



# A linear version of Dawson-Gärtner theorem – Applications to Cramér's theory

P Petit

## ► To cite this version:

P Petit. A linear version of Dawson-Gärtner theorem – Applications to Cramér's theory. 2019.  
hal-02279591

**HAL Id: hal-02279591**

**<https://hal.science/hal-02279591>**

Preprint submitted on 5 Sep 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A LINEAR VERSION OF DAWSON-GÄRTNER THEOREM APPLICATIONS TO CRAMÉR'S THEORY

BY PIERRE PETIT

*Institut de Mathématiques de Toulouse*

September 5, 2019

## Abstract

We prove a linear version of Dawson-Gärtner theorem: weak large deviation principles and the equality  $-s = p^*$  between the negentropy and the Fenchel-Legendre transform of the pressure are preserved through linear projective limits. As a result, the equality  $-s = p^*$  holds in great generality for empirical means of independent and identically distributed random variables (Cramér's theory), e.g. in any measurable normed space, and even in any projective limit of such spaces. Eventually, we give an original example where  $-s \neq p^*$  and discuss the dual equality.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Setting and main results</b>	<b>2</b>
2.1	Entropy and pressure . . . . .	2
2.2	Inverse image of a locally measurable topological vector space . . . . .	4
2.3	Projective limits of locally measurable topological vector spaces . . . . .	4
<b>3</b>	<b>Proofs of Theorems 1 and 2</b>	<b>5</b>
<b>4</b>	<b>Application to Cramér's theory</b>	<b>7</b>
4.1	Cramér's theorem in separable normed vector spaces . . . . .	8
4.2	Cramér's theorem in locally convex separable spaces . . . . .	8
4.3	Cramér's theorem in locally convex measurable spaces . . . . .	10
<b>5</b>	<b>A counterexample to <math>s = -p^*</math></b>	<b>12</b>
<b>6</b>	<b>About the equality <math>p = (-s)^*</math></b>	<b>13</b>

---

MSC2010 subject classifications: 60F10; 28E15.

Key words and phrases: Dawson-Gärtner theorem, projective limits, Cramér's theory, large deviations, Fenchel-Legendre transformation, measurable vector spaces, measurable cardinals.

E-mail: pierre.petit@normalesup.org

Address: UPS IMT, F-31062 Toulouse Cedex 9, France

# 1 Introduction

In 1987, Dawson and Gärtner (see [DG87]) formalized the fact that large deviation principles with good rate functions well behave with respect to projective limits. Originally designed to establish large deviation principles in weak topologies, especially in Sanov-type theorems, the result allows to extend large deviation principles to rather general contexts (see, *e.g.*, [DS89], [dA94a], [dA94b], [dA97], [DZ98]).

In the spirit of Sanov's theorem (see [San57]) and Varadhan's theory of large deviations (see [Var66]), via Freidlin and Wentzell's theory (see [FW84]), Dawson-Gärtner theorem relies on a compactness assumption: the goodness of the rate function. In order to handle generalizations of Cramér's theorem (see, *e.g.*, [Cra38], [BZ79], [DS89], [DZ98], [Cer07]), we weaken the role of compactness and exploit the role of linearity when dealing with empirical means of random variables. We prove that weak large deviation principles and the equality  $-s = p^*$  between the negentropy and the Fenchel-Legendre transform of the pressure well behave with respect to linear projective limits (see Theorem 2). We also prove a linear version of the inverse contraction principle (see Theorem 1).

In Section 4, we apply the result to Cramér's theory. In Section 4.3, we obtain a general framework in which the entropy  $s$  and the pressure  $p$  of a sequence of empirical means of independent and identically distributed random variables are well-defined. It can be seen as a constructive general model of Cramér's theory, closely related to the nonconstructive frameworks of [BZ79] and [Cer07], and generalizing both the setting of separable Banach spaces with their Borel  $\sigma$ -algebra and the setting of weak topologies with cylinder  $\sigma$ -algebra. It appears that Theorem 1 and Theorem 2 imply the equality  $-s = p^*$  in this general framework (see Theorem 7). In particular, the equality  $-s = p^*$  is true in any separable (not necessarily complete) normed vector space with its Borel  $\sigma$ -algebra (see Theorem 4), and even in any normed vector space with its Borel  $\sigma$ -algebra as soon as the question makes sense (see Theorem 6).

This analysis was motivated by a question of Raphaël Cerf: does there exist a sequence of empirical means of independent and identically distributed random variables for which  $-s \neq p^*$ ? We provide such an example, on the borders of Cramér's theory, in Section 5. And we discuss the dual equality  $p = (-s)^*$  in Section 6.

## 2 Setting and main results

### 2.1 Entropy and pressure

As in [LPS95], we want that the entropy of any sequence of random variables be well-defined. We add an extra hypothesis so that the same holds for the pressure. Let  $(\mathcal{X}, \tau)$  be a (not necessarily Hausdorff) topological vector space and let  $\mathcal{F}$  be a (not necessarily Borel)  $\sigma$ -algebra over  $\mathcal{X}$  such that:

(LM1) every point  $x$  in  $\mathcal{X}$  admits a local basis  $\mathcal{V}_x$  made up of measurable neighborhoods;

(LM2) every continuous linear functional over  $\mathcal{X}$  is a measurable map.

We will say that  $(\mathcal{X}, \tau, \mathcal{F})$  is a *locally measurable vector space* (*l.m.v.s.*). Throughout the text,  $\mathcal{X}^*$  denotes the topological dual of  $\mathcal{X}$ , i.e. the vector space of continuous linear functional over  $\mathcal{X}$ . Let  $(Z_n)_{n \geq 1}$  be a sequence of  $(\mathcal{X}, \mathcal{F})$ -valued random variables and  $(v_n)_{n \geq 1}$  a sequence of

positive numbers diverging to  $+\infty$ . Let us define three important functions in the theory of large deviations (further details can be found in [BZ79], [LPS95], and [LP95]). The *entropy* of the sequence  $(Z_n, v_n)_{n \geq 1}$  is the function  $s : \mathcal{X} \rightarrow [-\infty, 0]$  defined by

$$\forall x \in \mathcal{X} \quad s(x) := \inf_{V \in \mathcal{V}_x} \liminf_{n \rightarrow \infty} \frac{1}{v_n} \log \mathbb{P}(Z_n \in V),$$

the *entropy superior* of the sequence  $(Z_n, v_n)_{n \geq 1}$  is the function  $\bar{s} : \mathcal{X} \rightarrow [-\infty, 0]$  defined by

$$\forall x \in \mathcal{X} \quad \bar{s}(x) := \inf_{V \in \mathcal{V}_x} \limsup_{n \rightarrow \infty} \frac{1}{v_n} \log \mathbb{P}(Z_n \in V),$$

and the *pressure* of the sequence  $(Z_n, v_n)_{n \geq 1}$  is the function  $p : \mathcal{X}^* \rightarrow [-\infty, +\infty]$  defined by

$$\forall \lambda \in \mathcal{X}^* \quad p(\lambda) := \limsup_{n \rightarrow \infty} \frac{1}{v_n} \log \mathbb{E}(e^{v_n \lambda(Z_n)}).$$

Note that the functions  $s$  and  $\bar{s}$  do not depend on the local bases  $\mathcal{V}_x$  chosen in each point  $x \in \mathcal{X}$  (hypothesis (LM1)) and that  $p$  is well defined (hypothesis (LM2)). Remember that the *Fenchel-Legendre transform* of a function  $h : \mathcal{X}^* \rightarrow [-\infty, +\infty]$  is the function  $h^* : \mathcal{X} \rightarrow [-\infty, +\infty]$

$$\forall x \in \mathcal{X} \quad h^*(x) = \sup_{\lambda \in \mathcal{X}^*} (\lambda(x) - h(\lambda)).$$

(see [Mor67] for a detailed study of the notion). It is always true that

$$s \leq \bar{s} \leq -p^*, \tag{1}$$

the second inequality being a consequence of Chebyshev's inequality. It appears that equalities in (1) are of greatest interest in large deviation theory. Indeed, by definition, the entropy  $s$  is the greatest function that satisfies the lower bound:

$$(\text{LB}) \text{ for all subset } G \in \mathcal{F}, \liminf_{n \rightarrow \infty} \frac{1}{v_n} \log \mathbb{P}(Z_n \in G) \geq \sup_{x \in G} s(x).$$

And the entropy superior  $\bar{s}$  satisfies the compact upper bound:

$$(\text{UB}_k) \text{ for all relatively compact subset } K \in \mathcal{F}, \limsup_{n \rightarrow \infty} \frac{1}{v_n} \log \mathbb{P}(Z_n \in K) \leq \sup_{x \in \bar{K}} \bar{s}(x).$$

Recall that a set is compact if each of its open covers has a finite subcover (no separation axiom is required). And we say that a subset is relatively compact if its closure is compact. The proof of  $(\text{UB}_k)$  can be adapted from the one of [LP95, Lemma 2.5] as follows: if  $\sup \{\bar{s}(x) ; x \in \bar{K}\} < M$ , then, for all  $x \in \bar{K}$ , there exists  $V_x \in \mathcal{V}_x$  such that  $\limsup_{n \rightarrow \infty} \frac{1}{v_n} \log \mathbb{P}(Z_n \in V_x) < M$ ; extracting a finite subcover of  $\bar{K}$  by the  $V_x$  and using the principle of the largest term ([LP95, Lemma 2.3]), one gets the bound. Therefore, a natural sufficient condition so that  $(Z_n, v_n)_{n \geq 1}$  satisfies a weak large deviation principle (i.e.  $(\text{LB}) + (\text{UB}_k)$ ) is that  $s = \bar{s}$ . Furthermore, the function  $p^*$  is often easier to compute than  $s$ , which makes relevant the study of the equality  $s = -p^*$ . The purpose of the text is to show that both l.m.v.s. and equalities in (1) well behave with respect to inverse images and projective limits, in the spirit of Dawson-Gärtner theorem (see, e.g., [DZ98, Theorem 4.6.1]). Then we apply the result to the case of empirical means of independent and identically distributed random variables (Cramér's theory) and obtain a general weak large deviation principle with entropy  $s = -p^*$ .

