\{\omega : \|\tilde{Z}_n - Z_n\| > \eta\big\}$ are obviously measurable. Note that

$$|\tilde{Z}_n(t) - Z_n(t)| \leq \frac{|X_{\lfloor nt\rfloor+1}|}{n}.$$

Thus, for any $\eta > 0$ and any $\lambda > 0$,

$$\mathbb{P}\big(\|\tilde{Z}_n - Z_n\| > \eta\big) \leq n\mathbb{P}(|X_1| > n\eta) \leq n\mathbb{E}\big[e^{\lambda|X_1|}\big]e^{-\lambda n\eta}.$$

Since $\mathcal{D}_\Lambda = \mathbb{R}^d$, it follows, by considering first $n \to \infty$ and then $\lambda \to \infty$, that for any $\eta > 0$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(\|\tilde{Z}_n - Z_n\| > \eta\big) = -\infty.$$

Therefore, the probability measures μ_n and $\tilde{\mu}_n$ are exponentially equivalent. $\square$

Lemma 9.10. *Let $\mathcal{X}$ consist of all the maps from $[0,1]$ to $\mathbb{R}^d$ such that $t = 0$ is mapped to the origin, and equip $\mathcal{X}$ with the topology of pointwise convergence on $[0,1]$, namely, the product topology. Then the probability measures $\tilde{\mu}_n$ of Lemma 9.9 (defined on $\mathcal{X}$ by the natural embedding) satisfy the LDP in this Hausdorff topological space with the good rate function I of (9.1).*

Lemma 9.11. *The probability measures $\tilde{\mu}_n$ are exponentially tight in the space $C_0([0,1])$ of all continuous functions $f : [0,1] \to \mathbb{R}^d$ such that $f(0) = 0$, equipped with the supremum norm topology.*

Proof of Theorem 9.5. Note that $C_0([0,1])$ is a measurable subset of $\big(\mathcal{X}, \tau_{\text{prod}}\big)$; explicitly, $\big\{f : [0,1] \to \mathbb{R}^d : f(0) = 0\big\}$ is closed and $C([0,1])$ is Borel since

$$C([0,1]) = \bigcap_{x\in\mathbb{Q}\cap[0,1]} \bigcap_{m=1}^\infty \bigcup_{n=1}^\infty \left\{ f : [0,1] \to \mathbb{R}^d : \sup_{y,z\in B_{n^{-1}}(x)} |f(y) - f(z)| \leq \frac{1}{m} \right\}.$$

In addition, since $\tilde{\mu}_n(C_0([0,1])) = 1$, we have that $\{\tilde{\mu}_n\}$ satisfies the LDP in $C_0([0,1])$ (in $\mathcal{X}$). Since (a) $\tilde{\mu}_n$ is exponentially tight in $\mathcal{X}$ by Lemma 9.11 and (b) the product topology is coarser than the norm-induced topology on $C_0([0,1])$, we infer from Theorem 5.2 that $\{\tilde{\mu}_n\}$ satisfies the LDP in $(C_0([0,1]), \|\cdot\|_\infty)$. Finally, since $C_0([0,1])$ is closed in $L^\infty([0,1])$, $\{\tilde{\mu}_n\}$ satisfies the LDP in $L^\infty([0,1])$. $\square$

Lemma 9.12. *Let J denote the collection of all ordered finite subsets of $(0,1]$. For any $j = \Big\{0 < t_1 < t_2 < \cdots < t_{|j|} \leq 1\Big\} \in J$ and any $f : [0,1] \to \mathbb{R}^d$, let $p_j(f)$*

denote the vector $\left(f(t_1), f(t_2), ..., f\left(t_{|j|}\right)\right) \in \left(\mathbb{R}^d\right)^{|j|}$ *. Then, the sequence of laws* $\left\{\mu_n \circ p_j^{-1}\right\}$ *satisfies the LDP in* $\left(\mathbb{R}^d\right)^{|j|}$ *with the good rate function*

