The metric of large deviation convergence

TIEFENG JIANG[1] and GEORGE L. O'BRIEN [2]

June 21, 1999

Abstract.

We construct a metric space of set functions $(\mathcal{Q}(\mathcal{X}), d)$ such that a sequence $\{P_n\}$ of Borel probability measures on a metric space $(\mathcal{X}, d^*)$ satisfies the full Large Deviation Principle (LDP) with speed $\{a_n\}$ and good rate function I if and only if the sequence $\{P_n^{a_n}\}$ converges in $(\mathcal{Q}(\mathcal{X}), d)$ to the set function e^{-I}. Weak convergence of probability measures is another special case of convergence in $(\mathcal{Q}(\mathcal{X}), d)$. Properties related to the LDP and to weak convergence are then characterized in terms of $(\mathcal{Q}(\mathcal{X}), d)$.

Running head: The metric of large deviations.

[1]Department of Statistics, 370 Serra Mall, Stanford University, Stanford, CA94305-4065, tjiang@stat.stanford.edu

[2]Department of Mathematics and Statistics, York University, Toronto, Ontario, Canada M2N 3T5, obrien@yorku.ca. Research of this author supported in part by the Natural Sciences and Engineering Research Council of Canada.

Key Words: large deviations, metric spaces.

AMS (1991) subject classifications: 60F10, 28C15, 60B10.

1 INTRODUCTION

The Large Deviation Principle (LDP), formulated in Varadhan[14], has been the subject of extensive study in the last decades. For expositions of this area of research, and a vast collection of applications, see the books Varadhan[15], Freidlin and Wentzell[8], Deuschel and Stroock[5], Dembo and Zeitouni[4], and Dupuis and Ellis[7].

Let $\{P_n; n \geq 1\}$ be a sequence of probability measures on a metric space $(\mathcal{X}, d^*)$, equipped with its Borel σ-algebra $\mathcal{B}(\mathcal{X})$ and let a_n be a sequence in $(0,1]$ with $a_n \to 0$. A function $I : \mathcal{X} \to [0,\infty]$ is called a *good rate function* if $I^{-1}([0,\gamma]) = \{x : I(x) \leq \gamma\}$ is a non-empty compact set for all $\gamma \in [0,\infty)$. Hereafter we denote the set of all good rate functions by $\mathcal{I}_g(\mathcal{X})$ and the set of all lower semicontinuous $I : \mathcal{X} \to [0,\infty]$ by $\mathcal{I}(\mathcal{X})$. We say that $\{P_n; n \geq 1\}$ satisfies the full LDP with rate I and speed $\{a_n\}$ if $I \in \mathcal{I}_g(\mathcal{X})$,

$$\liminf_{n\to\infty} a_n \log P_n(G) \geq -\inf_{x\in G} I(x) \tag{1.1}$$

for every open set $G \subset \mathcal{X}$, and

$$\limsup_{n\to\infty} a_n \log P_n(F) \leq -\inf_{x\in F} I(x) \tag{1.2}$$

for every closed set $F \subset \mathcal{X}$. We say that $\{P_n\}$ satisfies the weak LDP with rate $I \in \mathcal{I}(\mathcal{X})$ and speed $\{a_n\}$ if (1.1) holds for all open $G \subset \mathcal{X}$, and (1.2) holds for every compact set $F \subset \mathcal{X}$.

In this context, Varadhan[14] established the following integral theorem: If $\{P_n\}$ satisfies the full LDP with rate I and speed a_n then for every $\Phi \in C_b(\mathcal{X})$

$$\lim_{n\to\infty} a_n \log\left(\int_{\mathcal{X}} e^{\Phi/a_n} dP_n\right) = \sup_{x\in\mathcal{X}}\{\Phi(x) - I(x)\}. \tag{1.3}$$

Let $\mathcal{Q}_1(\mathcal{X})$ be the collection of all set functions μ either of the form $\mu(A) = (P(A))^\alpha$, $A \in \mathcal{B}(\mathcal{X})$, for some probability measure P and $\alpha \in (0,1]$, or of the form $\mu(A) = \exp(-\inf_{x\in A} I(x))$, $A \in \mathcal{B}(\mathcal{X})$, for some $I \in \mathcal{I}_g(\mathcal{X})$. Hereafter, we denote by $\mu = P^\alpha$ the former set functions and by $\mu = e^{-I}$ the latter.

In this work we construct a metric d on $\mathcal{Q}_1(\mathcal{X})$ such that a sequence $\{P_n\}$ of probability measures satisfies the full LDP with rate I and speed a_n if and only if $d(P_n^{a_n}, e^{-I}) \to 0$ (see Theorems 2.1 and 2.2 below).

The space $\mathcal{Q}_1(\mathcal{X})$ may seem somewhat artificial. Hence, we construct d on the space $\mathcal{Q}(\mathcal{X})$ of all set functions obeying four natural properties, with $\mathcal{Q}_1(\mathcal{X})$ identified as a subspace of $\mathcal{Q}(\mathcal{X})$. A full LDP is then merely a special case of convergence in the metric space $(\mathcal{Q}(\mathcal{X}), d)$. While similar to the set SA of subadditive capacities defined in O'Brien and Vervaat[11] and O'Brien[10], $\mathcal{Q}(\mathcal{X})$ is in general different from SA. In particular, for any probability measure P, P^α is in SA if and only if P is tight.

Weak convergence of probability measures occurs in the space $(\mathcal{Q}_1(\mathcal{X}), d)$ by taking $a_n = 1$ for all n and replacing the limit point e^{-I} by a probability measure. This explains the many similarities between the tools involved in proving the LDP and those involved in proving weak convergence of probability measures (see Pukhalskii[13], Table 1, for a list of such similarities).

A sequence of probability measures $\{P_n; n \geq 1\}$ on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ is said to be *exponentially tight* with respect to speed $\{a_n\}$ if, for every $\epsilon > 0$, there exists a compact set $K \subset \mathcal{X}$ such that $\limsup_{n\to\infty} P_n(K^c)^{a_n} < \epsilon$. In Theorem 2.3, we show that a sequence of tight probability measures $\{P_n\}$ is exponentially tight with respect to speed a_n if and only if $\{P_n^{a_n}\}$ is totally bounded in the metric space $(\mathcal{Q}_1(\mathcal{X}), d)$. In Corollary 2.2, we show that $(\mathcal{Q}_1(\mathcal{X}), d)$ is complete and separable when $(\mathcal{X}, d^*)$ is complete and separable.

We define a subadditive-integral $f \mapsto \tilde{\int} f d\mu$ for $\mu \in \mathcal{Q}_1(\mathcal{X})$ and f Borel measurable (see Definition 2.6). Using this subadditive-integral, we present Theorem 2.4, which extends both the Portmanteau theorem of weak convergence and Dudley[6], Theorem 11.3.3, to the context of $(\mathcal{Q}_1(\mathcal{X}), d)$. Varadhan's integral theorem is another direct corollary of Theorem 2.4, which allows us to conclude that the convergence in Varadhan's integral theorem is uniform with respect to any class of functions which are uniformly bounded and equicontinuous (see Corollary 2.4).

Our main results are stated in Section 2, with the proofs provided in Section 3. The paper is self-contained except for the proof of Proposition 2.3.

2 MAIN RESULTS

Throughout this paper, we will assume that $(\mathcal{X}, d^*)$ is a metric space, with $\mathcal{B}(\mathcal{X})$ its Borel σ-algebra. Let $\mathcal{F}, \mathcal{G}$ and $\mathcal{K}$ be the collection of all closed, open and compact subsets of $\mathcal{X}$, respectively. We denote the set of all probability measures on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ by $\mathcal{P}(\mathcal{X})$. With $d^*(x, A) = \inf_{y\in A} d^*(x, y)$, for every $\epsilon > 0$ let $A^\epsilon = \{x \in \mathcal{X};\ d^*(x, A) < \epsilon\}$

and $A^{-\epsilon} = ((A^c)^\epsilon)^c = \{x \in \mathcal{X};\ d^*(x, A^c) \geq \epsilon\}$. In particular, $A^\epsilon \in \mathcal{G}$ while $A^{-\epsilon} \in \mathcal{F}$. We leave it to the reader to check the following immediate consequences of the definition of $A^{-\epsilon}$.

Lemma 2.1 *For $\epsilon > 0$, $\delta > 0$ and a set $A \subset \mathcal{X}$*

(i) $(A^\epsilon)^{-\epsilon} \supset A \supset (A^{-\epsilon})^\epsilon$; and

(ii) $(A^{-\epsilon})^{-\delta} \supset A^{-\epsilon-\delta}$.