## 2.2 Inverse image of a locally measurable topological vector space

Let  $\mathcal{X}$  be a vector space, let  $(\mathcal{X}_1, \tau_1, \mathcal{F}_1)$  be a l.m.v.s., and let  $f_1: \mathcal{X} \rightarrow \mathcal{X}_1$  be a linear map. Let  $\tau$  (resp.  $\mathcal{F}$ ) be the initial topology (resp. the initial  $\sigma$ -algebra) on  $\mathcal{X}$  with respect to the linear map  $f_1$ , also called *inverse image topology under  $f_1$*  (resp.  *$\sigma$ -algebra generated by  $f_1$* ). It appears to be that  $(\mathcal{X}, \tau, \mathcal{F})$  is a l.m.v.s. (see Propositions 1 and 2 below). We will say that  $(\mathcal{X}, \tau, \mathcal{F})$  is the *inverse image of the l.m.v.s.  $(\mathcal{X}_1, \tau_1, \mathcal{F}_1)$  under  $f_1$* .

**Theorem 1** (Linear inverse contraction principle). *Let  $\mathcal{X}$  be a vector space, let  $(\mathcal{X}_1, \tau_1, \mathcal{F}_1)$  be a l.m.v.s., and let  $f_1: \mathcal{X} \rightarrow \mathcal{X}_1$  be a linear map. Let  $(\mathcal{X}, \tau, \mathcal{F})$  be the inverse image of the l.m.v.s.  $(\mathcal{X}_1, \tau_1, \mathcal{F}_1)$  under  $f_1$ . Let  $(Z_n)_{n \geq 1}$  be a sequence of  $(\mathcal{X}, \mathcal{F})$ -valued random variables and  $(v_n)_{n \geq 1}$  a sequence of positive numbers diverging to  $+\infty$ . Define  $s$ ,  $\bar{s}$ , and  $p$  (resp.  $s_1$ ,  $\bar{s}_1$ , and  $p_1$ ) the entropy, the entropy superior, and the pressure of the sequence  $(Z_n, v_n)_{n \geq 1}$  (resp. of the sequence  $(f_1(Z_n), v_n)_{n \geq 1}$ ). Then,*

$$s = s_1 \circ f_1, \quad \bar{s} = \bar{s}_1 \circ f_1, \quad \text{and} \quad p^* = p_1^* \circ f_1. \quad (2)$$

1. *If  $s_1 = \bar{s}_1$ , then  $s = \bar{s}$ ; in particular,  $(Z_n, v_n)_{n \geq 1}$  satisfies a weak large deviation principle.*
2. *If  $\bar{s}_1 = -p_1^*$ , then  $\bar{s} = -p^*$ .*
3. *If  $(f_1(Z_n), v_n)_{n \geq 1}$  satisfies a weak large deviation principle and each compact subset of  $\mathcal{X}_1$  is measurable, then  $(Z_n, v_n)_{n \geq 1}$  satisfies a weak large deviation principle.*

## 2.3 Projective limits of locally measurable topological vector spaces

Let  $(I, \leq)$  be a directed preordered set. For all  $i \in I$ , let  $(\mathcal{X}_i, \tau_i, \mathcal{F}_i)$  be a l.m.v.s. and, for all  $(i, j) \in I^2$  with  $i \leq j$ , let  $f_{i,j}$  be a map from  $\mathcal{X}_j$  to  $\mathcal{X}_i$ . Suppose that

(PL1) for all  $i \in I$ ,  $f_{i,i} = \text{id}_{\mathcal{X}_i}$ ;

(PL2) for all  $(i, j, k) \in I^3$  with  $i \leq j \leq k$ ,  $f_{i,k} = f_{i,j} \circ f_{j,k}$ ;

(PL3) for  $(i, j) \in I^2$  with  $i \leq j$ ,  $f_{i,j}$  is a measurable, continuous, and linear map.

In particular,  $(\mathcal{X}_i, f_{i,j})_{i \leq j}$  is a projective system of sets (see, e.g., [Bou66, III.52]). Let  $\mathcal{X}$  be its projective limit and, for all  $i \in I$ , let  $f_i$  be the canonical projection from  $\mathcal{X}$  to  $\mathcal{X}_i$ . Let  $\tau$  (resp.  $\mathcal{F}$ ) be the initial topology (resp. the initial  $\sigma$ -algebra) on  $\mathcal{X}$  with respect to the family of maps  $(f_i)_{i \in I}$ . It appears to be that  $(\mathcal{X}, \tau, \mathcal{F})$  is also a l.m.v.s. (see Propositions 1 and 2 below). We will say that  $(\mathcal{X}, \tau, \mathcal{F})$  is the *projective limit of the projective system of l.m.v.s.  $(\mathcal{X}_i, \tau_i, \mathcal{F}_i, f_{i,j})_{i \leq j}$* .

**Theorem 2** (Linear Dawson-Gärtner). *Let  $(\mathcal{X}_i, \tau_i, \mathcal{F}_i, f_{i,j})_{i \leq j}$  be a projective system of l.m.v.s. and  $(\mathcal{X}, \tau, \mathcal{F})$  its projective limit. For all  $i \in I$ , let  $f_i$  be the canonical projection from  $\mathcal{X}$  to  $\mathcal{X}_i$ . Let  $(Z_n)_{n \geq 1}$  be a sequence of  $(\mathcal{X}, \mathcal{F})$ -valued random variables and  $(v_n)_{n \geq 1}$  a sequence of positive numbers diverging to  $+\infty$ . Define  $s$ ,  $\bar{s}$ , and  $p$  (resp.  $s_i$ ,  $\bar{s}_i$ , and  $p_i$ ) the entropy, the entropy superior, and the pressure of the sequence  $(Z_n, v_n)_{n \geq 1}$  (resp. of the sequence  $(f_i(Z_n), v_n)_{n \geq 1}$ ). Then,*

$$s = \inf_{i \in I} s_i \circ f_i, \quad \bar{s} = \inf_{i \in I} \bar{s}_i \circ f_i, \quad \text{and} \quad p^* = \sup_{i \in I} p_i^* \circ f_i. \quad (3)$$

1. If, for all  $i \in I$ ,  $s_i = \bar{s}_i$ , then  $s = \bar{s}$ ; in particular,  $(Z_n, v_n)_{n \geq 1}$  satisfies a weak large deviation principle.
2. If, for all  $i \in I$ ,  $\bar{s}_i = -p_i^*$ , then  $\bar{s} = -p^*$ .
3. If, for all  $i \in I$ ,  $(f_i(Z_n), v_n)_{n \geq 1}$  satisfies a weak large deviation principle and each compact subset of  $\mathcal{X}_i$  is measurable, then  $(Z_n, v_n)_{n \geq 1}$  satisfies a weak large deviation principle.

### 3 Proofs of Theorems 1 and 2

Let us kill two birds with one stone. Let  $(\mathcal{X}, \tau, \mathcal{F})$  be either an inverse image structure or a projective limit structure. The structure  $(\mathcal{X}, \tau, \mathcal{F})$  is defined as an initial structure with respect to a family of linear maps  $f_i: \mathcal{X} \rightarrow \mathcal{X}_i$  ( $i \in I$ ) such that  $(I, \leq)$  is a directed preordered set ( $I = \{1\}$  in the case of the inverse image). Moreover, we have the following key result, where, for all  $i \in I$  and all point  $x_i \in \mathcal{X}_i$ ,  $\mathcal{V}_{i, x_i}$  denotes a local basis of  $x_i$  made up of measurable neighborhoods.