$$I_j(\mathbf{z}) = \sum_{\ell=1}^{|j|} (t_\ell - t_{\ell-1}) \Lambda^* \left(\frac{z_\ell - z_{\ell-1}}{t_\ell - t_{\ell-1}} \right),$$

where $\mathbf{z} = \left(z_1, ..., z_{|j|}\right)$ *and* $t_0 = 0,\ z_0 = 0$*. In particular,* $\left\{\tilde{\mu}_n \circ p_j^{-1}\right\}$ *satisfies the LDP also with rate* I_j*.*

Proof. Note that $\mu_n \circ p_j^{-1}$ is the law of the vector

$$Z_n^j = \left(Z_n(t_1), Z_n(t_2), ..., Z_n\left(t_{|j|}\right)\right).$$

To study the LDP, we apply Gärtner–Ellis theorem (Theorem 3.5) and the contraction principle (Theorem 5.1) to the following transformed vectors

$$Y_n^j = \left(Z_n(t_1), Z_n(t_2) - Z_n(t_1), ..., Z_n\left(t_{|j|}\right) - Z_n\left(t_{|j|-1}\right)\right),$$

which satisfies the LDP with rate function (by writing $\mathbf{y} = \left(y_1, ..., y_{|j|}\right)$ and $\lambda = \left(\lambda_1, ..., \lambda_{|j|}\right)$):

$$\Lambda_j^*(\mathbf{y}) = \sum_{\ell=1}^{|j|} (t_\ell - t_{\ell-1}) \Lambda^* \left(\frac{y_\ell}{t_\ell - t_{\ell-1}} \right),$$

where

$$\Lambda_j(\lambda) = \sum_{\ell=1}^{|j|} (t_\ell - t_{\ell-1}) \Lambda(\lambda_\ell).$$

The LDP for $\left\{\tilde{\mu}_n \circ p_j^{-1}\right\}$ holds as a consequence of exponential equivalence. □

Corollary 9.13. *For any* $j \in J$*,* $\left\{\tilde{\mu}_n \circ p_j^{-1}\right\}$ *satisfies the LDP in* $\left(\mathbb{R}^d\right)^{|j|}$ *with the good rate function* I_j*.*

Proof of Lemma 9.10. By Dawson-Gärtner theorem (Theorem 5.10), the rate function governing the LDP is

$$I_{\mathcal{X}}(\varphi) = \sup_{0<t_0<t_1<...<t_k\le 1} \sum_{\ell=1}^{k} (t_\ell - t_{\ell-1}) \Lambda^* \left(\frac{\varphi(t_\ell) - \varphi(t_{\ell-1})}{t_\ell - t_{\ell-1}} \right).$$

Since $\Lambda^* \geq 0$, we may assume without loss of generality that $t_k = 1$.

We first demonstrate that $I_{\mathcal{X}} = I$. By convexity of Λ^* and Jensen's inequality, it is clear that $I_{\mathcal{X}} \leq I$. For the other inequality, first assume $\varphi \in \mathcal{AC}$ and write

$$\psi_k(t) = \begin{cases} \int_{\frac{\lfloor kt \rfloor}{k}}^{\frac{\lfloor kt \rfloor + 1}{k}} \dot{\varphi}(s) ds & \text{if } t \in [0,1), \\ \int_{1-\frac{1}{k}}^{1} \dot{\varphi}(s) ds & \text{if } t = 1, \end{cases}$$

so that

$$I_{\mathcal{X}}(\varphi) \geq \liminf_{k\to\infty} \sum_{\ell=1}^{k} \frac{1}{k} \Lambda^*\left(k\left[\varphi\left(\frac{\ell}{k}\right) - \varphi\left(\frac{\ell-1}{k}\right)\right]\right) = \liminf_{k\to\infty} \int_0^1 \Lambda^*(\psi_k(t))dt$$
$$\geq \int_0^1 \liminf_{k\to\infty} \Lambda^*(\psi_k(t))dt \geq \int_0^1 \Lambda^*(\dot\varphi(t))dt,$$