We first introduce the space $\mathcal{Q}(\mathcal{X})$.

Definition 2.1 *We define $\mathcal{Q}(\mathcal{X})$ to be the collection of set functions $\mu : \mathcal{B}(\mathcal{X}) \mapsto [0, \infty)$ which satisfy the following four conditions:*

(i) $\mu(\phi) = 0$;

(ii) for all A, $\mu(A) = \inf\{\mu(G); G \in \mathcal{G}, A \subset G\}$;

(iii) subadditivity: for any sequence $\{A_i, i \geq 1\}$,

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu(A_i);$$

(iv) for any $G \in \mathcal{G}$,

$$\mu(G) = \lim_{\delta \downarrow 0} \mu(G^{-\delta}).$$

Condition (ii) implies the monotonicity property $\mu(A) \leq \mu(B)$ when $A \subset B$. Also, to verify (iii), it is enough by (ii) to check it for open sets.

We construct a metric on $\mathcal{Q}(\mathcal{X})$ as follows.

Definition 2.2 *For any $\mu, \nu \in \mathcal{Q}(\mathcal{X})$, define*

$$\begin{aligned} d(\mu, \nu) = \inf\{\delta > 0; \quad & \mu(F) \leq \nu(F^\delta) + \delta \text{ for all } F \in \mathcal{F} \text{ and} \\ & \mu(G) \geq \nu(G^{-\delta}) - \delta \text{ for all } G \in \mathcal{G}\}. \end{aligned}$$

When restricted to $\mathcal{P}(\mathcal{X})$, d coincides with the Prohorov metric (see Billingsley[(3)], page 238). Indeed, let $\delta > 0$ be such that $\mu(F) \leq \nu(F^\delta) + \delta$ for all $F \in \mathcal{F}$. Hence, $\mu(G^c) \leq \nu((G^c)^\delta) + \delta$ for all $G \in \mathcal{G}$, that is $\mu(G) \geq \nu(G^{-\delta}) - \delta$.

Since $(\overline{A})^\delta = A^\delta$ and $(A^\circ)^{-\delta} = A^{-\delta}$ (here and in the sequel, $\overline{A}$ and A° are the closure and the interior of $A \subset \mathcal{X}$, respectively), it is easy to see that

$$\begin{aligned} d(\mu, \nu) = \inf\{\delta > 0; \quad & \mu(A) \leq \nu(A^\delta) + \delta \text{ and} \\ & \mu(A) \geq \nu(A^{-\delta}) - \delta \text{ for all } A \in \mathcal{B}(\mathcal{X})\}. \end{aligned} \tag{2.1}$$

It easily follows from part (i) of Lemma 2.1 that for any $\mu, \nu \in Q(\mathcal{X})$

$$\begin{aligned} &\{\delta > 0; \nu(F) \leq \mu(F^\delta) + \delta \text{ for all } F \in \mathcal{F}\} \\ &= \{\delta > 0;\ \mu(G) \geq \nu(G^{-\delta}) - \delta \text{ for all } G \in \mathcal{G}\}. \end{aligned} \tag{2.2}$$

This leads to a symmetric formulation with no sets of the form $A^{-\epsilon}$:

$$\begin{aligned} d(\mu, \nu) = \inf\{\delta > 0; \quad &\mu(F) \leq \nu(F^\delta) + \delta \text{ and} \\ &\nu(F) \leq \mu(F^\delta) + \delta \text{ for all } F \in \mathcal{F}\}. \end{aligned} \tag{2.3}$$

To complete the construction we will prove in Section 3 that

Theorem 2.1 *d is a metric on $\mathcal{Q}(\mathcal{X})$.*

We now define narrow convergence, which coincides with weak convergence when restricted to $\mathcal{P}(\mathcal{X})$ and is its natural extension to $\mathcal{Q}(\mathcal{X})$.

Definition 2.3 *For $\{\mu, \mu_n; n \geq 1\} \subset \mathcal{Q}(\mathcal{X})$, if*
(i) $\liminf_{n\to\infty} \mu_n(G) \geq \mu(G)$ *for every $G \in \mathcal{G}$ and*
(ii) $\limsup_{n\to\infty} \mu_n(F) \leq \mu(F)$ *for every $F \in \mathcal{F}$,*
then we say that $\{\mu_n\}$ converges narrowly to μ, which we denote by $\mu_n \Longrightarrow \mu$.

We say that $\{\mu_n\}$ converges vaguely to μ, denoted by $\mu_n \stackrel{v}{\Longrightarrow} \mu$, as soon as (i) holds and (ii) applies for every $F \in \mathcal{K}$.

Note that (i) and (ii) of Definition 2.3 are equivalent when μ and all μ_n are in $\mathcal{P}(\mathcal{X})$, but not in general.

Definition 2.4 *$\mu \in \mathcal{Q}(\mathcal{X})$ is said to be tight if for every $\epsilon > 0$ there exists $K \in \mathcal{K}$ such that $\mu(K^c) < \epsilon$.*

In analogy with the theory of weak convergence, the narrow convergence of Definition 2.3 characterizes the convergence in $(\mathcal{Q}(\mathcal{X}), d)$ to a tight limit point.

Theorem 2.2 *For μ_0 tight, $d(\mu_n, \mu_0) \to 0$ if and only if $\mu_n \Longrightarrow \mu_0$.*

The cone $\mathcal{Q}(\mathcal{X})$ contains all *sup-measures* $\mu(A) = \sup_{x \in A} \mu(\{x\})$ for which $x \mapsto \mu(\{x\})$ is upper semicontinuous; $\mathcal{Q}(\mathcal{X})$ also contains all set functions of the form $f(P(\cdot))$ with $P \in \mathcal{P}(\mathcal{X})$ and $f \in C_b([0,1])$ non-decreasing such that $f(0) = 0$ and $f(x+y) \leq f(x) + f(y)$ for all $x, y \in [0,1]$ for which $x + y \leq 1$. In particular, $\{P^\alpha, P \in \mathcal{P}(\mathcal{X}), 0 < \alpha \leq 1\} \cup \{e^{-I};\ I \in \mathcal{I}(\mathcal{X})\}$, is a subset of $\mathcal{Q}(\mathcal{X})$.

Comparing (1.1) and (1.2) with Definition 2.3 we see that for a sequence of probability measures $\{P_n\}$ and for $I \in \mathcal{I}(\mathcal{X})$, $\{P_n\}$ satisfies the weak LDP with rate I and speed $\{a_n\}$ if and only if $P_n^{a_n} \stackrel{v}{\Longrightarrow} e^{-I}$. Since $\mu = e^{-I}$ is tight for every $I \in \mathcal{I}_g(\mathcal{X})$, by Theorem 2.2, $\{P_n\}$ satisfies the full LDP with rate I and speed $\{a_n\}$ if and only if $I \in \mathcal{I}_g(\mathcal{X})$ and $d(P_n^{a_n}, e^{-I}) \to 0$.

Let $a_n = n^{-1}$ and P_n be the law of the empirical mean of n i.i.d. q-stable real-valued random variables (any $q \in (1,2)$). By Cramér's theorem, (1.1) and (1.2) hold for $I \equiv 0 \in \mathcal{I}(\mathcal{X})$ (c.f. Dembo and Zeitouni[4], Theorem 2.2.3) while, considering the sets $G_x = (x, \infty)$ for large x, we see that $d(P_n^{a_n}, e^{-I}) \to 1$. This example demonstrates that we can not dispense of tightness of e^{-I} when relating the metric d to the full LDP.

Topological issues related to those considered here were studied in O'Brien and Vervaat[11] and O'Brien[10]. Their class of set functions was somewhat different but the main difference between their results and the present work is that they did not present a metric corresponding to d. It turns out, however, that a similar construction to the above can be used to give a metric for the set of tight subadditive capacities, as defined in those papers.

The following is a straightforward corollary of Theorem 2.2 concerning uniform upper and lower bounds.