**Proposition 1.** *For all  $x \in \mathcal{X}$ , the set*

$$\mathcal{V}_x := \{f_i^{-1}(V_i); i \in I, V_i \in \mathcal{V}_{i, f_i(x)}\}$$

*is a local basis of neighborhoods of  $x$ . In particular,  $(\mathcal{X}, \tau, \mathcal{F})$  satisfies (LM1).*

*Proof.* In the case of an inverse image structure, it stems from the definition of the inverse image topology. And, as for a projective limit structure, it is proved in [Bou71, I.29].  $\square$

**Proposition 2.** *The topological dual of  $\mathcal{X}$  is*

$$\mathcal{X}^* = \{\lambda_i \circ f_i; i \in I, \lambda_i \in \mathcal{X}_i^*\}.$$

*In particular,  $(\mathcal{X}, \tau, \mathcal{F})$  satisfies (LM2).*

*Proof.* Let  $\lambda \in \mathcal{X}^*$ . Since  $\lambda$  is continuous, there is a set  $V \in \mathcal{V}_0$  such that  $V \subset \{\lambda < 1\}$ . According to Proposition 1, there exists  $i \in I$  and  $V_i \in \mathcal{V}_{i, 0}$  such that  $V = f_i^{-1}(V_i)$ . Then, for all  $x \in \mathcal{X}$ , the equality  $f_i(x) = 0$  entails  $\lambda(x) = 0$ : indeed, if  $\lambda(x) \neq 0$ , then there is a real number  $t$  such that  $\lambda(tx) > 1$ , so  $f_i(x) \notin V_i$ , whence  $f_i(x) \neq 0$ . Therefore, there exists a linear functional  $\tilde{\lambda}$  such that the diagram

$$\begin{array}{ccc}
 & & \mathbb{R} \\
 & \nearrow \lambda & \uparrow \tilde{\lambda} \\
 \mathcal{X} & \xrightarrow{q} & \mathcal{X} / \text{Ker}(f_i) \\
 & \searrow f_i & \downarrow \tilde{f}_i \\
 & & f_i(\mathcal{X}) \hookrightarrow \mathcal{X}_i
 \end{array}$$

is commutative, where  $q$  is the quotient map and  $\tilde{f}_i$  the isomorphism induced by  $f_i$ . The map  $\tilde{\lambda}_i = \tilde{\lambda} \circ \tilde{f}_i^{-1}$  is a linear functional on  $f_i(\mathcal{X})$  and  $V_i \cap f_i(\mathcal{X}) \subset \{\tilde{\lambda}_i < 1\}$ : indeed, for  $x_i \in V_i \cap f_i(\mathcal{X})$ , there exists  $x \in V = f_i^{-1}(V_i)$  such that  $f_i(x) = x_i$ , whence

$$\tilde{\lambda}_i(x_i) = \tilde{\lambda} \circ \tilde{f}_i^{-1} \circ f_i(x) = \tilde{\lambda} \circ \tilde{f}_i^{-1} \circ \tilde{f}_i \circ q(x) = \tilde{\lambda} \circ q(x) = \lambda(x) < 1.$$

So  $\tilde{\lambda}_i$  is a continuous linear functional on  $f_i(\mathcal{X})$ . According to the Hahn-Banach theorem, there exists a linear extension  $\lambda_i \in \mathcal{X}_i^*$  of  $\tilde{\lambda}_i$  and it satisfies  $\lambda = \lambda_i \circ f_i$ .  $\square$

*Proof of Theorems 1 and 2.* Let us begin with the proof of

$$s = \inf_{i \in I} s_i \circ f_i, \quad \bar{s} = \inf_{i \in I} \bar{s}_i \circ f_i, \quad \text{and} \quad p^* = \sup_{i \in I} p_i^* \circ f_i. \quad (4)$$

Using Proposition 1, for all  $x \in \mathcal{X}$ ,

$$\begin{aligned} s(x) &= \inf_{V \in \mathcal{V}_x} \liminf_{n \rightarrow \infty} \frac{1}{v_n} \log \mathbb{P}(Z_n \in V) \\ &= \inf_{i \in I} \inf_{V_i \in \mathcal{V}_{i, f_i(x)}} \liminf_{n \rightarrow \infty} \frac{1}{v_n} \log \mathbb{P}(f_i(Z_n) \in V_i) = \inf_{i \in I} s_i(f_i(x)). \end{aligned}$$

The proof for  $\bar{s}$  is similar. Finally, using Proposition 2,

$$p^*(x) = \sup_{\lambda \in \mathcal{X}^*} (\lambda(x) - p(\lambda)) = \sup_{i \in I} \sup_{\lambda_i \in \mathcal{X}_i^*} (\lambda_i \circ f_i(x) - p(\lambda_i \circ f_i)) = \sup_{i \in I} p_i^*(f_i(x)).$$

Assertions 1 and 2 follow immediately. Let us prove assertion 3. Let  $K \in \mathcal{F}$  be a relatively compact subset of  $\mathcal{X}$  and  $i \in I$ . Then, using the fact that  $f_i(\overline{K})$  is compact, thus measurable,

$$\limsup_{n \rightarrow \infty} \frac{1}{v_n} \log \mathbb{P}(Z_n \in K) \leq \limsup_{n \rightarrow \infty} \frac{1}{v_n} \log \mathbb{P}(f_i(Z_n) \in f_i(\overline{K})) \leq \sup_{f_i(\overline{K})} s_i = \sup_{\overline{K}} s_i \circ f_i$$

and it remains to prove the min-max inequality

$$\inf_{i \in I} \sup_{\overline{K}} s_i \circ f_i \leq \sup_{\overline{K}} \inf_{i \in I} s_i \circ f_i = \sup_{\overline{K}} s. \quad (5)$$

Let  $\alpha > 0$ . Using (4), for all  $x \in \overline{K}$ , there exists  $i(x) \in I$  such that

$$s_{i(x)} \circ f_{i(x)}(x) \leq \max(s(x) + \alpha/2, -1/\alpha - \alpha/2).$$

Since  $s_{i(x)} \circ f_{i(x)}$  is upper semi-continuous, there exists  $V(x) \in \mathcal{V}_x$  such that

$$\sup_{V(x)} s_{i(x)} \circ f_{i(x)} \leq \max(s_{i(x)} \circ f_{i(x)}(x) + \alpha/2, -1/\alpha).$$

Hence

$$\sup_{V(x)} s_{i(x)} \circ f_{i(x)} \leq \max(s(x) + \alpha, -1/\alpha).$$

The cover  $\{V(x) ; x \in \overline{K}\}$  of  $\overline{K}$  admits a finite subcover  $\{V(x_k) ; k \in \{1, \dots, r\}\}$ . Since  $I$  is a directed set, there exists an upper bound  $j \in I$  of  $\{i(x_1), \dots, i(x_r)\}$ . Let us check that

$$\forall k \in \{1, \dots, r\} \quad s_j \circ f_j \leq s_{i(x_k)} \circ f_{i(x_k)}. \quad (6)$$

For  $i \leq j$ ,

$$s_i \circ f_i = s_i \circ f_{i,j} \circ f_j,$$

so we only need to prove that  $s_j \leq s_i \circ f_{i,j}$ . For all  $x \in X_j$ , since  $f_{ij}^{-1}(V_i) \in \mathcal{V}_{j,x}$ ,

$$\begin{aligned} \inf_{V_j \in \mathcal{V}_{j,x}} \liminf_{n \rightarrow \infty} \frac{1}{v_n} \log \mathbb{P}(f_j(Z_n) \in V_j) &\leq \inf_{V_i \in \mathcal{V}_{i,f_{i,j}(x)}} \liminf_{n \rightarrow \infty} \frac{1}{v_n} \log \mathbb{P}(f_{ij} \circ f_j(Z_n) \in V_i) \\ &= \inf_{V_i \in \mathcal{V}_{i,f_{i,j}(x)}} \liminf_{n \rightarrow \infty} \frac{1}{v_n} \log \mathbb{P}(f_i(Z_n) \in V_i) \end{aligned}$$

and equation (6) is proved. Hence

$$\sup_{\overline{K}} s_j \circ f_j \leq \max \left( \max_{1 \leq k \leq r} s(x_k) + \alpha, -1/\alpha \right) \leq \max \left( \sup_{\overline{K}} s + \alpha, -1/\alpha \right).$$

Conclude to (5) by letting  $\alpha \rightarrow 0$ .  $\square$

## 4 Application to Cramér's theory

Let  $(\mathcal{X}, \tau, \mathcal{F})$  be a l.m.v.s. Let  $(X_n)_{n \geq 1}$  be a sequence of independent and identically distributed (i.i.d.) random variables over  $(\mathcal{X}, \mathcal{F})$ . In this section, we consider the case when  $Z_n$  is the empirical mean  $\overline{X}_n = (X_1 + X_2 + \dots + X_n)/n$ . In order that  $Z_n$  be well defined, we introduce extra assumptions:

(LM3) the addition  $(x, y) \in (\mathcal{X}^2, \mathcal{F}^{\otimes 2}) \mapsto x + y \in (\mathcal{X}, \mathcal{F})$  is measurable;

(LM4) the  $\sigma$ -algebra  $\mathcal{F}$  is dilatation-invariant.

