**Calculus of Variations**

# From dimension free concentration to the Poincaré inequality

**Nathael Gozlan · Cyril Roberto · Paul-Marie Samson**

**Abstract** We prove that a probability measure on an abstract metric space satisfies a non trivial dimension free concentration inequality for the $\ell_2$ metric if and only if it satisfies the Poincaré inequality. Under some additional assumptions, our result extends to convex sets situation.

**Mathematics Subject Classification (1991)** 60E15 · 32F32 · 26D10

## 1 Introduction

Throughout the paper $(\mathcal{X}, d)$ denotes a polish metric space and $\mathcal{P}(\mathcal{X})$ the set of Borel probability measures on $\mathcal{X}$. On the product space $\mathcal{X}^n$, we consider the following $\ell_p$ product distance $d_p$ defined by

N. Gozlan (✉) · P.-M. Samson
Université Paris Est Marne la Vallée, Laboratoire d'Analyse et de Mathématiques
Appliquées (UMR CNRS 8050), 5 bd Descartes, 77454 Marne la Vallée Cedex 2, France
e-mail: nathael.gozlan@univ-mlv.fr

P.-M. Samson
e-mail: paul-marie.samson@univ-mlv.fr

C. Roberto
Université Paris Ouest Nanterre la Défense, MODAL'X, EA 3454,
200 Avenue de la République, 92000 Nanterre, France
e-mail: croberto@math.cnrs.fr

$$d_p(x, y) = \left[ \sum_{i=1}^{n} d^p(x_i, y_i) \right]^{1/p}, \quad x, y \in \mathcal{X}^n, \quad p \geq 1.$$

(Note that the dependence on the dimension $n$ is understood.) If $A$ is a Borel subset of $\mathcal{X}^n$, we define its enlargement $A_{r,p}$ (simply denoted by $A_r$ when $n = 1$), $r \geq 0$ as follows

$$A_{r,p} = \left\{ x \in \mathcal{X}^n; d_p(x, A) \leq r \right\}.$$

Also, in all what follows, $\alpha : \mathbb{R}^+ \to \mathbb{R}^+$ will always be a non increasing function. One will say that $\mu \in \mathcal{P}(\mathcal{X})$ satisfies the *dimension free concentration property* with the concentration profile $\alpha$ and with respect to the $\ell_p$ product structure if

$$\mu^n(A_{r,p}) \geq 1 - \alpha(r), \quad \forall r \geq 0, \tag{1.1}$$

for all $A \subset \mathcal{X}^n$, with $\mu^n(A) \geq 1/2$. For simplicity, we will often say that $\mu$ satisfies the dimension free concentration inequality $\mathbf{CI}_p^\infty(\alpha)$, and, if $\mu$ satisfies (1.1) only for $n = 1$, we will say that $\mu$ satisfies $\mathbf{CI}(\alpha)$. We refer to [28] for an introduction on the notion of concentration of measure.

The general problem considered in this paper is to give a characterization of the class of probability measures satisfying $\mathbf{CI}_p^\infty(\alpha)$. Our main result shows that the class of probability measures satisfying $\mathbf{CI}_2^\infty(\alpha)$, for some non trivial $\alpha$, is always contained in the class of probability measures satisfying the Poincaré inequality. Moreover, these two classes coincide when $\alpha$ is exponential: $\alpha(r) = be^{-ar}$, for some $a, b > 0$.

Before stating this result, let us recall the definition of the Poincaré inequality: one says that $\mu \in \mathcal{P}(\mathcal{X})$ satisfies the *Poincaré inequality* with the constant $\lambda \in \mathbb{R}^+ \cup \{+\infty\}$, if

$$\lambda \operatorname{Var}_\mu(f) \leq \int |\nabla^- f|^2 \, d\mu, \tag{1.2}$$

for all Lipschitz function $f : \mathcal{X} \to \mathbb{R}$, where by definition

$$|\nabla^- f|(x) = \limsup_{y \to x} \frac{[f(y) - f(x)]_-}{d(y, x)}, \quad (\text{with } [X]_- := \max(-X, 0))$$

when $x$ is not isolated in $\mathcal{X}$ (we set $|\nabla^- f|(x) = 0$, when $x$ is isolated in $\mathcal{X}$). By convention $\infty \times 0 = 0$, so that $\lambda = +\infty$ if and only if $\mu$ is a Dirac measure.