Corollary 2.1 *Suppose $\{P_n; n \geq 1\}$ satisfies the full LDP with rate I and speed $\{a_n\}$. Then,*

$$\limsup_{n\to\infty} \sup_{F \in \mathcal{F}_1} (P_n(F)^{a_n} - e^{-I(F)}) \leq 0,$$

for every $\mathcal{F}_1 \subset \mathcal{F}$ such that

$$\lim_{M\to\infty} \lim_{\delta\to 0} \sup_{F \in \mathcal{F}_1} |I(F^\delta) \wedge M - I(F) \wedge M| = 0\ .$$

Also,

$$\liminf_{n\to\infty} \inf_{G \in \mathcal{G}_1} (P_n(G)^{a_n} - e^{-I(G)}) \geq 0,$$

for every $\mathcal{G}_1 \subset \mathcal{G}$ such that

$$\lim_{M\to\infty} \lim_{\delta\to 0} \sup_{G\in\mathcal{G}_1} |I(G^{-\delta}) \wedge M - I(G) \wedge M| = 0 \,.$$

In particular, if the rate function I is continuous (so that $I(\cdot) \wedge M$ is uniformly continuous), the above two conditions hold for $\mathcal{F}_1 = \mathcal{F}$ and $\mathcal{G}_1 = \mathcal{G}$, and therefore the large deviation upper and lower bounds hold uniformly over $\mathcal{F}$ and $\mathcal{G}$ respectively.

Definition 2.5 *A collection $\{\mu_\alpha; \alpha \in J\} \subset \mathcal{Q}(\mathcal{X})$ is said to be uniformly tight if for every $\epsilon > 0$ there exists $K \in \mathcal{K}$ such that $\sup_{\alpha\in J} \mu_\alpha(K^c) < \epsilon$.*

The following theorem relates uniform tightness with total boundedness in $(\mathcal{Q}(\mathcal{X}), d)$.

Theorem 2.3 *(i) If the collection $\{\mu_\alpha; \alpha \in J\} \subset \mathcal{Q}(\mathcal{X})$ is uniformly tight and if $\sup_\alpha \mu_\alpha(\mathcal{X}) < \infty$, then the collection is totally bounded.*
(ii) If $(\mathcal{X}, d^*)$ is complete, the collection $\{\mu_\alpha; \alpha \in J\} \subset \mathcal{Q}(\mathcal{X})$ is totally bounded in $(\mathcal{Q}(\mathcal{X}), d)$ and each μ_α is tight, then the collection is uniformly tight and $\sup_\alpha \mu_\alpha(\mathcal{X}) < \infty$.

The next two propositions broaden the scope of some results in O'Brien and Vervaat[11] and O'Brien[10]. Also, since the exponential tightness of $\{P_n\}$ is exactly the uniform tightness of $\{P_n^{a_n}\}$ (provided each P_n is tight), the next proposition generalizes both Le Cam's theorem (compare with Dudley[6], Theorem 11.5.3) and the exponential tightness Lemma 2.6 of Lynch and Sethuraman[9].

Proposition 2.1 *For $\{\mu_n, n \geq 0\} \subset \mathcal{Q}(\mathcal{X})$, $\mu_n \Longrightarrow \mu_0$ and μ_n is tight for every $n \geq 0$, if and only if $\mu_n \overset{v}{\Longrightarrow} \mu_0$ and $\{\mu_n; n \geq 1\}$ is uniformly tight.*

The following result allows us to characterize the possible limits of sequences in $(\mathcal{Q}_1(\mathcal{X}), d)$, en route to proving the completeness of this space.

Proposition 2.2 *Let $\{\mu_n; n \geq 0\} \subset \mathcal{Q}_1(\mathcal{X})$. Suppose $\mu_n \Longrightarrow \mu_0$ and μ_0 is tight.*
(i) If $\mu_n = P_n^{a_n}$ for $n > 0$, where $P_n \in \mathcal{P}(\mathcal{X})$ and $\liminf_{n\to\infty} a_n = 0$, then $\mu_0 = e^{-I_0}$ for some $I_0 \in \mathcal{I}_g(\mathcal{X})$.
(ii) If $\mu_n = e^{-I_n}$ for $n > 0$, where $I_n \in \mathcal{I}_g(\mathcal{X})$, then $\mu_0 = e^{-I_0}$ for some $I_0 \in \mathcal{I}_g(\mathcal{X})$.
(iii) If $\mu_n = P_n^{a_n}$ for $n > 0$, where $P_n \in \mathcal{P}(\mathcal{X})$ and $\liminf_{n\to\infty} a_n = \alpha > 0$, then

$\mu_0 = P^\alpha$ *for some* $P \in \mathcal{P}(\mathcal{X})$.

(iv) In (i) and (iii), unless μ_0 *is the Dirac measure, the sequence* $\{a_n\}$ *must actually converge and, in (iii), we must also have* $P_n \Longrightarrow P$.

Proposition 2.3 *Suppose* $\mathcal{X}$ *is complete,* $\{\mu_n\}$ *is a Cauchy sequence in* $(\mathcal{Q}_1(\mathcal{X}), d)$ *and each* μ_n *tight. Then* $\{\mu_n\}$ *converges to a tight limit in* $(\mathcal{Q}_1(\mathcal{X}), d)$ *and the conclusions of Proposition 2.2 apply.*

We next show that for $(\mathcal{X}, d^*)$ separable, I is the rate of some weak LDP as soon as $\inf_x I(x) = 0$. The LDP is full if I is good.

Lemma 2.2 *If* $(\mathcal{X}, d^*)$ *is separable and* $I \in \mathcal{I}(\mathcal{X})$ *is such that* $\inf_x I(x) = 0$, *then for any* $a_n \to 0$ *there exists* $P_n \in \mathcal{P}(\mathcal{X})$ *such that* $P_n^{a_n} \Longrightarrow e^{-I}$.

Completeness and separability of the metric space $(\mathcal{Q}_1(\mathcal{X}), d)$ now follow from the same properties for $(\mathcal{X}, d^*)$.

Corollary 2.2 *If* $(\mathcal{X}, d^*)$ *is complete and separable, then so is* $(\mathcal{Q}_1(\mathcal{X}), d)$, *which is the completion of the set* $\{P^\alpha, P \in \mathcal{P}(\mathcal{X}), 0 < \alpha \le 1\}$.

As shown below, the approximate comparison principle of Baxter and Jain(2) holds for any complete metric space $\mathcal{X}$ and is a direct consequence of Theorem 2.2 and part (ii) of Proposition 2.2.

Corollary 2.3 *Suppose* $\mathcal{X}$ *is complete. For* $a > 0$ *and* $P, Q \in \mathcal{P}(\mathcal{X})$ *define*

$$\rho_a(P, Q) := \sup_{F \in \mathcal{F}} \{P(F) - Q(F^a)\}.$$

Suppose that for k *fixed,* $P_{n,k} \in \mathcal{P}(\mathcal{X})$ *satisfy the full LDP with rate* I_k *and speed* a_n, *and that* $Q_n \in \mathcal{P}(\mathcal{X})$ *satisfy for every* $a > 0$,

$$\lim_{k\to\infty} \limsup_{n\to\infty} a_n \log \rho_a(P_{n,k}, Q_n) = -\infty. \tag{2.4}$$

Then, $e^{-I_k} \Longrightarrow e^{-I}$ *for some* $I \in \mathcal{I}_g(\mathcal{X})$ *and* $\{Q_n\}$ *satisfies the full LDP with rate* I *and speed* a_n.

In particular, the large deviation comparison principle of Baxter and Jain[1] corresponds to the special case of Corollary 2.3 in which $P_{n,k} = P_n$ is independent of k.

We now define the subadditive-integral with respect to $\mu \in \mathcal{Q}_1(\mathcal{X})$ en route to our extended Portmanteau theorem.

Definition 2.6 *The* subadditive-*integral of a Borel function f on $A \in \mathcal{B}(\mathcal{X})$ is defined by*

$$\begin{aligned} \widetilde{\int}_A f d\mu &= (\int_A |f|^{1/\alpha} dP)^\alpha \text{ if } \mu = P^\alpha, \text{ where } P \in \mathcal{P}(\mathcal{X}),\ 0 < \alpha \le 1; \\ &= \sup_{x \in A} |f(x)| \mu(\{x\}) \text{ if } \mu = e^{-I}. \end{aligned}$$

Note that $\mu = P^\alpha = e^{-I}$ for some $P \in \mathcal{P}(\mathcal{X})$, $\alpha \in (0,1]$ and $I \in \mathcal{I}_g(\mathcal{X})$ only when $\mu = \delta_{x_0}$ for some x_0, in which case $\widetilde{\int}_A f d\mu = |f(x_0)| 1_A(x_0)$. Indeed, $\mu = P^\alpha = e^{-I}$ and $I(x_1) \le I(x_2) < \infty$ for some $x_1 \ne x_2$ together imply that $e^{-I(x_1)} = e^{-I(x_1) \wedge I(x_2)} = (e^{-I(x_1)/\alpha} + e^{-I(x_2)/\alpha})^\alpha$, which is impossible.