As before, we define the entropy  $s$ , the entropy superior  $\overline{s}$ , and the pressure  $p$  associated to the sequence  $(\overline{X}_n, n)_{n \geq 1}$ , or shortly to the distribution  $\mu$  of  $X_1$ . In this context, the pressure simplifies like this:

$$p(\lambda) = \log \mathbb{E}[e^{\lambda(X_1)}].$$

*Remark.* All the spaces we will consider will be locally convex spaces in which each point admits a local basis made up of open convex measurable neighborhoods (see [Cer07, Chapter 4] who introduces this natural assumption). In this context, we also have  $s = \overline{s}$ : the proof of [Pet18, Proposition 2] adapts without change.

It can be found in standard texts that any separable Banach space  $(\mathcal{B}, \tau_{\|\cdot\|}, \sigma(\tau_{\|\cdot\|}))$  endowed with its Borel  $\sigma$ -algebra satisfies (LM1), (LM2), (LM3), and (LM4). Moreover, in this setting:

**Theorem 3.** *Let  $(X_n)_{n \geq 1}$  be a sequence of i.i.d. random variables on a separable Banach space  $(\mathcal{B}, \tau_{\|\cdot\|}, \sigma(\tau_{\|\cdot\|}))$ . The sequence  $(\overline{X}_n, n)_{n \geq 1}$  satisfies a weak large deviation principle with entropy  $s = -p^*$  where  $p(\lambda) = \log \mathbb{E}(e^{\lambda(X_1)})$ .*

*Proof.* See, e.g., [DZ98, Theorem 6.1.3], [Cer07, Corollary 16.4], or [Pet18, Theorems 12.1 and 12.4] for different proofs.  $\square$

In this section we provide general explicit locally convex spaces which satisfy (LM1), (LM2), (LM3), and (LM4), in which the equality  $s = -p^*$  is always true. In Section 5, we provide an example where the space satisfy (LM1), (LM2), (LM3), and (LM4), but in which  $s = \overline{s} \neq -p^*$ .



#### 4.1 Cramér's theorem in separable normed vector spaces

Here, we relax the hypothesis of completeness in Theorem 3. Let  $(\mathcal{N}, \tau_{\|\cdot\|}, \sigma(\tau_{\|\cdot\|}))$  be a separable normed vector space (not necessarily complete) endowed with its Borel  $\sigma$ -algebra. Obviously,  $(\mathcal{N}, \tau_{\|\cdot\|}, \sigma(\tau_{\|\cdot\|}))$  satisfies (LM1), (LM2) and (LM4). And, since  $\mathcal{N}$  is second countable, (LM3) is satisfied.

**Theorem 4.** *Let  $(X_n)_{n \geq 1}$  be a sequence of i.i.d. random variables on a separable normed vector space  $(\mathcal{N}, \tau_{\|\cdot\|}, \sigma(\tau_{\|\cdot\|}))$ . The sequence  $(\overline{X}_n, n)_{n \geq 1}$  satisfies a weak large deviation principle with entropy  $s = -p^*$  where  $p(\lambda) = \log \mathbb{E}(e^{\lambda(X_1)})$ .*

*Proof.* Just consider the completion  $(\hat{\mathcal{N}}, \hat{\tau})$  of  $(\mathcal{N}, \tau_{\|\cdot\|})$ , which is a separable Banach space. Let  $i: \mathcal{N} \hookrightarrow \hat{\mathcal{N}}$  be the canonical inclusion. By construction,  $\tau_{\|\cdot\|}$  is the inverse image topology of  $\hat{\tau}$  with respect to  $i$ . And the Borel  $\sigma$ -algebra satisfies

$$\sigma(\tau_{\|\cdot\|}) = \sigma(i^{-1}(\hat{\tau})) = i^{-1}(\sigma(\hat{\tau})).$$

So  $(\mathcal{N}, \tau_{\|\cdot\|}, \sigma(\tau_{\|\cdot\|}))$  is the inverse image of the l.m.v.s.  $(\hat{\mathcal{N}}, \hat{\tau}, \sigma(\hat{\tau}))$  with respect to  $i$ . According to Theorem 3,  $(i(\overline{X}_n), n)_{n \geq 1}$  satisfies a weak large deviation principle with entropy  $\hat{s} = -\hat{p}^*$  where  $\hat{p}(\lambda) = \log \mathbb{E}(e^{\lambda(i(X_1))})$ . Eventually apply Theorem 1 to conclude with  $s = \hat{s}|_{\mathcal{N}}$  and  $p^* = \hat{p}^*|_{\mathcal{N}}$ .  $\square$

#### 4.2 Cramér's theorem in locally convex separable spaces

The following simple construction makes explicit the settings developped in [BZ79], [DZ98], and [Cer07], and generalizes both Cramér's theorem in separable Banach spaces and Sanov's theorem. It will be generalized even more in the next section, at the expense of technical notions of set theory. Let  $\mathcal{X}$  be a vector space and consider a family  $\mathcal{C}$  of subsets of  $\mathcal{X}$  satisfying

(LCS) every  $C \in \mathcal{C}$  is a symmetric algebraically open convex set containing 0 such that the locally convex topology generated by  $C$  is separable (or equivalently second-countable).

For all  $x \in \mathcal{X}$ , define

$$\mathcal{C}_x = \left\{ x + \bigcap_{i=1}^r \lambda_i C_i ; r \geq 0, C_i \in \mathcal{C}, \lambda_i > 0 \right\}.$$

Then,  $\mathcal{C}_0$  is a local basis of 0 for a locally convex topology  $\tau(\mathcal{C})$  on  $\mathcal{X}$  (see, e.g., [Bou81, II.25]). As for the  $\sigma$ -algebra, let

$$\mathcal{F}(\mathcal{C}) = \sigma\left(\bigcup_{x \in \mathcal{X}} \mathcal{C}_x\right).$$

Let us make several points on the setting. First, if  $C$  is an algebraically open convex subset containing 0, then  $C \cap (-C)$  is a symmetric algebraically open convex subset containing 0, so the symmetry can be assumed without any restriction, but it is handy when introducing seminorms. Secondly,  $\mathcal{F}(\mathcal{C})$  is not necessarily the Borel  $\sigma$ -algebra  $\sigma(\tau(\mathcal{C}))$ : it is the case when  $\tau(\mathcal{C})$  is second-countable. Thirdly, each  $C \in \mathcal{C}_0$  is symmetric, and the locally convex topology generated by  $C$  is separable: indeed, the upper bound of finitely many second-countable topologies is still second-countable. However,  $\tau(\mathcal{C})$  is not necessarily separable (a fortiori not second-countable).

For instance, take  $\mathcal{X} = \mathbb{R}^{\mathbb{R}}$  endowed with the product topology  $\tau$  and the cylinder  $\sigma$ -algebra  $\mathcal{F}$ . The topology  $\tau$  is not separable and the space  $(\mathcal{X}, \tau, \mathcal{F})$  is obtained by the previous construction with  $\mathcal{C}$  being the set of all open half-spaces  $C = \{f \in \mathcal{X} \mid f(x) < 1\}$  ( $x \in \mathbb{R}$ ).

In the following, we will denote by  $M_C$  the Minkowski functional of any convex subset  $C$  of  $\mathcal{X}$  containing 0 (see, *e.g.*, [Zäl02]). In particular, if  $C \in \mathcal{C}$ ,  $M_C$  is a seminorm.

**Proposition 3.**  $(\mathcal{X}, \tau(\mathcal{C}), \mathcal{F}(\mathcal{C}))$  satisfies (LM1), (LM2), (LM3), and (LM4).

As a result, we will say that  $(\mathcal{X}, \tau(\mathcal{C}), \mathcal{F}(\mathcal{C}))$  is a *locally convex separable space* (l.c.s.s.).

*Proof.* By construction, (LM1) and (LM4) are satisfied. To prove (LM2), let  $\lambda \in \mathcal{X}^*$  and let  $H$  be the open half-space  $\{x \in \mathcal{X} \mid \lambda(x) < 1\}$ . We only need to prove that  $H$  is measurable. Since  $H$  is a neighbourhood of 0, there exists  $C \in \mathcal{C}_0$  such that  $C \subset H$ . Let us show that  $H$  is a countable union of homothetic transforms of  $C$ , so that  $H$  is measurable. Let  $Q$  be a countable subset of  $\mathcal{X}$  which is dense with respect to the locally convex topology  $\tau_C$  generated by  $C$ . We have

$$H \supseteq \bigcup_{\substack{u \in Q, r \in \mathbb{Q}_+ \\ u + rC \subset H}} u + rC \quad (7)$$

and we want to show the converse inclusion. Let  $x \in H$ . Since  $H$  is open,  $M_H(x) < 1$  and there exists  $r \in \mathbb{Q}_+$  such that  $0 < r < 1 - M_H(x)$ ; so  $x + rC \subset H$ . The set  $x + rC/2$  is open with respect to  $\tau_C$ , so contains a point  $u$  of  $Q$ . Then,  $x \in u + rC/2 \subset x + rC \subset H$  and the converse inclusion of (7) is proved.