*Remark 1.1* Let us make a few comments about the definition of the Poincaré inequality.

(1) Since the right hand side of (1.2) is always finite when $f$ is Lipschitz, it follows in particular that Lipschitz functions always have finite variance when Poincaré inequality holds.

(2) When $(\mathcal{X}, d)$ is a smooth Riemannian manifold equipped with its geodesic distance and $f : \mathcal{X} \to \mathbb{R}$, it is not difficult to check that if $f$ is differentiable at a point $x$, then $|\nabla^- f|(x)$ coincides with the norm of the *vector* $\nabla f(x)$ (belonging to the tangent space at $x$). If $(\mathcal{X}, d) = (B, \| \cdot \|)$ is a Banach space, and $f : B \to \mathbb{R}$ is differentiable at $x \in B$, then $|\nabla^- f|(x)$ is equal to $\|Df(x)\|_*$, the dual norm of the differential $Df(x) \in B^*$ of $f$ at $x$. So (1.2) gives back the usual definitions in a smooth context.

(3) To a Lipschitz function $f$ on $\mathcal{X}$, one can also associate $|\nabla^+ f|$ and $|\nabla f|$, which are defined by replacing $[\cdot]_-$ by $[\cdot]_+$ or by $|\cdot|$ respectively. Since $|\nabla^+ f| = |\nabla^-(-f)|$, replacing $f$ by $-f$, we observe that the Poincaré inequality can be equivalently restated as

$$\lambda \operatorname{Var}_\mu(f) \leq \int |\nabla^+ f|^2 \, d\mu, \tag{1.3}$$

for all Lipschitz function $f : \mathcal{X} \to \mathbb{R}$. Moreover, since $|\nabla f| = \max(|\nabla^- f|; |\nabla^+ f|)$, we see that (1.2) and (1.3) both imply yet another version of Poincaré inequality (considered for instance in [28] or [10]):

$$\lambda \operatorname{Var}_\mu(f) \leq \int |\nabla f|^2 \, d\mu, \tag{1.4}$$

for all Lipschitz function $f : \mathcal{X} \to \mathbb{R}$. That (1.4) also implies (1.2) is not obvious. A proof of this fact can be found in [17, Proposition 5.1] (the result is stated for the logarithmic Sobolev inequality but the same conclusion holds for the Poincaré inequality). The proof relies on a technique, developed in [2], consisting in relaxing the right hand side of (1.4) and yielding to the notion of Cheeger energy.

## 1.1 Main results

Denote by $\overline{\Phi}$ the tail distribution function of the standard Gaussian measure $\gamma(dx) = (2\pi)^{-1/2} e^{-x^2/2} \, dx$ on $\mathbb{R}$ defined by

$$\overline{\Phi}(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{x}^{+\infty} e^{-u^2/2} \, du, \quad x \in \mathbb{R}.$$

The main result of this paper is the following theorem.

**Theorem 1.2** *If $\mu$ satisfies the dimension free concentration property $\mathbf{CI}_2^\infty(\alpha)$, then $\mu$ satisfies the Poincaré inequality* (1.2) *with the constant $\lambda$ defined by*

$$\sqrt{\lambda} = \sup \left\{ \frac{\overline{\Phi}^{-1}(\alpha(r))}{r}; \quad r > 0 \text{ s.t } \alpha(r) \leq 1/2 \right\}.$$

*Moreover, if $\alpha$ is convex decreasing and such that $\alpha(0) = 1/2$, then $\lambda = (2\pi \alpha_+'(0)^2)$, where $\alpha_+'(0) \in [-\infty, 0)$ is the right derivative of $\alpha$ at 0.*

Conversely, it is well known—since the work by Gromov and Milman [25] (see also [1, 10,19,39] for related results)—that a probability measure $\mu$ verifying the Poincaré inequality satisfies a dimension free concentration property with a profile of the form $\alpha(r) = be^{-ar}$, for some $a, b > 0$. We recall this property in the following theorem and refer to Sect. 5 for a proof.