Let $\widetilde{\int}$ denote $\widetilde{\int}_{\mathcal{X}}$. The following properties of the subadditive-integral are elementary:

(a) $\widetilde{\int}_A f d\mu = \widetilde{\int} f 1_A d\mu$ for every A and f;

(b) $\widetilde{\int}_A c d\mu = |c| \mu(A)$ for every constant c;

(c) $\widetilde{\int} f d\mu \le \widetilde{\int} (f+g) d\mu \le \widetilde{\int} f d\mu + \widetilde{\int} g d\mu$ for every $f, g \ge 0$.

We denote the bounded Lipschitz norm on the space of all bounded, uniformly Lipschitz continuous functions by $\|f\|_{\mathrm{BL}} := \sup_{x \ne y} |f(x) - f(y)|/d^*(x,y) + \sup_x |f(x)|$, with the corresponding unit ball denoted by $B := \{f; \|f\|_{\mathrm{BL}} \le 1\}$.

Our last theorem is an extension of the Portmanteau theorem of weak convergence to the context of $(\mathcal{Q}_1(\mathcal{X}), d)$, establishing also the equivalence of the metrics $\beta(\mu,\nu) := \sup_{f \in B} |\widetilde{\int} f d\mu - \widetilde{\int} f d\nu|$ and $d(\mu,\nu)$ on the set $\mathcal{Q}_1(\mathcal{X})$ (c.f. (3.22) for $\beta(\mu,\nu) = 0$ implying $\mu = \nu$).

Theorem 2.4 *Let $\mu_n \in \mathcal{Q}_1(\mathcal{X})$ for $n \ge 0$ and suppose μ_0 is tight. The following are equivalent.*

(i) $\mu_n \Longrightarrow \mu_0$*;*

(ii) $d(\mu_n, \mu_0) \to 0$*;*

(iii) $\widetilde{\int} f d\mu_n \to \widetilde{\int} f d\mu_0$ *for every* $f \in C_b(\mathcal{X})$*;*

(iv) $\beta(\mu_n, \mu_0) \to 0$.

For μ_0 tight, (i) and (ii) are equivalent by Theorem 2.2. For $\mu_n \in \mathcal{P}(\mathcal{X})$, the fact that (i) implies (iii) is the classical Portmanteau theorem whereas, for $\mu_n = P_n^{a_n}$, $a_n \to 0$ and $\mu_0 = e^{-I}$, it is Varadhan's integral theorem. The equivalence of (ii) and (iv) provides an equivalent definition of d via bounded Lipschitz functions, with Dudley[6] corresponding to the special case of $\mu_n \in \mathcal{P}(\mathcal{X})$ tight.

The following easy corollary of Theorem 2.4 extends Dudley[6], Corollary 11.3.4, to the setting of $\mathcal{Q}_1(\mathcal{X})$, yielding in particular information about uniform convergence in Varadhan's integral theorem.

Corollary 2.4 *Let $\widetilde{B}$ be a uniformly bounded, equicontinuous collection of real-valued functions on the topological space $(\mathcal{X}, \tau)$. Suppose $\{\mu_n; n \geq 0\} \subset \mathcal{Q}_1(\mathcal{X})$ is such that $\mu_n \Longrightarrow \mu_0$, with μ_0 tight. Then,*

$$\limsup_{n\to\infty} \sup_{f\in\widetilde{B}} |\widetilde{\int} f d\mu_n - \widetilde{\int} f d\mu_0| = 0. \tag{2.5}$$

3 THE PROOFS.

Proof of Theorem 2.1. By (2.3), d is symmetric.

Now, suppose that $d(\mu, \nu) = 0$. Then, $\mu(G) \geq \nu(G^{-\delta}) - \delta$ for every $\delta > 0$ and $G \in \mathcal{G}$. Taking $\delta \downarrow 0$, by (iv) of Definition 2.1, $\mu(G) \geq \nu(G)$. By symmetry, $\mu(G) = \nu(G)$ for every $G \in \mathcal{G}$ and by (ii) of Definition 2.1 we conclude that $\mu = \nu$.

Turning to the proof of the triangle inequality, fix $\mu, \nu, \omega \in \mathcal{Q}(\mathcal{X})$ and $x > d(\mu, \omega)$, $y > d(\omega, \nu)$. Then, we have that

$$\mu(A) \leq \omega(A^x) + x, \quad \omega(A) \leq \nu(A^y) + y \tag{3.1}$$

and

$$\mu(A) \geq \omega(A^{-x}) - x, \quad \omega(A) \geq \nu(A^{-y}) - y \tag{3.2}$$

for all $A \in \mathcal{B}(\mathcal{X})$ (see (2.1)). From (3.1), for any $F \in \mathcal{F}$

$$\mu(F) \leq \omega(F^x) + x \leq \nu(F^{x+y}) + x + y. \tag{3.3}$$

Similarly, by Lemma 2.1 and (3.2), for any $G \in \mathcal{G}$,

$$\mu(G) \geq \omega(G^{-x}) - x \geq \nu(G^{-x-y}) - x - y. \tag{3.4}$$

Combining (3.3) and (3.4) we obtain $d(\mu,\nu) \le x+y$. Taking $x \downarrow d(\mu,\omega)$ and $y \downarrow d(\omega,\nu)$ we conclude that $d(\mu,\nu) \le d(\mu,\omega)+d(\omega,\nu)$. ∎

Proof of Theorem 2.2. As $\epsilon \downarrow 0$, by (iv) of Definition 2.1, $\mu_0(G^{-\epsilon}) \to \mu_0(G)$ for any $G \in \mathcal{G}$. By (ii) of Definition 2.1, $\mu_0(K^\epsilon) \to \mu_0(K)$ for any $K \in \mathcal{K}$. Since μ_0 is tight, also $\mu_0(F^\epsilon) \to \mu_0(F)$ for any $F \in \mathcal{F}$. Hence, by Definitions 2.2 and 2.3, $d(\mu_n,\mu_0) \to 0$ implies that $\mu_n \Longrightarrow \mu_0$.

Now, assume μ_0 is tight and $\mu_n \Longrightarrow \mu_0$. We will show first that for any $\delta > 0$

$$\limsup_{n\to\infty} \sup_{F\in\mathcal{F}} (\mu_n(F) - \mu_0(F^{2\delta})) \le 0. \tag{3.5}$$

Let $\epsilon > 0$ and let K be compact with $\mu_0(K^c) < \epsilon$. For $x \in K$, let $G_x := \{x\}^{\delta/2}$. Let $\{G_1,\cdots,G_m\}$ be a subcover of K and let $\Gamma_0 = \cup_{i=1}^m G_i$. There exists n_0 such that for all $n > n_0$,

$$\mu_n(\Gamma_0^c) \le \mu_0(\Gamma_0^c) + \epsilon \le \mu_0(K^c) + \epsilon < 2\epsilon.$$

For $F \in \mathcal{F}$, let $\Gamma_F :=$ union of those G_i's which intersect F. Then

$$F \cap \Gamma_0 \subset \Gamma_F \subset \overline{\Gamma}_F \subset F^{2\delta}.$$

Then, for $n > n_0$,

$$\begin{aligned} \mu_n(F) &\le \mu_n(F\cap\Gamma_0) + \mu_n(\Gamma_0^c) \\ &\le \mu_n(\overline{\Gamma}_F) + 2\epsilon + [\mu_0(F^{2\delta}) - \mu_0(\overline{\Gamma}_F)]. \end{aligned}$$

Therefore,

$$\mu_n(F) - \mu_0(F^{2\delta}) \le \mu_n(\overline{\Gamma}_F) - \mu_0(\overline{\Gamma}_F) + 2\epsilon \le \sup_{F\in\mathcal{F}}\{\mu_n(\overline{\Gamma}_F) - \mu_0(\overline{\Gamma}_F)\} + 2\epsilon. \tag{3.6}$$

Since $\limsup_n\{\mu_n(\overline{\Gamma}_F) - \mu_0(\overline{\Gamma}_F)\} \le 0$ for each $F \in \mathcal{F}$ and there are only finitely many possibilities for $\overline{\Gamma}_F$, we see from (3.6) that

$$\limsup_{n\to\infty} \sup_F \{\mu_n(F) - \mu_0(F^{2\delta})\} \le 2\epsilon. \tag{3.7}$$

Letting $\epsilon \downarrow 0$, we get (3.5).