Now we prove (LM3), i.e. the addition  $(x, y) \in (\mathcal{X}^2, \mathcal{F}(\mathcal{C})^{\otimes 2}) \mapsto x + y \in (\mathcal{X}, \mathcal{F}(\mathcal{C}))$  is measurable. Let  $C \in \mathcal{C}_0$ . We only need to prove that  $\{(x, y) \in \mathcal{X}^2 \mid x + y \in C\} \in \mathcal{F}(\mathcal{C})^{\otimes 2}$ . Remember that the locally convex topology  $\tau_C$  generated by  $C$  is separable. Let  $Q$  be a countable subset of  $\mathcal{X}$  which is dense with respect to  $\tau_C$ . Let us show that

$$\{(x, y) \in \mathcal{X}^2 \mid x + y \in C\} = \bigcup_{\substack{u \in Q \\ r \in \mathbb{Q}}} \{(x, y) \in \mathcal{X}^2 \mid M_C(x - u) < r, M_C(y + u) < 1 - r\}.$$

To show  $\subset$ , let  $(x, y) \in \mathcal{X}^2$  such that  $x + y \in C$ . Let  $a := M_C(x + y) < 1$ . Let  $r \in \mathbb{Q} \cap (0, (1 - a)/2)$ . The nonempty open subset  $x + rC$  contains a point  $u$  of  $Q$ . Then  $M_C(x - u) < r$  (since  $C$  is symmetric) and

$$M_C(y + u) \leq M_C(y + x) + M_C(u - x) < a + r < 1 - r.$$

To show  $\supset$ , just notice that, for each  $(x, y, u) \in \mathcal{X}^3$ ,

$$M_C(x + y) \leq M_C(x - u) + M_C(y + u).$$

□

**Theorem 5.** Let  $(X_n)_{n \geq 1}$  be a sequence of i.i.d. random variables on a l.c.s.s.  $(\mathcal{X}, \tau(\mathcal{C}), \mathcal{F}(\mathcal{C}))$ . The sequence  $(\overline{X}_n, n)_{n \geq 1}$  satisfies a weak large deviation principle with entropy  $s = -p^*$  where  $p(\lambda) = \log \mathbb{E}(e^{\lambda(X_1)})$ .

*Proof.* It appears to be that the previous construction of  $(\mathcal{X}, \tau(\mathcal{C}), \mathcal{F}(\mathcal{C}))$  is equivalent to the following one. For each  $C \in \mathcal{C}_0$ , let  $\tau_C$  (resp.  $\mathcal{F}_C$ ) be the locally convex topology generated by  $C$  (resp. the  $\sigma$ -algebra generated by all homothetic transforms of  $C$ ). Then,  $(\mathcal{X}, \tau(\mathcal{C}), \mathcal{F}(\mathcal{C}))$  is the projective limit of the projective system of l.m.v.s.  $(\mathcal{X}, \tau_C, \mathcal{F}_C, \text{id}_{\mathcal{X}})_{C \supseteq B}$ . Now, for  $C \in \mathcal{C}_0$ , consider the Hausdorff space  $(\mathcal{N}_C, \tilde{\tau}_C)$  canonically associated to  $(\mathcal{X}, \tau_C)$  and  $f_C: \mathcal{X} \rightarrow \mathcal{N}_C$  the canonical projection. Then, since  $\tau_C$  is separable,  $\mathcal{N}_C$  is a separable normed vector space, and  $(\mathcal{X}, \tau_C, \mathcal{F}_C)$  is the inverse image of the l.m.v.s.  $(\mathcal{N}_C, \tilde{\tau}_C, \sigma(\tilde{\tau}_C))$  under  $f_C$  (notice that, since  $(\mathcal{N}_C, \tilde{\tau}_C)$  is a separable normed space, the  $\sigma$ -algebra generated by the open balls is equal to the Borel  $\sigma$ -algebra  $\sigma(\tilde{\tau}_C)$ ). By Theorem 4, the conclusion holds in  $(\mathcal{N}_C, \tilde{\tau}_C, \sigma(\tilde{\tau}_C))$ . And the conclusion holds in  $(\mathcal{X}, \tau(\mathcal{C}), \mathcal{F}(\mathcal{C}))$  through theorems 1 and 2.  $\square$

The model of Theorem 5 is closely related to the models of [BZ79] and [Cer07]. In [BZ79], they consider a Hausdorff locally convex topological vector space  $(\mathcal{X}, \tau)$  endowed with the Borel  $\sigma$ -algebra  $\sigma(\tau)$  and assume conditions of measurability so that  $\overline{X}_n$  be measurable. Remark that  $p$  is automatically defined when the  $\sigma$ -algebra is the Borel  $\sigma$ -algebra. Here, we do not suppose the  $\sigma$ -algebra to be the Borel  $\sigma$ -algebra. In [Cer07], the  $\sigma$ -algebra is not supposed to be the Borel  $\sigma$ -algebra either, but it is assumed that the vector space operations are measurable, that each point of  $\mathcal{X}$  admits a local basis of open convex measurable subsets of  $\mathcal{X}$ , and that the continuous linear functionals are measurable. Here, we do not require the scalar multiplication to be measurable. More importantly, Theorem 5 provides a simple constructive model for Cramér's theory: separability appears to be the crux of the matter of measurability (even in the generalization exposed in section 4.3). And the model embraces standard applications (see, *e.g.*, [Cer07, chap. 16]). Indeed, we obtain the weak large deviation principle and the equality  $s = -p^*$  when:

- $\mathcal{X}$  is a separable normed vector space endowed with the Borel  $\sigma$ -algebra ( $\mathcal{C} = \{B(0, 1)\}$ ), in particular when  $\mathcal{X}$  is a separable Banach space (standard Cramér's theorem);
- $\mathcal{X}$  is a vector space in duality with another one  $\mathcal{Y}$ , and  $\mathcal{X}$  is endowed with the weak-topology and the cylinder  $\sigma$ -algebra ( $\mathcal{C}$  is the collection of cylinders  $\{|\langle \cdot, y \rangle| < 1\}$  for  $y \in \mathcal{Y}$ ), which allows to recover Sanov's theorem in the  $\tau$ -topology.

### 4.3 Cramér's theorem in locally convex measurable spaces

In the previous section, we have obtained Cramér's theorem in any projective limit of separable normed spaces. Here, we relax the hypothesis of separability. Let  $(\mathcal{N}, \tau_{\|\cdot\|})$  be a normed space. We say that  $(\mathcal{N}, \tau_{\|\cdot\|}, \sigma(\tau_{\|\cdot\|}))$  is a *measurable normed space* if the addition  $(x, y) \mapsto x + y$  is measurable with respect to the  $\sigma$ -algebras  $\sigma(\tau_{\|\cdot\|}) \otimes \sigma(\tau_{\|\cdot\|})$  and  $\sigma(\tau_{\|\cdot\|})$ . In particular,  $(\mathcal{N}, \tau_{\|\cdot\|}, \sigma(\tau_{\|\cdot\|}))$  satisfies (LM1), (LM2), (LM3), and (LM4).

**Theorem 6.** *Let  $(X_n)_{n \geq 1}$  be a sequence of i.i.d. random variables on a measurable normed space  $(\mathcal{N}, \tau_{\|\cdot\|}, \sigma(\tau_{\|\cdot\|}))$ . The sequence  $(\overline{X}_n, n)_{n \geq 1}$  satisfies a weak large deviation principle with entropy  $s = -p^*$  where  $p(\lambda) = \log \mathbb{E}(e^{\lambda(X_1)})$ .*

*Remark.* At this point, we can say that the equality  $s = -p^*$  for empirical means of i.i.d. random variables is always true in the context of normed spaces with their Borel  $\sigma$ -algebra, as soon as the question makes sense. That is the reason why the example given in Section 5 requires a little more effort.

As in the construction in Section 4.2, let us call *locally convex measurable space* (l.c.m.s.) any  $(\mathcal{X}, \tau(\mathcal{C}), \mathcal{F}(\mathcal{C}))$ ,  $\mathcal{C}$  being any family of subsets of  $\mathcal{X}$  satisfying

(LCM) every  $C \in \mathcal{C}$  is a symmetric algebraically open convex set containing 0 such that  $(\mathcal{N}_C, \tilde{\tau}_C)$  is a measurable normed space,

where we recall that  $(\mathcal{N}_C, \tilde{\tau}_C)$  is the Hausdorff space canonically associated to  $(\mathcal{X}, \tau_C)$ ,  $\tau_C$  being the locally convex topology generated by  $C$ . Following the proof of Theorem 5, we see that a l.c.m.s. is a projective limit of inverse images of measurable normed spaces and we deduce the following result.

**Theorem 7.** *Let  $(X_n)_{n \geq 1}$  be a sequence of i.i.d. random variables on a l.c.m.s.  $(\mathcal{X}, \tau(\mathcal{C}), \mathcal{F}(\mathcal{C}))$ . The sequence  $(\bar{X}_n, n)_{n \geq 1}$  satisfies a weak large deviation principle with entropy  $s = -p^*$  where  $p(\lambda) = \log \mathbb{E}(e^{\lambda(X_1)})$ .*

The proof of Theorem 6 requires some technical notions of set theory. Essentially, we show that any probability measure over a measurable normed space is supported by a separable subspace and apply Theorem 4. Let  $\kappa$  be the *density* of  $\mathcal{N}$ , i.e. the smallest cardinal  $\lambda$  such that there exists a dense subset  $D$  of  $X$  of cardinality  $\lambda$ . We denote by  $\mathfrak{P}(E)$  the power set of the set  $E$ .