where the second line follows from Fatou's lemma and lower semicontinuity of Λ^*. If $\varphi \notin \mathcal{AC}$, there exist $\delta > 0$ and $s_1^n < t_1^n \leq \ldots \leq s_{k_n}^n < t_{k_n}^n$ such that

$$\lim_{n\to 0} \sum_{\ell=1}^{k_n} t_\ell^n - s_\ell^n = 0 \quad \text{and} \quad \sum_{\ell=1}^{k_n} \|\varphi(t_\ell^n) - \varphi(s_\ell^n)\| \geq \delta.$$

Observe that by non-negativity of Λ^*,

$$I_{\mathcal{X}}(\varphi) \geq \sup_{0\leq s_1<t_1\leq s_2<t_2\leq\ldots\leq s_k<t_k} \sum_{\ell=1}^{k} (t_\ell - s_\ell)\Lambda^*\left(\frac{\varphi(t_\ell)-\varphi(s_\ell)}{t_\ell - s_\ell}\right)$$
$$= \sup_{\substack{0\leq s_1<t_1\leq s_2<t_2\leq\ldots\leq s_k<t_k \\ \lambda_1,\ldots,\lambda_k\in\mathbb{R}^d}} \sum_{\ell=1}^{k} [\langle \lambda_\ell, \varphi(t_\ell) - \varphi(s_\ell)\rangle - (t_\ell - s_\ell)\Lambda(\lambda_\ell)].$$

If one chooses, for any $\rho > 0$,

$$\lambda_\ell = \rho \cdot \frac{\varphi(t_\ell) - \varphi(s_\ell)}{\|\varphi(t_\ell) - \varphi(s_\ell)\|} \quad \text{and} \quad M = \sup_{\|\lambda\|=\rho} \Lambda(\lambda),$$

it follows that $I_{\mathcal{X}}(\varphi) \geq \rho\delta \to \infty$ as $\rho \to \infty$. □

Proof of Lemma 9.11. We first claim that given any real-valued random variable $X \sim \nu$, we have $E\big[e^{\delta\Lambda_\nu^*(X)}\big] < \infty$ for all $\delta < 1$. Indeed, by Markov's inequality and Lemma 3.4,

$$E\big[e^{\delta\Lambda_\nu^*(X)}\big] = \int_{-\infty}^{\overline{x}} e^{\delta\Lambda_\nu^*(x)} d\nu(x) + \int_{\overline{x}}^{\infty} e^{\delta\Lambda_\nu^*(x)} d\nu(x)$$
$$\leq \int_{-\infty}^{\overline{x}} \frac{d\nu(x)}{\nu((-\infty,x])^\delta} + \int_{\overline{x}}^{\infty} \frac{d\nu(x)}{\nu([x,\infty))^\delta},$$

By integration by part, the former integral can be evaluated as

$$\int_{-\infty}^{\overline{x}} \frac{d\nu(x)}{\nu((-\infty,x])^\delta} = \nu((-\infty,\overline{x}])^{1-\delta} + \delta\int_{-\infty}^{\overline{x}} \frac{d\nu(x)}{\nu((-\infty,x])^\delta},$$

and thus is bounded from above by $(1-\delta)^{-1}$. Since a similar estimates holds also for the latter integral, the claim is proved.