Similarly, since each Γ_F is open,

$$\limsup_{n\to\infty} \sup_F \{\mu_0(F) - \mu_n(F^{2\delta})\} \le \limsup_n \sup_F \{\mu_0(\Gamma_F) - \mu_n(\Gamma_F)\} + \epsilon \le \epsilon.$$

Letting $\epsilon \downarrow 0$, we get

$$\limsup_{n\to\infty} \sup_{F\in\mathcal{F}} (\mu_0(F) - \mu_n(F^{2\delta})) \le 0. \tag{3.8}$$

Combining (3.5) and (3.8), we see that $d(\mu_n, \mu_0) \le 2\delta$ for all n large enough. Thus $d(\mu_n, \mu_0) \to 0$. ∎

Proof of Theorem 2.3. Suppose first that $(\mathcal{X}, d^*)$ is complete, μ_α is tight for every $\alpha \in J$ and $\{\mu_\alpha; \alpha \in J\}$ is totally bounded in $(\mathcal{Q}(\mathcal{X}), d)$. Then, fixing $\delta > 0$, the collection $\{\mu_\alpha; \alpha \in J\}$ has a finite cover by open balls of radius δ, centered at $\{\mu_\beta; \beta \in J'\}$ (in particular, $J' \subset J$ is a finite set). This shows immediately that $\sup_\alpha \mu_\alpha(\mathcal{X}) < \infty$. Moreover, the finite collection $\{\mu_\beta; \beta \in J'\}$ is uniformly tight, so $\sup_{\beta\in J'} \mu_\beta(K^c) < \delta$ for some $K := K(\delta) \in \mathcal{K}$. By Lemma 2.1,

$$\mu_\alpha((K^\delta)^c) \le \mu_\beta(((K^\delta)^c)^\delta) + \delta \le \mu_\beta(K^c) + \delta < 2\delta. \tag{3.9}$$

Now fix $\epsilon > 0$ and $\delta_n > 0$ such that $\sum_n \delta_n = \epsilon$. Let L be the closure of $\cap_{n\ge1} K(\delta_n)^{\delta_n}$, noting that $\mathcal{X}$ is complete and therefore $L \in \mathcal{K}$. By (3.9), for each $\alpha \in J$,

$$\mu_\alpha(L^c) \le \sum_{n\ge1} \mu_\alpha((K(\delta_n)^{\delta_n})^c) \le \sum_n 2\delta_n = 2\epsilon\ .$$

With $\epsilon > 0$ arbitrarily small we have the claimed uniform tightness of $\{\mu_\alpha; \alpha \in J\}$.

Assume now that $A := \sup_\alpha \mu_\alpha(\mathcal{X}) < \infty$ and that $\{\mu_\alpha; \alpha \in J\}$ is uniformly tight but is not totally bounded. Then, there exists $\delta > 0$ and $\{\mu_n; n \ge 1\} \subset \{\mu_\alpha; \alpha \in J\}$ such that $d(\mu_n, \mu_m) > 2\delta$ for all $n \ne m$, while $\sup_{n\ge1} \mu_n(K^c) < \delta$ for some $K \in \mathcal{K}$. By (2.3) there exist $F'_{n,m} \in \mathcal{F}$ such that, for $F_{n,m} = F'_{n,m} \cap K$,

$$\begin{aligned} \delta + \mu_n(F_{n,m}) > \mu_n(F'_{n,m}) &> \mu_m((F_{n,m})^{2\delta}) + 2\delta \quad \text{or} \\ \delta + \mu_m(F_{n,m}) > \mu_m(F'_{n,m}) &> \mu_n((F_{n,m})^{2\delta}) + 2\delta. \end{aligned} \tag{3.10}$$

Since $\{F_{n,m}; n \ne m\} \subset \mathcal{K}$, there exist $M = M(\delta) < \infty$ and $\{\Gamma_i; i = 1, \ldots, M\} \subset \mathcal{G}$ such that for every $n \ne m$ and some $i = i(n, m) \in \{1, \ldots, M\}$

$$F_{n,m} \subset \Gamma_i \subset (F_{n,m})^{2\delta}$$

(cf. the proof of Theorem 2.2 above). For $v_n = (\mu_n(\Gamma_1), \ldots, \mu_n(\Gamma_M)) \in R^M$, equipping R^M with the supremum norm $\|\cdot\|$, it follows from (3.10) that $\|v_n - v_m\| > \delta$ while $\sup_n \|v_n\| \le A$. This contradicts the compactness of $[-A, A]^M$. ∎

Proof of Proposition 2.1. Suppose that $\mu_n \Longrightarrow \mu_0$ and μ_n is tight for every $n \geq 0$. Fix $\epsilon > 0$ and $K_n \in \mathcal{K}$, $n \geq 0$ such that $\mu_n(K_n^c) < \epsilon$. Since $\mu_n \Longrightarrow \mu_0$ it follows that for every $\delta > 0$

$$\limsup_{n\to\infty} \mu_n((K_0^\delta)^c) \leq \mu_0((K_0^\delta)^c) \leq \mu_0(K_0^c) < \epsilon.$$

In particular, $b_n := 2\inf\{\delta > 0; \mu_n((K_0^\delta)^c) < \epsilon\} \to 0$. For $n > 0$, let $L_n := \overline{K_0^{b_n} \cap K_n}$ and let $L = \cup_{n=0}^\infty L_n$. Since each L_n is compact and every open cover of K_0 also covers all but finitely many of the L_n's, L is compact. By subadditivity of μ_n,

$$\mu_n(L^c) \leq \mu_n((K_0^{b_n} \cap K_n)^c) \leq \mu_n((K_0^{b_n})^c) + \mu_n(K_n^c) < 2\epsilon\ .$$

Since $\epsilon > 0$ is arbitrary, we have established the uniform tightness of $\{\mu_n; n \geq 0\}$.

Conversely, suppose that $\mu_n \overset{v}{\Longrightarrow} \mu_0$ and $\{\mu_n; n \geq 1\}$ is uniformly tight. Fix $F \in \mathcal{F}$, $\epsilon > 0$ and $K \in \mathcal{K}$ such that $\mu_n(K^c) < \epsilon$ for all $n \geq 1$. Since $\mu_n \overset{v}{\Longrightarrow} \mu_0$, by subadditivity and monotonicity of μ_n,

$$\limsup_{n\to\infty} \mu_n(F) \leq \epsilon + \limsup_{n\to\infty} \mu_n(F \cap K) \leq \epsilon + \mu_0(F \cap K) \leq \epsilon + \mu_0(F).$$

Taking $\epsilon \downarrow 0$, we conclude that $\mu_n \Longrightarrow \mu_0$. The tightness of μ_0 then follows by noting that $\mu_0(K^c) \leq \liminf \mu_n(K^c) \leq \epsilon$. ∎

Proof of Proposition 2.2. (i) Restrict attention to a subsequence along which $a_n \to 0$. Now suppose to the contrary that $\mu_0 = P^\alpha$, $\alpha \in (0,1]$ and that P is not a Dirac measure. We first claim that there exist $F_i \in \mathcal{F}$, $i = 1,2$ such that $P(F_1)P(F_2) > 0$ and $d^*(F_1, F_2) > 0$. Actually, if the claim is not true, it is not difficult to see that, for all $x \in \mathcal{X}$, $P(B(x, r_x)) = 0$ for some $r_x > 0$, where $B(x, r_x) = \{y \in \mathcal{X};\ d^*(y,x) < r_x\}$. Hence $P(K) = 0$ for every compact set K. Since this contradicts tightness, our claim is true. Since $P_n^{b_n} \Longrightarrow P$ as $b_n = a_n/\alpha \to 0$, we have for $\delta > 0$ sufficiently small

$$P(F_1) + P(F_2) \leq P(F_1^\delta \cup F_2^\delta) \leq \liminf_{n\to\infty} (P_n(F_1^\delta) + P_n(F_2^\delta))^{b_n} \leq \max\{P(\overline{F_1^\delta}), P(\overline{F_2^\delta})\}.$$