**Lemma 1.**  $\mathfrak{P}(\kappa \times \kappa) = \mathfrak{P}(\kappa) \otimes \mathfrak{P}(\kappa)$ .

*Proof.* The result is proved in [Tal79]. □

**Lemma 2.** *There is no probability measure over  $(\kappa, \mathfrak{P}(\kappa))$  such that every singleton has measure zero.*

*Proof.* Lemma 1 implies (see [Kun68, Lemma 12.2]) that  $\kappa \leq 2^{\aleph_0}$  and, for all  $\lambda \leq \kappa$ ,  $\mathfrak{P}(\lambda \times \lambda) = \mathfrak{P}(\lambda) \otimes \mathfrak{P}(\lambda)$ . Then, applying [Kun68, Theorem 12.9], we get that, for all  $\lambda \leq \kappa$ ,  $\lambda$  is not real-valued measurable (see [Kun68, Definition 11.1]). And the conclusion stems from [Dra74, Theorem 1.4]. □

Now, let  $\mu$  be a probability measure over  $(\mathcal{N}, \tau_{\|\cdot\|}, \sigma(\tau_{\|\cdot\|}))$ .

**Lemma 3.** *There exists a closed separable subspace  $\tilde{\mathcal{N}}$  of  $\mathcal{N}$  such that  $\mu(\tilde{\mathcal{N}}) = 1$ .*

*Proof.* Since there is no diffuse probability measure over  $(\kappa, \mathfrak{P}(\kappa))$  (see Lemma 2), we can apply [MS48, Theorem III] and deduce that  $\mathcal{N} = S \cup Z$  where  $S$  is a separable subset of  $\mathcal{N}$ , and  $\mu(Z) = 0$  ( $Z$  is the reunion of the negligible open subsets of  $\mathcal{N}$  and  $S = \mathcal{N} \setminus Z$ ). Then, define  $\tilde{\mathcal{N}} = \overline{\text{Vect}(S)}$  the closed linear span of  $S$ .  $\tilde{\mathcal{N}}$  is indeed separable since, if  $D$  is a countable dense subset of  $S$ , then  $\text{Vect}_{\mathbb{Q}}(D)$  is a countable dense subset of  $\tilde{\mathcal{N}}$ . □

*Proof of Theorem 6.* Let  $\mu$  be the distribution of  $X_1$ . Let  $\tilde{\mathcal{N}}$  be a closed separable subspace of  $\mathcal{N}$  such that  $\mu(\tilde{\mathcal{N}}) = 1$  (see Lemma 3). Let  $\tilde{s}$  (resp.  $\tilde{p}$ ) be the entropy (resp. the pressure) associated to  $\tilde{\mu} := \mu|_{\tilde{\mathcal{N}}}$ . Theorem 4 implies that  $\tilde{s} = \tilde{p}$ . Now,

$$s(x) = \begin{cases} \tilde{s}(x) & \text{if } x \in \tilde{\mathcal{N}} \\ -\infty & \text{if } x \in \mathcal{N} \setminus \tilde{\mathcal{N}}. \end{cases} \quad (8)$$

Moreover, if  $x \in \tilde{\mathcal{N}}$ , since  $\mu(\mathcal{N} \setminus \tilde{\mathcal{N}}) = 0$  and since the restriction  $\tilde{\lambda}$  of any  $\lambda \in \mathcal{N}^*$  to  $\tilde{\mathcal{N}}$  is in  $\tilde{\mathcal{N}}^*$ ,

$$\begin{aligned} p^*(x) &= \sup_{\lambda \in \mathcal{N}^*} \left( \lambda(x) - \log \int_{\tilde{\mathcal{N}}} e^{\lambda(y)} d\mu(y) \right) \\ &= \sup_{\tilde{\lambda} \in \tilde{\mathcal{N}}^*} \left( \tilde{\lambda}(x) - \log \int_{\tilde{\mathcal{N}}} e^{\tilde{\lambda}(y)} d\tilde{\mu}(y) \right) \\ &= \tilde{p}^*(x). \end{aligned} \tag{9}$$

And, if  $x \in \mathcal{N} \setminus \tilde{\mathcal{N}}$ , the Hahn-Banach theorem provides a  $\lambda \in \mathcal{N}^*$  such that  $\lambda(x) = 1$  and  $\lambda$  is zero on  $\tilde{\mathcal{N}}$ ; therefore, since  $\mu(\mathcal{N} \setminus \tilde{\mathcal{N}}) = 0$ ,

$$p^*(x) \geq \sup_{t \geq 0} \left( t\lambda(x) - \log \int_{\tilde{\mathcal{N}}} e^{t\lambda(y)} d\mu(y) \right) = \sup_{t \geq 0} t = +\infty. \tag{10}$$

The conclusion follows from eqs. (8) to (10).  $\square$

## 5 A counterexample to $s = -p^*$

In this section, we define a sequence  $(\overline{X}_n, n)_{n \geq 1}$  for which  $s = \overline{s} \neq -p^*$ . The requirements on the space are quite precise: the topology must be rich enough so that there is no (hidden) convex-tension; the  $\sigma$ -algebra must be rich enough, so that the space be a l.m.v.s., but not too big, so that the addition be measurable. We will work in the space  $\ell^\infty$  of bounded real sequences, equipped with the topology  $\tau_{\|\cdot\|}$  of the norm. The open ball of radius  $r$  and centered in  $x$  will be denoted by  $B(x, r)$ . We denote by  $\mathcal{C}(\ell^\infty, (\ell^\infty)^*)$  the *cylinder  $\sigma$ -algebra*, i.e. the  $\sigma$ -algebra over  $\ell^\infty$  generated by all the continuous linear functionals on  $\ell^\infty$  (see [VTC87]).

**Proposition 4.**  $(\ell^\infty, \tau_{\|\cdot\|}, \mathcal{C}(\ell^\infty, (\ell^\infty)^*))$  satisfies (LM1), (LM2), (LM3), and (LM4).

*Proof.* By construction, (LM2) and (LM4) are satisfied. To prove (LM1), it suffices to see that the open balls of  $\ell^\infty$  are in  $\mathcal{C}(\ell^\infty, (\ell^\infty)^*)$ : indeed, for all  $x \in \ell^\infty$  and  $r > 0$ ,

$$B(x, r) = \bigcap_{i=1}^{\infty} \{y \in \ell^\infty \mid |e_i^*(y - x)| < r\},$$

where  $(e_i^*)_{i \geq 1}$  is the family of coordinate linear functionals. Finally, (LM3) holds since we have  $\mathcal{C}(\ell^\infty, (\ell^\infty)^*) = \mathcal{F}(\mathcal{C})$ ,  $\mathcal{C}$  being the set of cylinders  $\{|\lambda(\cdot)| < 1\}$  for  $\lambda \in (\ell^\infty)^*$ , and repeating the same proof as that of Proposition 3 (see also [VTC87, Proposition 2.3]).  $\square$

Now, we consider the Cantor space  $C = \{0, 1\}^{\mathbb{N}} \subset \ell^\infty$ , equipped with the product  $\sigma$ -algebra  $\mathfrak{P}(\{0, 1\})^{\otimes \mathbb{N}}$ , and the measure  $\mu = \mathcal{B}(1/2)^{\otimes \mathbb{N}}$ . The measure  $\mu$  is the pushforward measure of the Lebesgue measure  $\lambda$  over the Borel  $\sigma$ -algebra  $\mathcal{B}([0, 1])$  through the dyadic development  $\varphi: [0, 1] \rightarrow C$ .

**Proposition 5.** *The Lebesgue measure extends to the  $\sigma$ -algebra  $\mathfrak{P}([0, 1])$  if and only if there exists a real-valued measurable cardinal not greater than  $2^{\aleph_0}$ .*

*Proof.* See [Jec06, p. 131] (see also Fremlin, Real-valued-measurable cardinals (2009) (1D)).  $\square$

Suppose that there exists a real-valued measurable cardinal not greater than  $2^{\aleph_0}$ . Let  $\tilde{\lambda}$  be an extension of  $\lambda$  to  $\mathfrak{P}([0, 1])$ . We can suppose that  $\tilde{\lambda}$  is symmetric, i.e.  $\sigma: x \mapsto 1 - x$  is  $\tilde{\lambda}$ -invariant: take an extension  $\tilde{\lambda}$  of  $\lambda$  over  $\mathfrak{P}([0, 1/2])$  and define  $\tilde{\lambda}$  over  $\mathfrak{P}((1/2, 1])$  by  $\tilde{\lambda}(A) = \tilde{\lambda}(\sigma(A))$ . Now, the pushforward measure  $\tilde{\mu} = \tilde{\lambda} \circ \varphi^{-1}$  is an extension of  $\mu$  to  $\mathfrak{P}(C)$ . Obviously,  $\tilde{\mu}$  extends to a measure over  $\mathfrak{P}(\ell^\infty)$ , and we consider the restriction of the latter to  $\mathcal{C}(\ell^\infty, (\ell^\infty)^*)$ , which will still be denoted by  $\tilde{\mu}$ . By construction,  $\tilde{\mu}$  is symmetric, i.e. if  $X$  is a random variable with distribution  $\tilde{\mu}$ , then the distribution of  $1 - X$  is also  $\tilde{\mu}$  (here 1 denotes the constant sequence  $(1)_{n \geq 1}$ ).