We now apply the inverse contraction principle to $\tilde\mu_n$ in $C_0([0,1])$ when equipped with the supremum norm topology. Denote by X_1^j the j-th component of X_1, by Λ_j the logarithmic moment generating function X_1^j, and by Λ_j^* the Fenchel-Legendre transform of Λ_j. We claim that $\tilde\mu_n$ is exponentially tight via the closure of the sets

$$K_\alpha = \cap_{j=1}^d K_\alpha^j \text{ where } K_\alpha^j := \left\{\varphi \in \mathcal{AC} : \varphi(0) = 0, \int_0^1 \Lambda_j^*\big(\dot\varphi_j(\theta)\big)d\theta \leq \alpha\right\}.$$

To this end, note that $\frac{d\tilde Z_n}{dt}(t) = X_{\lfloor nt\rfloor+1}$ and hence

$$\tilde{\mu}_n(K_\alpha^c) \le d \max_{1\le j\le d} \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n \Lambda_j^*\big(X_i^j\big) > \alpha\right).$$

By Markov's inequality and our claim,

$$\frac{1}{n}\log\tilde{\mu}_n(K_\alpha^c) \le -\delta\alpha + \log d + \max_{1\le j\le d}\log\mathbb{E}\Big[e^{\delta\Lambda_j^*(X_1^j)}\Big] \to \infty \text{ as } \alpha\to\infty.$$

It remains to show that K_α is precompact by the Arzelà–Ascoli theorem (uniformly bounded and equicontinuous). To this end, observe that by convexity of Λ_j^* and Jensen's inequality, for any $\varphi \in \mathcal{AC}$ and any $0 \le s < t \le 1$,

$$\Lambda_j^*\left(\frac{\varphi_j(t)-\varphi_j(s)}{t-s}\right) \le \frac{1}{t-s}\int_s^t \Lambda_j^*\big(\dot{\varphi}_j(\theta)\big)d\theta \le \frac{\alpha}{t-s}$$

Since $\Lambda_j^*(x) \ge M\|x\| - \max\big\{\Lambda_j(M), \Lambda_j(-M)\big\}$ for all $M > 0$, we deduce from the inequality above that for all $0 < t-s < \delta$,

$$\|\varphi_j(t)-\varphi_j(s)\| \le \frac{1}{M}\big(\alpha + \delta\max\big\{\Lambda_j(M),\Lambda_j(-M)\big\}\big),$$

from which the equicontinuity follows naturally, and so does the uniform boundedness due to the fact $\varphi(0) = 0$. ◻

Mogulskii's theorem (Theorem 9.5) extends to the laws ν_ε of

$$Y_\varepsilon(t) = \varepsilon\sum_{i=1}^{\lfloor \frac{t}{\varepsilon}\rfloor} X_i, 0 \le t \le 1.$$

Theorem 9.14. *The probability measures ν_ε induced on $L^\infty([0,1])$ by Y_ε satisfy the LDP with the good rate function I in Theorem 9.5.*

Proof. It suffices to prove the LDP on arbitrary subsequences $\varepsilon_m \searrow 0$. To this end, let $\varepsilon_{n_m} = \Big\lfloor \frac{1}{\varepsilon_m}\Big\rfloor$ so that μ_{n_m} satisfies the LDP with rate I. It suffices to show the exponential equivalence, namely,

$$\frac{1}{n_m}\log\mathbb{P}\Big(\|Y_{\varepsilon_m} - Z_{n_m}\| \ge \delta\Big) \to -\infty.$$

Indeed,

$$\|Y_{\varepsilon_m} - Z_{n_m}\| \le 2\varepsilon_m \max_{1\le i\le n_m}|X_i|,$$

implying, due to the fact $0 \in \mathring{\mathcal{D}}_\Lambda$,

$$\frac{1}{n_m}\log\mathbb{P}\Big(\|Y_{\varepsilon_m} - Z_{n_m}\| \ge \delta\Big) \le \frac{1}{n_m}\log n_m + \frac{1}{n_m}\log\mathbb{P}\left(|X_1| \ge \frac{\delta}{2\varepsilon_m}\right) \to -\infty.$$

◻

Now consider the process $w_\varepsilon(t) = w_t$ with w_t the standard Brownian motion. We study the LDP of w_ε in $C_0([0,1])$. To this end, consider the space $H_1 := \Big\{\int_0^t f(s)ds : f \in L^2([0,1])\Big\}$ equipped with the norm $\|\varphi\|_{H_1} = \|\dot{\varphi}\|_{L^2}$.