**Theorem 1.3** (Gromov and Milman) *Suppose that $\mu$ satisfies the Poincaré inequality* (1.2) *with a constant $\lambda > 0$, then it satisfies the dimension free concentration property with the profile*

$$\alpha(r) = b \exp(-a\sqrt{\lambda} r), \quad r \geq 0,$$

*where $a, b$ are universal constants.*

Thus, Theorems 1.2 and 1.3 give a full description of the set of probability distributions verifying a dimension free concentration property with a concentration profile $\alpha$ such that $\{r : \alpha(r) < 1/2\} \neq \emptyset$ : this set coincides with the set of probability measures verifying the Poincaré inequality. An immediate corollary of Theorems 1.2 and 1.3 (see Corollary 4.1 below) is that any type of dimension free concentration inequality can always be improved into a dimension free concentration inequality with an exponential profile (up to universal constants). This was already noticed by Talagrand [41]. See Sect. 4.3 for a further discussion.

*Remark 1.4* Let us make some comments on the constant $\lambda$ appearing in Theorem 1.2.

(1) We observe first that $\lambda > 0$ if and only if there is some $r_o > 0$ such that $\alpha(r_o) < 1/2$. In particular, Theorem 1.2 applies in the case of the following "minimal" profile $\alpha = \beta_{a_o,r_o}$, defined as follows

$$\beta_{a_o,r_o}(r) = 1/2, \ \text{if } r < a_o \ \text{ and } \ \beta_{a_o,r_o}(r) = a_o, \ \text{if } r \geq r_o, \qquad (1.5)$$

where $a_o \in [0, 1/2)$, $r_o > 0$. If a probability measure satisfies $\mathbf{CI}_2^\infty(\beta_{a_o,r_o})$, then it satisfies the Poincaré inequality with the constant

$$\sqrt{\lambda_{a_o,r_o}} := \frac{\overline{\Phi}^{-1}(a_o)}{r_o}$$

(2) Then we notice that any non increasing function $\alpha : \mathbb{R}^+ \to \mathbb{R}^+$, with $\alpha(0) = 1/2$, can be written as an infimum of minimal profiles:

$$\alpha = \inf_{r>0} \beta_{\alpha(r),r}.$$

Therefore, the constant $\lambda$ given in Theorem 1.2 is the supremum of the constants $\lambda_{\alpha(r),r}$ $r > 0$ defined above. This shows that the information contained in the concentration profile $\alpha$ is treated pointwise, and that the global behavior of $\alpha$ is not taken into account.

(3) It is well known that the standard Gaussian measure $\gamma$ satisfies the dimension free concentration property with the profile $\alpha = \overline{\Phi}$ (this follows from the Isoperimetric theorem in Gauss space due to Sudakov and Cirelson [40] and Borell [11], see e.g. [28]). Hence, applying Theorem 1.2, we conclude that $\gamma$ satisfies the Poincaré inequality with the constant $\lambda = 1$, which is well known to be optimal (see e.g. [3, Chapter 1]).

(4) Finally we observe that, if the concentration profile $\alpha(r)$ goes to zero too fast when $r \to \infty$, then $\lambda = +\infty$ and so $\mu$ is a Dirac measure. This happens for instance when $\alpha(r) = be^{-ar^k}$, $r \geq 0$, with $k > 2$ and $a, b > 0$.

Theorem 1.2 completes a previous result obtained by the first author [18] (see also [21]), namely that the Gaussian dimension free concentration is characterized by a transport-entropy inequality. We now state this result and start by recalling some notation. The Kantorovich–Rubinstein distance $W_p$, $p \geq 1$, between $\nu, \mu \in \mathcal{P}(\mathcal{X})$ is defined by

$$W_p^p(\nu, \mu) = \inf_{\Pi(\nu,\mu)} \mathbb{E}[d^p(X, Y)],$$

where the infimum runs over the set $\Pi(\nu, \mu)$ of couples of random variables $(X, Y)$ such that $\text{Law}(X) = \mu$ and $\text{Law}(Y) = \nu$. Then, a probability measure is said to satisfy the *p-Talagrand transport-entropy inequality* with constant $C > 0$, if it holds

$$W_p^p(\nu, \mu) \leq C H(\nu|\mu), \quad \forall \nu \in \mathcal{P}(\mathcal{X}), \qquad (1.6)$$

where the relative entropy functional is defined by $H(\nu|\mu) = \int \log \frac{d\nu}{d\mu} \, d\nu$ if $\nu$ is absolutely continuous with respect to $\mu$, and $H(\nu|\mu) = +\infty$ otherwise. Inequalities of this type were introduced by Marton and Talagrand in the nineties [31,43]. We refer to the survey [20] for more informations on this topic.