In particular, for all  $f \in (\ell^\infty)^*$ ,  $f(X)$  has the same distribution as  $f(1 - X) = f(1) - f(X)$ , whence

$$\forall f \in (\ell^\infty)^* \quad \mathbb{E}[f(X)] = f(1/2).$$

Now, let  $(X_n)_{n \geq 1}$  be a sequence of i.i.d. random variables on  $\ell^\infty$  with distribution  $\tilde{\mu}$ . Let  $(\bar{X}_n)_{n \geq 1}$  be the associate sequence of empirical means. Let  $p$  be the pressure of  $(\bar{X}_n, n)_{n \geq 1}$ . We have

$$\begin{aligned} p^*(1/2) &= - \inf_{f \in (\ell^\infty)^*} \log \mathbb{E}[e^{f(1/2) - f(X)}] \\ &= - \inf_{f \in (\ell^\infty)^*} \inf_{t \in \mathbb{R}} \log \mathbb{E}[e^{t(f(1/2) - f(X))}] \\ &= 0. \end{aligned}$$

Now, let us evaluate the entropy superior  $\bar{s}$  of  $(\bar{X}_n, n)_{n \geq 1}$ . For any  $x \in \ell^\infty$ , we have

$$\bar{s}(x) = \inf_{\varepsilon > 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^{\infty} \log \mathbb{P}(|X_{1,i} + \dots + X_{n,i} - nx_i| < n\varepsilon) = -\infty$$

since  $\log \mathbb{P}(|X_{1,i} + \dots + X_{n,i} - nx_i| < n\varepsilon)$  is negative and independent of  $i$ . In particular, we have proved that  $s = \bar{s} = -\infty \neq -p^*$ .

*Remark.* We cannot simply consider the probability space  $(\ell^\infty, \mathfrak{P}(\ell^\infty), \tilde{\mu})$ : indeed, as we have supposed that there exists a real-valued measurable cardinal not greater than  $2^{\aleph_0}$ , we can show that the addition on  $(\ell^\infty, \mathfrak{P}(\ell^\infty))$  is not measurable (see [Tal79] and [Kun68, Lemma 12.2 and Theorem 12.9]).

*Question.* The existence of a real-valued measurable cardinal not greater than  $2^{\aleph_0}$  is incompatible with the Continuum Hypothesis. Is there nevertheless a counterexample compatible with the Continuum Hypothesis?

## 6 About the equality $p = (-s)^*$

Throughout this section,  $(X_n)_{n \geq 1}$  is a sequence of i.i.d. random variables on a l.c.m.s.  $(\mathcal{X}, \tau(\mathcal{C}), \mathcal{F}(\mathcal{C}))$ , and  $s$  (resp.  $p$ ) denotes the entropy (resp. pressure) of the sequence  $(\bar{X}_n, n)_{n \geq 1}$ . Notice that our proof of  $s = -p^*$  (see Theorem 7) basically relies on Theorem 3, so on the dual equality  $p = (-s)^*$  valid in any separable Banach space (see [Pet18, Theorem 12.4]), where the Fenchel-Legendre transform of a function  $g: \mathcal{X} \rightarrow [-\infty, +\infty]$  is the function  $g^*: \mathcal{X}^* \rightarrow [-\infty, +\infty]$  defined by

$$\forall \lambda \in \mathcal{X}^* \quad g^*(\lambda) = \sup_{x \in \mathcal{X}} (\lambda(x) - g(x)).$$

That being said, the equality  $p = (-s)^*$  is generally stronger than the dual equality.

**Proposition 6.** *One has  $p = (-s)^*$  if and only if  $s = -p^*$  and  $p$  is  $\sigma(\mathcal{X}^*, \mathcal{X})$ -lower semi-continuous.*

*Proof.* Just take the convex conjugate and notice that  $s$  is always upper semi-continuous (cf. [Mor67, §5.d]).  $\square$

**Example 1.** Here is an example of a l.c.s.s., borrowed from [Dud69, Section 2], where  $p \neq (-s)^*$  (and  $s = -p^*$  by Theorem 5). In particular, it illustrates the fact that, contrary to the equality  $s = -p^*$ , the equality  $p = (-s)^*$  is not preserved by projective limits. Let  $\mathcal{Y} = \mathcal{C}([0, \omega_1]; \mathbb{R})$  be the space of continuous functions on the ordinal space  $\omega_1$ . Endow  $\mathcal{Y}$  with the topology  $\tau$  of uniform convergence on compact subsets of  $[0, \omega_1)$ . Recall that  $\mathcal{Y}^*$  consists of the bounded regular Borel measures on  $[0, \omega_1)$  with compact support.

Each  $f \in \mathcal{Y}$  is eventually constant (see, e.g., [SS95, p. 70]). Denote by  $E(f)$  this constant.  $E$  is a discontinuous linear functional on  $(\mathcal{Y}, \tau)$ : any compact subset of  $[0, \omega_1)$  is included in some  $[0, \alpha]$  with  $\alpha < \omega_1$ , so  $E$  is not represented by an element of  $\mathcal{Y}^*$ . Now consider the algebraic dual  $\mathcal{Y}^a$  of  $\mathcal{Y}$  endowed with the cylinder  $\sigma$ -algebra  $\mathcal{C}(\mathcal{Y}^a, \mathcal{Y})$  and the Dirac mass  $\nu$  at  $E$ . It is shown in [Dud69, Section 2] that  $\nu^*(\mathcal{Y}^*) = 1$ , so  $\mu := \nu^*$  defines a probability measure on  $(\mathcal{Y}^*, \mathcal{C}(\mathcal{Y}^*, \mathcal{Y}))$  (since  $\mathcal{C}(\mathcal{Y}^*, \mathcal{Y}) \subset \mathcal{C}(\mathcal{Y}^a, \mathcal{Y})|_{\mathcal{Y}^*}$ ). And, applying known results on outer measures, one shows that, for any  $f \in \mathcal{Y}$ ,  $p(f) = E(f)$ . So,  $p$  is not  $\sigma(\mathcal{Y}, \mathcal{Y}^*)$ -lower semi-continuous, which provides the announced counterexample in the l.c.s.s.  $\mathcal{X} = \mathcal{Y}^*$  endowed with the weak-star topology  $\sigma(\mathcal{Y}^*, \mathcal{Y})$  and with the cylinder  $\sigma$ -algebra  $\mathcal{C}(\mathcal{Y}^*, \mathcal{Y})$ .

*Question.* Is there an example where  $p \neq (-s)^*$  on a separable normed space?

Now we give a standard sufficient condition for the equality  $p = (-s)^*$ . A probability measure  $\nu$  on  $(\mathcal{X}, \tau(\mathcal{C}), \mathcal{F}(\mathcal{C}))$  is *convex tight* if, for all  $\alpha > 0$ , there exists a measurable relatively compact convex subset  $K$  of  $\mathcal{X}$  such that  $\nu(K) > 1 - \alpha$ .

**Theorem 8.** *If the distribution of  $X_1$  is convex-tight, then  $p = (-s)^*$ .*

*Proof.* The proof of [Pet18, Theorem 12.4] adapts directly as follows. Let  $\lambda \in \mathcal{X}^*$  and let  $\alpha > 0$ . Since the distribution of  $X_1$  is convex-tight and using Fatou's lemma, there exists a measurable relatively compact convex subset  $K$  of  $\mathcal{X}$  such that

$$\min(p(\lambda) - \alpha, 1/\alpha) \leq \log \mathbb{E}(e^{\lambda(X_1)} \mathbf{1}_K(X_1)) \leq \inf_{n \geq 1} \frac{1}{n} \log \mathbb{E}(e^{n\lambda(\bar{X}_n)} \mathbf{1}_K(\bar{X}_n)).$$

The inequality  $p \leq (-s)^*$  follows from a version of Varadhan's lemma (see [Pet18, Lemma 12.2], which extends to the present setting).  $\square$

*Remark.* According to theorem 7 and proposition 6, an alternate proof consists in showing that  $p$  is  $\sigma(\mathcal{X}^*, \mathcal{X})$ -lower semi-continuous. The following proof is a slight generalization of the one of [Cer07, Lemma 12.1]. First, denote by  $(\tilde{\mathcal{X}}, \tilde{\tau})$  the Hausdorff topological vector space canonically associated to  $(\mathcal{X}, \tau(\mathcal{C}))$ , where  $\tilde{\mathcal{X}} = \mathcal{X}/\overline{\{0\}}$  and  $\tilde{\tau}$  is the quotient topology. Also denote by  $\pi: \mathcal{X} \rightarrow \tilde{\mathcal{X}}$  the canonical projection. By Hahn-Banach theorem,  $\overline{\{0\}} = \{x \in \mathcal{X} \mid \forall \lambda \in \mathcal{X}^* \quad \lambda(x) = 0\}$ , so the topological dual of  $\tilde{\mathcal{X}}$  is isomorph to  $\mathcal{X}^*$ : we will identify both, writing, for  $(\lambda, x) \in \mathcal{X}^* \times \mathcal{X}$ ,  $\lambda(x) = \lambda(\pi(x))$ . Moreover, we have  $\sigma(\mathcal{X}^*, \mathcal{X}) = \sigma(\mathcal{X}^*, \tilde{\mathcal{X}})$ . Since  $p$  is convex, it suffices to show that  $p$  is  $\tau(\mathcal{X}^*, \tilde{\mathcal{X}})$ -lower semi-continuous where  $\tau(\mathcal{X}^*, \tilde{\mathcal{X}})$  is the Mackey topology, that is to say the locally convex topology

admitting for local basis of 0 the sets  $\{\lambda \in X^* \mid \forall \tilde{x} \in \tilde{K} \quad |\lambda(\tilde{x})| < 1\}$  where  $\tilde{K}$  ranges over the absolutely convex  $\sigma(\tilde{\mathcal{X}}, \mathcal{X}^*)$ -compact subsets of  $\tilde{\mathcal{X}}$ .