Lemma 9.15. *For any $d \in \mathbb{N}$ and any $\tau, \varepsilon, \delta > 0$,*

$$\mathbb{P}\left(\sup_{0\le t\le \tau}\|w_\varepsilon(t)\|\ge\delta\right)\le 4de^{-\frac{\delta^2}{2d\tau\varepsilon}}.$$

Proof. Observe that

$$\mathbb{P}\left(\sup_{0\le t\le \tau}\|w_\varepsilon(t)\|\ge\delta\right)=\mathbb{P}\left(\sup_{0\le t\le \tau}\|w_\varepsilon\|^2\ge\varepsilon^{-1}\delta^2\right)\le d\cdot\mathbb{P}\left(\sup_{0\le t\le \tau}(w_\varepsilon)_1^2\ge\frac{\delta^2}{d\varepsilon}\right)$$

By time rescaling,

$$P\left(\sup_{0\le t\le \tau}\|w_\varepsilon(t)\|\ge\delta\right)\le d\cdot\mathbb{P}\left(\sup_{0\le t\le 1}\left\|(w_\varepsilon)_1\right\|\right)\ge\frac{\delta}{\sqrt{d\tau\varepsilon}}.$$

Now for one-dimensional Brownian motion,

$$\mathbb{P}\left(\sup_{0\le t\le 1}\left\|(w_\varepsilon)_1\right\|\ge\eta\right)\le 2\mathbb{P}\left(\sup_{0\le t\le 1}(w_\varepsilon)_1\ge\eta\right)=4\mathbb{P}\left(\sup_{0\le t\le 1}(w_\varepsilon)_1\ge\eta\right)\le 4e^{-\frac{\eta^2}{2}},$$

where the equality follows from the Désiré André's reflection principle. $\square$

Theorem 9.16 (Schilder)**.** *The measure $\{\nu_\varepsilon\}$ satisfies, in $C_0([0,1])$, an LDP with good rate function*

$$I_w(\varphi)=\begin{cases}\frac{1}{2}\int_0^1|\dot\varphi(t)|^2\,dt \text{ if } & \varphi\in H_1,\\ \infty & \text{otherwise.}\end{cases}$$

Proof. Observe that $\hat w_\varepsilon(t)\coloneqq w_\varepsilon\big(\varepsilon\lfloor\frac{t}{\varepsilon}\rfloor\big)$ is the process of Y_ε with X_i the standard normal random variables in $\mathbb{R}^d$ with

$$\Lambda(\lambda)=\log\mathbb{E}\left[e^{\langle\lambda,X_1\rangle}\right]=\frac{\|\lambda\|^2}{2}\quad\text{and}\quad\Lambda^*(x)=\frac{\|x\|^2}{2}.$$

Hence, by Theorem 9.14, we have that the $\hat w_\varepsilon$ satisfies the LDP on $L^\infty([0,1])$ with rate function I_w. Now that $\hat w_\varepsilon$ and w_ε are exponentially equivalent:

$$\mathbb{P}(\|\hat w_\varepsilon-w_\varepsilon\|\ge\delta)\le\left(\left\lfloor\frac{1}{\varepsilon}\right\rfloor+1\right)\mathbb{P}\left(\sup_{0\le t\le\varepsilon}|w_\varepsilon(t)|\ge\delta\right)\le 4d\varepsilon^{-1}(1+\varepsilon)e^{-\frac{\delta^2}{2d\varepsilon^2}}.$$

Hence, w_ε satisfies the LDP with rate I_w on $L^\infty([0,1])$, which carries naturally to $C_0([0,1])$. $\square$