Taking $\delta \downarrow 0$, we arrive at the contradiction $P(F_1) + P(F_2) \leq \max\{P(F_1), P(F_2)\}$.
(ii) Since $I_n/\alpha \in \mathcal{I}_g(\mathcal{X})$ for all $\alpha > 0$, we easily adapt the proof of (i) above to the case at hand.
(iii) By restricting attention to a subsequence, we may suppose $a_n \to \alpha$. In this case,

clearly $P_n \Longrightarrow \hat{\mu}_0 = \mu_0^{1/\alpha}$. Without loss of generality we assume that $\hat{\mu}_0$ is not a Dirac measure. Thus there exist $F_i \in \mathcal{F}$, $i = 1, 2$ such that $\hat{\mu}_0(F_1)\hat{\mu}_0(F_2) > 0$ and $d^*(F_1, F_2) > 0$. We have

$$\begin{aligned}\hat{\mu}_0(F_1) + \hat{\mu}_0(F_2) &\le \lim_{\delta\downarrow 0}\liminf_{n\to\infty}(P_n(F_1^\delta) + P_n(F_2^\delta)) \\ &\le \lim_{\delta\downarrow 0}\hat{\mu}_0(\overline{F_1^\delta} \cup \overline{F_2^\delta}) = \hat{\mu}_0(F_1 \cup F_2) \le \hat{\mu}_0(F_1) + \hat{\mu}_0(F_2),\end{aligned}$$

so $\hat{\mu}_0(F_1 \cup F_2) = \hat{\mu}_0(F_1) + \hat{\mu}_0(F_2)$. Thus $\hat{\mu}_0$ must be in $\mathcal{P}(\mathcal{X})$.
(iv) This statement follows from the fact that all subsequences of a convergent sequence have the same limit. As noted below Definition 2.6, we cannot have $P^\alpha = e^{-I}$ for $P \in \mathcal{P}(\mathcal{X})$ and $I \in \mathcal{I}_g$, unless both are the same Dirac measure. With the same exception, we cannot have $P^\alpha = Q^\beta$ for $P \in \mathcal{P}(\mathcal{X})$, $Q \in \mathcal{P}(\mathcal{X})$, and $0 < \alpha < \beta \le 1$. To see this, note that if $0 < P(A) < 1$ we would need $Q(A) > P(A)$ and $Q(A^c) > P(A^c)$, which is impossible. ∎

Proof of Proposition 2.3. Since $\{\mu_n\}$ is Cauchy, the set $\{\mu_n : n \ge 1\}$ is totally bounded and hence uniformly tight (by Theorem 2.3). By Theorem 3.1(b) of O'Brien[(10)], $\{\mu_n\}$ has a convergent subsequence, whose limit is tight by Proposition 2.1. By the Cauchy property, the full sequence therefore has this tight limit. ∎

Proof of Lemma 2.2. Let $\mu = e^{-I}$ and let $\mathcal{U}$ be a countable base of the metric topology of $\mathcal{X}$. There exists a countable set $\mathcal{Y} = \{y_\ell; \ell = 1, 2, \ldots\} \subset \mathcal{X}$ such that $\mu(\mathcal{Y} \cap U) = \mu(U)$ for every $U \in \mathcal{U}$. Thus, for every $G \in \mathcal{G}$,

$$\mu(G) = \sup_{x\in G}\mu(\{x\}) = \sup_{U\in\mathcal{U}, U\subset G}\mu(U) = \sup_{U\in\mathcal{U}, U\subset G}\mu(\mathcal{Y}\cap U) = \mu(\mathcal{Y}\cap G)\ . \qquad (3.11)$$

In particular, $\mu(\mathcal{Y}) = \mu(\mathcal{X}) = 1$. Define $P_n \in \mathcal{P}(\mathcal{X})$ via $P_n(\{y_\ell\}) = c_n^{-1}2^{-\ell}\mu(\{y_\ell\})^{a_n^{-1}}$ for $y_\ell \in \mathcal{Y}$ and $c_n = \sum_\ell 2^{-\ell}\mu(\{y_\ell\})^{a_n^{-1}}$. Then, for every $A \in \mathcal{B}(\mathcal{X})$ and $y_j \in A$,

$$(2^{-j}\mu(\{y_j\})^{1/a_n})^{a_n} \le (\sum_{y_l\in A} 2^{-l}(\mu(\{y_l\}))^{1/a_n})^{a_n} \le \sup_{y_l\in A}\mu(\{y_l\}) = \mu(A\cap\mathcal{Y}).$$

Taking the limit as $n \to \infty$ of the left side and then taking the supremum over j such that $y_j \in A$, we see that

$$\lim_{n\to\infty}(\sum_{y_l\in A} 2^{-l}(\mu(\{y_l\}))^{1/a_n})^{a_n} = \mu(A\cap\mathcal{Y}).$$

Applying this to $\mathcal{X}$, we see that $c^{a_n} \to 1$, so

$$\lim_{n\to\infty} P_n^{a_n}(A) = \mu(\mathcal{Y} \cap A) \le \mu(A) \ . \tag{3.12}$$

Combining (3.11) and (3.12) we conclude that $P_n^{a_n} \Longrightarrow \mu$. ∎

Proof of Corollary 2.2. For $(\mathcal{X}, d^*)$ complete and separable, every $\mu \in \mathcal{Q}_1(\mathcal{X})$ is tight and, by Proposition 2.3, each Cauchy sequence in $(\mathcal{Q}_1(\mathcal{X}), d)$ has a subsequence for which one of the three cases of Proposition 2.2 applies. By Theorem 2.2, the metric space $(\mathcal{Q}_1(\mathcal{X}), d)$ is then complete. In this space, each $\mu = e^{-I}$, $I \in \mathcal{I}_g(\mathcal{X})$, is the limit of some sequence $\{P_n^{a_n}\}$ with $P_n \in \mathcal{P}(\mathcal{X})$ (by Lemma 2.2 and Theorem 2.2). Hence, $\mathcal{Q}_1(\mathcal{X})$ is the completion of $\{P^\alpha, P \in \mathcal{P}(\mathcal{X}), 0 < \alpha \le 1\}$.

Let $\mathcal{Y}$ be the countable collection of all Q^b for $b \in (0,1]$ rational and Q in a fixed countable dense subset of $(\mathcal{P}(\mathcal{X}), d)$. For $P \in \mathcal{P}(\mathcal{X})$ and $\alpha \in (0,1]$ there exists $P_n^{a_n} \in \mathcal{Y}$ such that $P_n \Longrightarrow P$ and $a_n \to \alpha$. With $\alpha > 0$, it follows that $P_n^{a_n} \Longrightarrow P^\alpha$ tight. Hence, $\mathcal{Y}$ is dense in $\{P^\alpha; P \in \mathcal{P}(\mathcal{X}), 0 < \alpha \le 1\}$ with respect to the metric d. By the triangle inequality, $\mathcal{Y}$ is thus dense in $(\mathcal{Q}_1(\mathcal{X}), d)$. ∎

Proof of Corollary 2.3. Fix $\alpha \in (0,1]$, $a > 0$ and $P, Q \in \mathcal{P}(\mathcal{X})$. Then, for $F \in \mathcal{F}$,

$$P(F)^\alpha \le Q(F^a)^\alpha + \rho_a(P,Q)^\alpha,$$

while for $G \in \mathcal{G}$,

$$Q(G^{-a})^\alpha = (1 - Q((G^c)^a))^\alpha \le (1 - P(G^c) + \rho_a(P,Q))^\alpha \le P(G)^\alpha + \rho_a(P,Q)^\alpha.$$

Consequently, we obtain that for every $a > 0$, $P, Q \in \mathcal{P}(\mathcal{X})$ and $\alpha \in (0,1]$

$$d(P^\alpha, Q^\alpha) \le \rho_a(P,Q)^\alpha \vee a. \tag{3.13}$$

Since (2.4) holds, by (3.13) also $\lim_k \limsup_n d(P_{n,k}^{a_n}, Q_n^{a_n}) = 0$. By the assumed full LDP for $\{P_{n,k}\}_n$ and Theorem 2.2, $d(P_{n,k}^{a_n}, e^{-I_k}) \to 0$ for each k fixed. Applying the triangle inequality for d then yields $\lim_k \limsup_n d(Q_n^{a_n}, e^{-I_k}) = 0$. In particular, $\{e^{-I_k}\}_k$ is a Cauchy sequence in $(\mathcal{Q}_1(\mathcal{X}), d)$. Note that each e^{-I_k} is tight. By Proposition 2.3 and part (ii) of Proposition 2.2, $\lim_k d(e^{-I_k}, e^{-I}) = 0$ for some $I \in \mathcal{I}_g(\mathcal{X})$. Applying the triangle inequality once more, we conclude that $d(Q_n^{a_n}, e^{-I}) \to 0$ as required. ∎

The following three lemmas are part of the proof of Theorem 2.4.

Lemma 3.1 *In cases (i)-(iii) of Proposition 2.2,* $\widetilde{\int} f d\mu_n \to \widetilde{\int} f d\mu_0$ *for every* $f \in C_b(\mathcal{X})$.