Let  $\lambda \in \mathcal{X}^*$  and  $\alpha > 0$ . Since the distribution of  $X_1$  is convex-tight, let  $K$  be a measurable relatively  $\tau(\mathcal{C})$ -compact convex subset of  $\mathcal{X}$  such that

$$\log \mathbb{E}(e^{\lambda(X_1)} \mathbf{1}_K(X_1)) \geq \min(p(\lambda), 1/\alpha) - \alpha.$$

The absolutely convex hull  $K_1$  of  $\overline{K}$  is  $\tau(\mathcal{C})$ -compact since it is the image of the compact set  $[0, 1] \times \overline{K} \times (-\overline{K})$  by the continuous map  $(t, x, y) \mapsto (1 - t)x + ty$ . Finally,  $\tilde{K} = \pi(K_1)$  is an absolutely convex  $\tilde{\tau}$ -compact (a fortiori  $\sigma(\tilde{\mathcal{X}}, \mathcal{X}^*)$ -compact) subset of  $\tilde{\mathcal{X}}$ . Considering

$$V := \left\{ \rho \in \mathcal{X}^* \mid \forall \tilde{x} \in \tilde{K} \quad |\rho(\tilde{x}) - \lambda(\tilde{x})| < \alpha \right\},$$

one has, for any  $\rho \in V$  and  $x \in K$ ,  $|\rho(x) - \lambda(x)| < \alpha$  (because  $\pi(K) \subset \tilde{K}$ ), hence

$$p(\rho) \geq \log \mathbb{E}(e^{\rho(X_1)} \mathbf{1}_K(X_1)) \geq \min(p(\lambda), 1/\alpha) - 2\alpha,$$

which yields the  $\tau(\mathcal{X}^*, \tilde{\mathcal{X}})$ -lower semi-continuity of  $p$ .

However, the convex-tightness of the distribution of  $X_1$  is not a necessary condition for the equality  $s = -p^*$  as shown in the following example.

**Example 2.** Let  $\mathcal{X}$  be the l.c.s.s.  $\mathbb{R}^{\mathbb{R}}$  endowed with the product topology  $\tau$  and the cylinder  $\sigma$ -algebra  $\mathcal{F}$ . There is no convex-tight probability measure on  $\mathcal{X}$ . Indeed, every compact subset of  $\mathcal{X}$  is weakly bounded, whereas no nonempty measurable subset is weakly bounded (any measurable subset being in the  $\sigma$ -algebra generated by countably many cylinder sets based on finitely many points): so, the only measurable relatively compact subset is the empty set. But, on  $\mathcal{X}$ , any Dirac mass and even any Gaussian measure has a continuous pressure, of the form  $L(\lambda) + Q(\lambda)$  where  $L$  is a linear functional on  $\mathcal{X}^* = \mathbb{R}^{(\mathbb{R})}$  (automatically  $\sigma(\mathcal{X}^*, \mathcal{X})$ -continuous) and  $Q$  is a  $(\sigma(\mathcal{X}^*, \mathcal{X})$ -continuous) nonnegative symmetric bilinear form on  $\mathcal{X}^*$  (see, *e.g.*, [Bog98, Proposition 2.3.9]).

## References

- [Bog98] Vladimir I. Bogachev. *Gaussian measures*. Mathematical Surveys and Monographs. AMS, 1998.
- [Bou66] N. Bourbaki. *Théorie des ensembles*. Hermann, Paris, 1966.
- [Bou71] N. Bourbaki. *Livre III, Topologie générale*. Hermann & Cie, Edition de 1971.
- [Bou81] N. Bourbaki. *Espaces vectoriels topologiques*. Masson, 1981.
- [BZ79] R. R. Bahadur and S. L. Zabell. Large deviations of the sample mean in general vector spaces. *Ann. Prob.*, 7(4):587–621, 1979.
- [Cer07] R. Cerf. *On Cramér's theory in infinite dimensions*. Panoramas et Synthèses 23. Société Mathématique de France, Paris, 2007.



- [Cra38] H. Cramér. Sur un nouveau théorème-limite de la théorie des probabilités. *Actual. Sci. Indust.*, 736:5–23, 1938.
- [dA94a] A. de Acosta. On large deviations of empirical measures in the  $\tau$ -topology. *J. Appl. Probab.*, 31A:41–47, 1994.
- [dA94b] A. de Acosta. Projective systems in large deviation theory ii: Some applications. In Jørgen Hoffmann-Jørgensen, James Kuelbs, and Michael B. Marcus, editors, *Probability in Banach spaces. 9*, volume 35 of *Progress in Probability*, pages 241–250. Birkhäuser Boston, Inc., Boston, MA, 1994. Papers from the Ninth International Conference held in Sandjberg, August 16–21, 1993.
- [dA97] Alejandro de Acosta. Exponential tightness and projective systems in large deviation theory. In *Festschrift for Lucien Le Cam*, pages 143–156. Springer, New York, 1997.
- [DG87] D. A. Dawson and J. Gärtner. Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics*, 20:247–308, 1987.
- [Dra74] Frank R. Drake. *Set theory — An introduction to large cardinals*, volume 76 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam, 1974.
- [DS89] J.-D. Deuschel and D. W. Stroock. *Large Deviations*. Academic Press, 1989.
- [Dud69] R. M. Dudley. Random linear functionals. *Trans. Amer. Math. Soc.*, 136:1–24, 1969.
- [DZ98] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 1998.
- [FW84] M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems*. Springer-Verlag, 1984.
- [Jec06] Thomas Jech. *Set theory*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2006. The third millennium edition, revised and expanded (corrected 4th printing).
- [Kun68] Kenneth Kunen. *Inaccessibility properties of cardinals*. ProQuest LLC, Ann Arbor, MI, 1968. Thesis (Ph.D.)—Stanford University.
- [LP95] J. T. Lewis and C.-E. Pfister. Thermodynamic probability theory. *Russian Math. Surveys*, 50:279–317, 1995.
- [LPS95] J. T. Lewis, C.-E. Pfister, and W. G. Sullivan. Entropy, concentration of probability and conditional limit theorems. *Markov Processes Relat. Fields*, 1(3):319–386, 1995.
- [Mor67] J.-J. Moreau. *Fonctionnelles convexes*. Séminaire sur les Equations aux Dérivées Partielles, Collège de France, 1966–67.
- [MS48] E. Marczewski and R. Sikorski. Measures in non-separable metric spaces. *Colloquium Math.*, 1:133–139, 1948.

- [Pet18] Pierre Petit. Cramér’s theorem in Banach spaces revisited. In *Séminaire de Probabilités XLIX*, volume 2215 of *Lecture Notes in Math.*, pages 455–473. Springer, Cham, 2018.
- [San57] I. N. Sanov. . *Mat. Sb.*, 42(1):11–44, 1957. English translation: On the probability of large deviations of random variables, *Sel. Transl. Math. Statist. Prob. I*: 213–244, 1961.
- [SS95] Lynn Arthur Steen and J. Arthur Seebach. *Counterexamples in topology*. Dover Publications, 1995.
- [Tal79] Michel Talagrand. Est-ce que  $l^\infty$  est un espace mesurable ? *Bull. Sci. Math. (2)*, 103(3):255–258, 1979.
- [Var66] S. R. S. Varadhan. Asymptotic probabilities and differential equations. *Comm. Pure Applied Math.*, 19:261–286, 1966.
- [VTC87] N. N. Vakhania, V. I. Tarieladze, and S. A. Chobanyan. *Probability distributions on Banach spaces*, volume 14 of *Mathematics and its Applications (Soviet Series)*. D. Reidel Publishing Co., Dordrecht, 1987. Translated from the Russian and with a preface by Wojbor A. Woyczynski.
- [Zăl02] C. Zălinescu. *Convex analysis in general vector spaces*. World Scientific Publishing Co. Inc., River Edge, NJ, 2002.