**Theorem 1.5** [18] *Fix $p \geq 2$ and $C > 0$. Then, a probability measure $\mu$ satisfies the p-Talagrand transport inequality (1.6) with constant $C$ if and only if it satisfies the dimension free concentration inequality $\mathbf{CI}_p^\infty(\alpha)$, with a concentration profile of the form*

$$\alpha(r) = \exp\left(-\frac{1}{C}[r - r_o]_+^p\right), \quad r \geq 0,$$

*for some $r_o \geq 0$ (with $[X]_+ := \max(X, 0)$).*

As we will see, the proofs of Theorem 1.2 and 1.5 are very different. Both make use of probability limit theorems, but not at the same scale: Theorem 1.5 used Sanov's large deviations theorem, whereas Theorem 1.2 is an application of the central limit theorem. Moreover, contrary to what happens in Theorem 1.2 (see item (2) of Remark 1.4), the global behavior of the concentration profile is used in Theorem 1.5.

In view of Theorems 1.2 and 1.5, it is natural to formulate the following general question:

(**Q**) Which functional inequality is equivalent to $\mathbf{CI}_p^\infty(\alpha)$ for a concentration profile of the form

$$\alpha(r) = \exp\left(-a[r - r_o]_+^k\right), \quad r \geq 0,$$

where $a > 0$, $r_o \geq 0$ and $k > 0$ ?

*Remark 1.6* It is easy to see, using the central limit theorem, that for $p \in [1, 2)$ the only probability measures verifying $\mathbf{CI}_p^\infty(\alpha)$, for some $\alpha$ such that $\alpha(r_o) < 1/2$ for at least one $r_o > 0$, are Dirac masses. Thus the question (**Q**) is interesting only for $p \geq 2$.

To summarize, Theorem 1.5 shows that the answer to (**Q**) is the $p$-Talagrand inequality for $k = p$ and $p \geq 2$. Theorem 1.2 shows that the answer is the Poincaré inequality for $p = 2$ and for $k \in (0, 1]$. Moreover point (4) of Remark 1.4 above shows that for $p = 2$, the question is interesting only for $k \in [1; 2]$. The case $k \in (1; 2)$ is still open.

Finally, we mention that some partial results are known for $p = \infty$. Indeed, Bobkov and Houdré [9] characterized the set of probability measures on $\mathbb{R}$ satisfying $\mathbf{CI}_\infty^\infty(\beta_{a_o, r_o})$, with $a_o \in [0, 1/2)$, where $\beta_{a_o, r_o}$ is the minimal concentration profile defined by (1.5). They showed that a probability measure $\mu$ belongs to this class if and only if the map $U_\mu$ defined by

$$U_\mu(x) = F_\mu^{-1}\left(\frac{1}{1 + e^{-x}}\right), \quad x \in \mathbb{R},$$

where $F_\mu(x) = \mu((-\infty, x])$ and $F_\mu^{-1}(p) = \inf\{x \in \mathbb{R}; F_\mu(x) \geq p\}$, $p \in (0, 1)$, satisfies the following inequality on the interval where it is defined:

$$|U_\mu(x) - U_\mu(y)| \leq a + b|x - y|,$$

for some $a, b \geq 0$.

Finally, we mention that, under additional assumptions, Theorem 1.2 extends to convex sets. More precisely, we will prove that, *inter alia*, in geodesic spaces in which the distance function is convex (Busemann spaces), a probability measure satisfies $\mathbf{CI}_2^\infty(\alpha)$ restricted to *convex* sets, with a non trivial profile $\alpha$, if and only if it satisfies the Poincaré inequality restricted to *convex* functions (see Sect. 6 for a precise statement). This generalizes a previous result by Bobkov and Götze that was only valid for probability on the real line [7, Theorem 4.2].