Proof of Lemma 3.1. Without loss of generality fix $f \geq 0$, $f \in C_b(\mathcal{X})$.
(i) and (ii). From Definition 2.6, for $A_i \in \mathcal{B}(\mathcal{X})$ such that $\mathcal{X} = \cup_{i=1}^M A_i$,

$$\inf_{x \in A_1} f(x)\mu_n(A_1) \leq \widetilde{\int} f d\mu_n \leq M^{a_n} \max_{i=1}^{M}\{\sup_{x \in A_i} f(x)\mu_n(A_i)\} \tag{3.14}$$

(taking $a_n := 0$ in case (ii)). Fix $\epsilon > 0$, and let $\{F_i; i = 1, \ldots, M = M(\epsilon)\} \subset \mathcal{F}$ be a finite cover of $\mathcal{X}$ such that $\sup_{y \in F_i} |f(y) - f(x_i)| < \epsilon$ for some $x_i \in F_i$. From (3.14), since $a_n \to 0$ and $\mu_n \Longrightarrow \mu_0 = e^{-I}$,

$$\limsup_{n \to \infty} \widetilde{\int} f d\mu_n \leq \max_{i=1}^{M}\{f(x_i)\mu_0(F_i)\} + \epsilon \leq \widetilde{\int} f d\mu_0 + 2\epsilon \; . \tag{3.15}$$

For any $x \in \mathcal{X}$, let $\delta = \delta(x, \epsilon) > 0$ be such that $\sup_{y \in \{x\}^\delta} |f(y) - f(x)| < \epsilon$. Then, from (3.14),

$$\liminf_{n \to \infty} \widetilde{\int} f d\mu_n \geq f(x)\mu_0(\{x\}^\delta) - \epsilon \geq f(x)\mu_0(\{x\}) - \epsilon, \tag{3.16}$$

which together with (3.15) implies the stated conclusion since $x \in \mathcal{X}$ and $\epsilon > 0$ are arbitrary.
(iii) Fix $\epsilon > 0$. Since $\{c; P(f = c) > 0\}$ is at most countable, there exist $\{c_k \geq 0; k = 1, \ldots, M = M(\epsilon)\}$ and P-continuity sets $\{A_k; k = 1, \ldots, M\}$ such that $\|f - f_\epsilon\|_\infty < \epsilon$ for $f_\epsilon = \sum_{i=1}^M c_k 1_{A_k}$. By the triangle inequality for subadditive-integrals,

$$|\widetilde{\int} f d\mu_n - \widetilde{\int} f d\mu_0| \leq 2\epsilon + |\widetilde{\int} f_\epsilon d\mu_n - \widetilde{\int} f_\epsilon d\mu_0|. \tag{3.17}$$

Since $\mu_0(A_k^o) = \mu_0(\overline{A}_k)$ and $a_n \to \alpha > 0$, it follows that

$$\lim_{n \to \infty} \widetilde{\int} f_\epsilon d\mu_n = \lim_{n \to \infty} (\sum_{k=1}^M (c_k \mu_n(A_k))^{1/a_n})^{a_n} = (\sum_{k=1}^M (c_k \mu_0(A_k))^{1/\alpha})^\alpha = \widetilde{\int} f_\epsilon d\mu_0.$$

As $\epsilon > 0$ is arbitrary, (3.17) implies the stated conclusion. ∎

Lemma 3.2 *For every* $c > 0$ *there exists* $r(c) < \infty$ *such that*

$$d(\mu, \nu) \leq r(c)[d(\mu^\alpha, \nu^\gamma) + |\gamma - \alpha|], \tag{3.18}$$

for every $\alpha, \gamma \in [c, 1]$ *and every* $\mu, \nu \in \mathcal{Q}(\mathcal{X})$ *for which* $\mu(\mathcal{X}) \leq 1$, $\nu(\mathcal{X}) \leq 1$.

Proof of Lemma 3.2. Noting that $\mu^\alpha, \nu^\gamma \in \mathcal{Q}(\mathcal{X})$, we fix $\delta > d(\mu^\alpha, \nu^\gamma)$ such that $\delta \leq 2$. One can check that for some $r(c) \in [1, \infty)$

$$(A^\gamma + \delta)^{1/\alpha} \leq A + r(c)[\delta + |\gamma - \alpha|]$$

for all $A, \delta \in [0,2]$ and $\alpha, \gamma \in [c, 1]$. One way is to prove the statement with $\delta = 0$ and then note that for the variables in the indicated ranges, the derivative with respect to δ has an upper bound depending only on c. In particular, for every $F \in \mathcal{F}$,

$$\mu(F) \leq (\nu(F^\delta)^\gamma + \delta)^{1/\alpha} \leq \nu(F^\delta) + r(c)[\delta + |\gamma - \alpha|] ,$$

and

$$\nu(F) \leq (\mu(F^\delta)^\alpha + \delta)^{1/\gamma} \leq \mu(F^\delta) + r(c)[\delta + |\gamma - \alpha|] .$$

Hence, $d(\mu, \nu) \leq r(c)[\delta + |\gamma - \alpha|]$. Taking $\delta \downarrow d(\mu^\alpha, \nu^\gamma)$ completes the proof. ∎

Lemma 3.3 *For every $P, Q \in \mathcal{P}(\mathcal{X})$ and $\alpha \in (0, 1]$,*

$$\beta(P^\alpha, Q^\alpha) \leq (2\alpha^{-1} d(P, Q))^\alpha . \tag{3.19}$$

Proof of Lemma 3.3. Fix $\epsilon > d(P, Q)$ and $f \in B$. Since $||f(x)|^p - |f(y)|^p| \leq pd^*(x, y)$ for all $x, y \in \mathcal{X}$ and $p \geq 1$, it follows that

$$\{x; |f(x)|^{1/\alpha} \geq t\}^\epsilon \subset \{x; |f(x)|^{1/\alpha} > t - \epsilon/\alpha\} \quad \forall t .$$

Hence

$$\begin{aligned} \widetilde{\int} f dP^\alpha &\leq (\int_0^1 P(|f|^{1/\alpha} \geq t) dt)^\alpha \leq (\int_0^1 Q(|f|^{1/\alpha} > t - \epsilon/\alpha) dt + \epsilon)^\alpha \\ &\leq (\int_0^1 Q(|f|^{1/\alpha} > t) dt)^\alpha + (2\alpha^{-1}\epsilon)^\alpha = \widetilde{\int} f dQ^\alpha + (2\alpha^{-1}\epsilon)^\alpha . \end{aligned}$$

Exchanging the roles of P and Q, taking the supremum over $f \in B$ followed by $\epsilon \downarrow d(P, Q)$ we obtain (3.19). ∎

Proof of Theorem 2.4. We will prove that (i) and (iii) are equivalent, and then that (ii) and (iv) are equivalent. Fix $\{\mu_n, n \geq 0\} \subset \mathcal{Q}_1(\mathcal{X})$.

First assume (i) and suppose that

$$\liminf_{k\to\infty} |\widetilde{\int} f d\mu_{n_k} - \widetilde{\int} f d\mu_0| > 0, \tag{3.20}$$

for some $f \in C_b(\mathcal{X})$ and a subsequence n_k. By definition of $\mathcal{Q}_1(\mathcal{X})$, passing to a subsequence of n_k, one of the cases (i)–(iii) of Proposition 2.2 applies, hence (3.20) contradicts Lemma 3.1. In conclusion, $\mu_n \Longrightarrow \mu_0$ implies that $\widetilde{\int} f d\mu_n \to \widetilde{\int} f d\mu_0$ for every $f \in C_b(\mathcal{X})$.

Now assume (iii). Fix $F \in \mathcal{F}$, $G \in \mathcal{G}$ and $\epsilon > 0$. Set $f_\epsilon(x) = 0 \vee (1 - d^*(x, F)/\epsilon)$. Since $1_F \le f_\epsilon \le 1_{F^\epsilon}$ and $f_\epsilon \in C_b(\mathcal{X})$, by the properties of the subadditive-integral,

$$\limsup_{n\to\infty} \mu_n(F) \le \limsup_{n\to\infty} \widetilde{\int} f_\epsilon d\mu_n = \widetilde{\int} f_\epsilon d\mu_0 \le \mu_0(F^\epsilon).$$

Similarly, since $1_{G^{-\epsilon}} \le g_\epsilon \le 1_G$ for $g_\epsilon(x) = 1 \wedge (d^*(x, G^c)/\epsilon) \in C_b(\mathcal{X})$, also

$$\liminf_{n\to\infty} \mu_n(G) \ge \mu_0(G^{-\epsilon}).$$

With $\mu_0 \in \mathcal{Q}_1(\mathcal{X})$ we have that $\mu_0(F^\epsilon) \to \mu_0(F)$ for $\epsilon \downarrow 0$ and every $F \in \mathcal{F}$. Taking $\epsilon \downarrow 0$, by the arbitrariness of $F \in \mathcal{F}$ and $G \in \mathcal{G}$ we conclude that $\mu_n \Longrightarrow \mu_0$.