## 1.2 Alternative formulation in terms of observable diameters

It is possible to give an alternative formulation of Theorems 1.2 and 1.3 using the notion of observable diameter introduced by Gromov [24, Chapter 3.1/2]. Recall that, if $(\mathcal{X}, d, \mu)$ is a metric space equipped with a probability measure and $t \in [0, 1]$, the partial diameter of $(\mathcal{X}, d)$ is defined as the infimum of the diameters of all the subsets $A \subset \mathcal{X}$ satisfying

$\mu(A) \geq 1 - t$. It is denoted by Part Diam$(\mathcal{X}, d, \mu, t)$. If $f : \mathcal{X} \to \mathbb{R}$ is some 1-Lipschitz function, we denote by $\mu_f \in \mathcal{P}(\mathbb{R})$ the push forward of $\mu$ under $f$. Then, the *observable diameter* of $(\mathcal{X}, d, \mu)$ is defined as follows

$$\text{Obs Diam}(\mathcal{X}, d, \mu, t) = \sup_{f \ 1-\text{Lip}} \text{Part Diam}(\mathbb{R}, |\cdot|, \mu_f, t) \in \mathbb{R}^+ \cup \{+\infty\}.$$

We define accordingly the observable diameters of $(\mathcal{X}^n, d_2, \mu^n, t)$ for all $n \in \mathbb{N}^*$.

The observable diameters are related to concentration profiles by the following lemma (see e.g. [16, Lemma 2.22]).

**Lemma 1.7** *If $\mu$ satisfies* **CI**$(\alpha)$*, then*

$$\text{Obs Diam}(\mathcal{X}, d, \mu, 2\alpha(r)) \leq 2r,$$

*for all $r \geq 0$ such that $\alpha(r) \leq 1/2$.*
*Conversely, for all $t \in [0, 1/2]$, for all $A \subset \mathcal{X}$, with $\mu(A) \geq 1/2$, it holds*

$$\mu(A_{r(t)}) \geq 1 - t$$

*with $r(t) = \text{Obs Diam}(\mathcal{X}, d, \mu, t)$.*

The following corollary gives an interpretation of the Poincaré inequality in terms of the boundedness of the observable diameters of the sequence of metric probability spaces $(\mathcal{X}^n, d_2, \mu^n)_{n \in \mathbb{N}^*}$.

**Corollary 1.8** *A probability measure $\mu$ on $(\mathcal{X}, d)$ satisfies the Poincaré inequality* (1.2) *with the optimal constant $\lambda$ if and only if for some $t \in (0, 1/2)$*

$$r_\infty(t) := \sup_{n \in \mathbb{N}^*} \text{Obs Diam}(\mathcal{X}^n, d_2, \mu^n, t) < \infty.$$

*Moreover,*

$$\overline{\Phi}^{-1}(t) \leq r_\infty(t)\sqrt{\lambda} \leq a \log\left(\frac{b}{t}\right), \quad \forall t \in (0, 1/2)$$

*where $a > 0$ and $b \geq 1$ are some universal constants.*

1.3 Tools

In this section, we briefly introduce the main tools that will be used in the proof of Theorem 1.2: inf-convolution operators, related to both concentration and to Hamilton–Jacobi equations, and the central limit theorem.

The first main tool in the proof of Theorem 1.2 is a new alternative formulation of concentration of measure **CI**$_p^\infty(\alpha)$ in terms of deviation inequalities for inf-convolution operators that was introduced in [21]. Recall that for all $t > 0$, the infimum convolution operator $f \mapsto Q_t f$ is defined for all $f : \mathcal{X}^n \to \mathbb{R} \cup \{+\infty\}$ bounded from below as follows

$$Q_t f(x) = \inf_{y \in \mathcal{X}^n} \left\{ f(y) + \frac{1}{t^{p-1}} d_p^p(x, y) \right\}, \quad x \in \mathcal{X}^n \tag{1.7}$$

(we should write $Q_t^{p,(n)}$, but we will omit, for simplicity, the superscripts $p$ and $(n)$ in the notation).

In the next proposition, we recall a result from [21] that gives a new way to express concentration of measure (our first main tool).

**Proposition 1.9** *Let $\mu \in \mathcal{P}(\mathcal{X})$; $\mu$ satisfies $\mathbf{CI}_p^\infty(\alpha)$ if and only if for all $n \in \mathbb{N}^*$ and for all measurable function $f : \mathcal{X}^n \to \mathbb{R} \cup \{+\infty\}$ bounded from below and such that $\mu^n(f = +\infty) < 1/2$, it holds*

$$\mu^n(Q_t f > m(f) + r) \leq \alpha(r^{1/p} t^{1-1/p}), \quad \forall r, t > 0, \tag{1.8}$$

*where $m(f) = \inf\{m \in \mathbb{R}; \mu^n(f \leq m) \geq 1/2\}$.*

The second main tool is the well known fact that the function $u : (t, x) \mapsto Q_t f(x)$ is, in some weak sense, solution of the Hamilton–Jacobi equation

$$\frac{\partial u}{\partial t} = -\frac{1}{4} |\nabla_x u|^2.$$

This result is very classical on $\mathbb{R}^k$ (see e.g [14]); extensions to metric spaces were proposed in [2,5,30] or [22]. This will be discussed in the next section.

The third tool is the central limit theorem for triangular arrays of independent random variables (see e.g. [15, p. 530]).

**Theorem 1.10** *For each $n$, let $X_{1,n}, X_{2,n}, \ldots, X_{n,n}$ be independent real random variables and define $T_n = X_{1,n} + \cdots + X_{n,n}$. Assume that $\mathbb{E}[T_n^2] = 1$, $\mathbb{E}[X_{k,n}] = 0$ for all $k \in \{0, \ldots, n\}$, and that the following Lindeberg condition holds, for all $t > 0$*

$$\sum_{k=1}^n \mathbb{E}\left[X_{k,n}^2 \mathbb{1}_{|X_{k,n}|>t}\right] \to 0 \quad as \ n \to +\infty. \tag{1.9}$$

*Then the distribution of $T_n$ converges weakly to the standard normal law.*

We end this introduction with a short roadmap of the paper. In Sect. 2 we make some comments on Theorem 1.2. In particular, we compare Theorem 1.2 to a result by E. Milman on the Poincaré inequality in spaces with non-negative curvature and show, as an immediate consequence of our main result as well as E. Milman's result, that the celebrated KLS conjecture for isotropic log-concave probability measures can be reduced to some universal concentration inequalities (for isotropic log-concave probability measures). In Sect. 3, we recall some properties of the infimum convolution operators that will be used in the proofs. Section 4 is dedicated to the proof of Theorem 1.2 and Sect. 5 to the proof of Theorem 1.3, by means of the Herbst argument (the latter is somehow classical (see e.g. [3,28]) but requires some care due to our general framework). Finally Sect. 6 deals with the case of convex sets and the Poincaré inequality restricted to convex functions.

## 2 Comparison with other results

In this section we collect some remarks and consequences of our main theorem. First we shall compare our result to one of E. Milman, in Riemannian setting. Then, we state the celebrated KLS conjecture and give an equivalent formulation in terms of dimension free concentration property. Finally, we may comment on other type of dimension free concentration properties involving a different definition of enlargement.

2.1 Dimension free concentration vs. non negative curvature.

In Riemannian setting, Theorem 1.2 is reminiscent of the following recent result by E. Milman showing that under non-negative curvature the Poincaré constant of a probability measure can be expressed through very weak concentration properties of the measure [34,35].

We recall that the Minkowski content of a set $A \subset \mathcal{X}$ is defined as follows

$$\mu^+(A) := \liminf_{r \to 0} \frac{\mu(A_r) - \mu(A)}{r}.$$

**Theorem 2.1** (Milman [35]) *Let $\mu(dx) = e^{-V(x)} dx$ be an absolutely continuous probability measure on a smooth complete separable Riemannian manifold $M$ equipped with its geodesic distance $d$. Suppose that $V : M \to \mathbb{R}$ is a function of class $\mathcal{C}^2$ such that*