Now assume (iv). Fix $F \in \mathcal{F}$, $\epsilon > 0$ and f_ϵ as above. Since $1_F \le f_\epsilon \le 1_{F^\epsilon}$ and $\|f_\epsilon\|_{\mathrm{BL}} \le 1 + \epsilon^{-1}$ it follows that for every $\mu, \nu \in \mathcal{Q}_1(\mathcal{X})$,

$$\mu(F) \le \widetilde{\int} f_\epsilon d\mu \le \widetilde{\int} f_\epsilon d\nu + (1+\epsilon^{-1})\beta(\mu,\nu) \le \nu(F^\epsilon) + (1+\epsilon^{-1})\beta(\mu,\nu). \qquad (3.21)$$

Clearly, (3.21) holds with μ and ν exchanging positions. By (2.3) and the arbitrariness of $F \in \mathcal{F}$ in (3.21) it follows that $d(\mu,\nu) \le \epsilon \vee (1+\epsilon^{-1})\beta(\mu,\nu)$. Taking $\epsilon \downarrow \beta(\mu,\nu)^{1/2} \le 1$, we conclude that

$$d(\mu,\nu) \le 2\beta(\mu,\nu)^{1/2} \quad \forall \mu,\nu \in \mathcal{Q}_1(\mathcal{X}) \, . \qquad (3.22)$$

In particular, this implies $d(\mu_n, \mu_0) \to 0$.

Assume now that (ii) holds and consider the case where μ_0 is tight. Fix $\epsilon > 0$ and $K \in \mathcal{K}$ such that $\mu_0(K^c) < \epsilon$. By the Arzelá-Ascoli theorem, the set of functions B restricted to K is a compact subset of $(C(K), \|\cdot\|_\infty)$. In particular, there exist $M = M(\epsilon) < \infty$ and $\{f_i; i = 1, \dots, M\} \subset B$ such that

$$\min_{i=1}^{M} \sup_{x\in K} |f(x) - f_i(x)| < \epsilon \quad \forall f \in B.$$

If $x \in K$ and $d^*(x, y) < \epsilon$, then $|f(y) - g(y)| \le 2\epsilon + |f(x) - g(x)|$ for every $f, g \in B$. Therefore,

$$\min_{i=1}^{M} \sup_{y\in K^\epsilon} |f(y) - f_i(y)| < 3\epsilon \quad \forall f \in B.$$

Hence, by the properties of the subadditive-integral, for every $f \in B$, $\mu \in \mathcal{Q}_1(\mathcal{X})$,

$$\begin{aligned}
|\widetilde{\int} f d\mu - \widetilde{\int}_{K^\epsilon} f d\mu| &\leq \mu((K^\epsilon)^c) \\
\min_{i=1}^{M} |\widetilde{\int}_{K^\epsilon} f d\mu - \widetilde{\int}_{K^\epsilon} f_i d\mu| &< 3\epsilon\mu(K^\epsilon) \leq 3\epsilon.
\end{aligned}$$

In particular, for all $f \in B$, $n \geq 1$,

$$|\widetilde{\int} f d\mu_n - \widetilde{\int} f d\mu_0| \leq 2\mu_n((K^\epsilon)^c) + 2\mu_0((K^\epsilon)^c) + 6\epsilon + \max_{i=1}^{M} |\widetilde{\int} f_i d\mu_n - \widetilde{\int} f_i d\mu_0| \quad (3.23)$$

Since $\mu_n \Longrightarrow \mu_0$ and $(K^\epsilon)^c \in \mathcal{F}$, for all n large enough $\mu_n((K^\epsilon)^c) \leq \epsilon$. By part (i) above, $\max_{i=1}^{M} |\tilde{\int} f_i d\mu_n - \tilde{\int} f_i d\mu_0| \to 0$, hence from (3.23)

$$\limsup_{n\to\infty} \beta(\mu_n, \mu_0) = \limsup_{n\to\infty} \sup_{f\in B} |\widetilde{\int} f d\mu_n - \widetilde{\int} f d\mu_0| \leq 10\epsilon.$$

Taking $\epsilon \downarrow 0$ completes the proof in case μ_0 is tight.

If $\mu_0 \in \mathcal{Q}_1(\mathcal{X})$ is non-tight, then necessarily $\mu_0 = P^\alpha$ for some $P \in \mathcal{P}(\mathcal{X})$ non-tight, $\alpha > 0$. Suppose that $d(\mu_n, \mu_0) \to 0$ but $\limsup_n \beta(\mu_n, \mu_0) > \eta > 0$. Since $d(\mu_n, \mu_0) \to 0$ implies $\mu_n \Longrightarrow \mu_0$, passing to a subsequence, by Proposition 2.2 we may assume that $\mu_n = P_n^{a_n}$ with $P_n \in \mathcal{P}(\mathcal{X})$, $a_n \to \alpha$ while $\beta(\mu_n, \mu_0) \geq \eta$. Applying Lemma 3.2 followed by Lemma 3.3, we deduce that $\beta(\mu_n, P^{a_n}) \to 0$. Since also $\beta(P^{a_n}, P^\alpha) \to 0$, we see that $\beta(\mu_n, \mu_0) \to 0$. ∎

Proof of Corollary 2.4. Define the pseudometric $d^*(x, y) = \sup_{f\in\widetilde{B}} |f(x) - f(y)|$ on $\mathcal{X}$. Replacing $\mathcal{X}$ if needed by the space of d^*-equivalence classes, we may assume without loss of generality that d^* is a metric (for details see Dudley[(6)]). Since $\widetilde{B}$ is equicontinuous, d^* is jointly continuous on $(\mathcal{X} \times \mathcal{X}, \tau \times \tau)$. Hence, the metric topology of $(\mathcal{X}, d^*)$ is coarser than τ. Our assumptions that $\mu_n \Longrightarrow \mu_0$ and μ_0 is tight in $(\mathcal{X}, \tau)$, are thus also valid in the metric space $(\mathcal{X}, d^*)$. Defining Lipschitz norms with respect to d^* let $A := \sup_{f\in\widetilde{B}} \|f\|_{\mathrm{BL}}$. We obtain (2.5) since $A \leq 1 + \sup_{f\in\widetilde{B}} \|f\|_\infty < \infty$ and $\beta(\mu_n, \mu_0) \to 0$ (see Theorems 2.2 and 2.4). ∎

Acknowledgment. Tiefeng Jiang would like to thank his adviser Amir Dembo for many inspiring discussions.

References

[1.] Baxter, J. R. and Jain, N. C. A comparison principle for large deviations, Proc. Amer. Math. Soc. 103:925-960, 1988.

[2.] Baxter, J. R. and Jain, N. C. An approximation condition for large deviations and an application to stationary sequences. Preprint, 1993.

[3.] Billingsley P., *Convergence of probability measures*, John Wiley and Sons, NY, 1968.

[4.] Dembo, A. and Zeitouni, O. *Large Deviations Techniques and Applications*, Springer-Verlag 1998.

[5.] Deuschel J. D. and Stroock D. W. *Large Deviations.* Academic Press, Boston, 1989.

[6.] Dudley R. M., *Real Analysis and Probability*, Wadsworth and Brooks/Cole, Pacific Grove, CA, 1989.

[7.] Dupuis P. and Ellis R. S. *A Weak Convergence Approach to the Theory of Large Deviations.* Wiley, New York, 1996.

[8.] Freidlin M. I. and Wentzell A. D. *Random Perturbations of Dynamical Systems.* Springer-Verlag, New York, 1984.

[9.] J. Lynch and J. Sethuraman. Large deviations for processes with independent increments. *Ann. Probab.*, 15:610–627, 1987.

[10.] O'Brien G. L. Sequences of capacities, with connections to large-deviation theory, Journal of Theoretical Probability, Vol.9, No.1, page 19-35, 1996.

[11.] O'Brien G. L. and Vervaat, W. Capacities, large deviations and loglog laws. In S. Cambanis, G. Samorodnitsky and M. Taqqu, editors, *Stable Processes and Related Topics*, pages 43–84. Birkhäuser, 1991.

[12.] Pukhalskii A. A. On functional principle of large deviations. In V. Sazonov and T. Shervashidze, editors, *New Trends in Probability and Statistics*, pages 198–218. VSP Moks'las, Moskva, 1991.

[13.] Pukhalskii, A. A. The method of stochastic exponentials for large deviations, Stoch. Proc. Appl. 54:45–70, 1994.

[14.] Varadhan S. R. S. Asymptotic probabilities and differential equations. *Comm. Pure Appl. Math.*, 19:261–286, 1966.

[15.] Varadhan S. R. S. *Large Deviations and Applications.* SIAM, Philadelphia, 1984.