$$\text{Ric} + \text{Hess } V \geq 0,$$

*and that $\mu$ satisfies the following concentration of measure inequality*

$$\mu(A_r) \geq 1 - \alpha(r), \quad \forall r \geq 0,$$

*with $\alpha : [0, \infty) \to [0, 1/2]$ such that $\alpha(r_o) < 1/2$, for some $r_o > 0$. Then $\mu$ satisfies the following Cheeger's inequality*

$$\mu^+(A) \geq D \min(\mu(A); 1 - \mu(A)), \quad \forall A \subset M,$$

*with*

$$D = \sup \left\{ \frac{\Psi(\alpha(r))}{r}; \quad r > 0 \text{ s.t } \alpha(r) < 1/2 \right\},$$

*where $\Psi : [0, 1/2)$ is some universal function.*

We recall that Cheeger's inequality with the constant $D$ implies the Poincaré inequality (1.2) with the constant $\lambda = D^2/4$ [13,33]. In our result the non-negative curvature assumption of Milman's result is replaced by the assumption that the concentration is dimension free.

*Remark 2.2* Notice that, if $M$ has non-negative Ricci curvature and $\mu(dx) = \frac{1}{|K|} \mathbf{1}_K(x) dx$ is the normalized restriction of the Riemanian volume to a geodesically convex set $K$, then Milman [36] also obtains that

$$D = \sup \left\{ \frac{1 - 2\alpha(r)}{r}; \quad r > 0 \right\}.$$

This bound is optimal (see [36]).

2.2 A remark on the KLS conjecture

In this section, $\mathcal{X} = \mathbb{R}^k$ is always equipped with its standard Euclidean norm $| \cdot |$.

Let us recall the celebrated conjecture by Kannan et al. [26]. Recall first that a probability measure $\mu$ on $\mathbb{R}^k$ is isotropic if $\int x \, \mu(dx) = 0$ and $\int x_i x_j \, \mu(dx) = \delta_{ij}$ for all $1 \leq i, j \leq n$. It is log-concave if it has a density of the form $e^{-V}$, where $V : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ is a convex function.

**Conjecture 2.3** (Kannan et al. [26]) *There is a universal constant $D > 0$ such that for all $k \in \mathbb{N}^*$, any log-concave and isotropic probability measure $\mu$ on $\mathbb{R}^k$ satisfies the following Cheeger inequality*

$$\mu^+(A) \geq D \min(\mu(A); 1 - \mu(A)), \quad \forall A \subset \mathbb{R}^k.$$

*Equivalently, there is a universal constant $\lambda > 0$ such that for all $k \in \mathbb{N}^*$, any log-concave and isotropic probability measure $\mu$ on $\mathbb{R}^k$ satisfies the following Poincaré inequality*

$$\lambda \text{Var}_\mu(f) \leq \int |\nabla f|^2 \, d\mu,$$

*for all $f : \mathbb{R}^k \to \mathbb{R}$ Lipschitz.*

Note that, in the statement above, the converse implication from Poincaré to Cheeger inequality is due to Buser [12] and Ledoux [27,29] and is in fact true more generally on Riemannian manifolds with non-negative Ricci curvature [29].

According to E. Milman's Theorem 2.1, the above conjecture can be reduced to a statement about universal concentration inequalities for log-concave isotropic probabilities.

**Corollary 2.4** *The KLS conjecture is equivalent to the following statement. There exists $r_o > 0, a_o \in [0, 1/2)$ such that for any $m \in \mathbb{N}^*$, any log-concave and isotropic probability $\nu$ on $\mathbb{R}^m$ satisfies*

$$\nu(A + r_o B_2) \geq 1 - a_o, \quad \forall A \subset \mathbb{R}^m \text{ s.t. } \nu(A) \geq 1/2, \qquad (2.1)$$

*where $B_2$ is the Euclidean unit ball of $\mathbb{R}^m$.*

This corollary follows immediately from Theorem 2.1. Below, we propose an alternative proof based on our main result (Theorem 1.2).

*Proof of Corollary 2.4* According to Theorem 1.3, it is clear that the KLS conjecture implies uniform exponential concentration estimates for isotropic log-concave probability measures.

Conversely, let $\mu$ be isotropic and log-concave on $\mathbb{R}^k$. For all $n \in \mathbb{N}^*$, the probability $\mu^n$ is still isotropic and log-concave on $\left(\mathbb{R}^k\right)^n$. So applying (2.1) to $\nu = \mu^n$ on $\left(\mathbb{R}^k\right)^n$, for all $n \in \mathbb{N}^*$, we conclude that $\mu$ satisfies $\mathbf{CI}_2^\infty(\beta_{a_o,r_o})$, where the concentration profile $\beta_{a_o,r_o}$ is defined by (1.5). According to Theorem 1.2, we conclude that $\mu$ satisfies Poincaré inequality with the constant $\lambda = \left(\overline{\Phi}^{-1}(a_o)/r_o\right)^2$. Since this holds for any isotropic log-concave probability measure in any dimension, this ends the proof. $\qquad \square$

### 2.3 Euclidean vs. Talagrand type enlargements

Theorem 1.2 improves a preceding result by the first author [18] where a stronger form of exponential dimension free concentration, introduced by Talagrand [41,42], was shown to be equivalent to a transport-entropy inequality which was known to be also equivalent to the Poincaré inequality. These equivalences, together with our main result, will allow us to prove that the two notions of exponential dimension free concentration properties are actually equivalent (see Theorem 2.9 below).

In order to present these equivalences, we need some notation and definition.

Given $n \in \mathbb{N}^*$ and $A \subset \mathcal{X}^n$, consider the following family of enlargements of $A$:

$$\widetilde{A}_{a,r} = \left\{ x \in \mathcal{X}^n; \exists y \in A \text{ s.t } \sum_{i=1}^n \theta(ad(x_i, y_i)) \leq r \right\}, \quad a > 0, \quad r \geq 0$$

where $\theta(t) = t^2$, if $t \in [0, 1]$ and $\theta(t) = 2t - 1$, if $t \geq 1$.

In the next definition, we recall the dimension free concentration property introduced by Talagrand.

**Definition 2.5** A probability measure $\mu$ on $\mathcal{X}$ is said to satisfy the Talagrand exponential type dimension free concentration inequality with constants $a, b \geq 0$ if for all $n \in \mathbb{N}^*$, for all $A \subset \mathcal{X}^n$ with $\mu^n(A) \geq 1/2$, it holds

$$\mu^n(\widetilde{A}_{a,r}) \geq 1 - be^{-r}, \quad \forall r \geq 0. \qquad (2.2)$$

*Remark 2.6* Using elementary algebra, one can compare the Talagrand concentration inequality (2.2) with the dimension free concentration inequality (1.1) under investigation in this paper. More precisely we may prove that the former is stronger than the latter. Indeed, since $t \mapsto \theta(\sqrt{t})$ is concave and vanishes at 0, it is thus sub-additive. In turn, the following inequality holds

$$\sum_{i=1}^{n} \theta(ad(x_i, y_i)) \geq \theta\left(\sqrt{\sum_{i=1}^{n} a^2 d^2(x_i, y_i)}\right) = \theta(ad_2(x, y)), \quad \forall x, y \in \mathcal{X}^n.$$

Therefore,

$$\widetilde{A}_{a,\theta(ar)} \subset A_{r,2},$$

and so, if $\mu$ satisfies the Talagrand concentration inequality (2.2), then it obviously verifies the dimension free concentration inequality with the profile $\alpha(u) = be^{-\theta(au)} \leq ebe^{-2au}$, $u \geq 0$.

The following theorem summarizes the known links between Talagrand exponential type dimension free concentration and the Poincaré inequality.

**Theorem 2.7** *Let $\mu$ be a probability measure on $\mathcal{X}$. The following statements are equivalent*

*(1) $\mu$ satisfies the Poincaré inequality (1.2) with a constant $\lambda > 0$.*
*(2) $\mu$ satisfies the Talagrand exponential type dimension free concentration inequality (2.2) with constants $a, b > 0$.*
*(3) $\mu$ satisfies the following transport-entropy inequality*

$$\inf_{(X,Y)\in\Pi(\mu,\nu)} \mathbb{E}\left(\theta(Cd(X, Y))\right) \leq H(\nu|\mu), \quad \forall \nu \in \mathcal{P}(\mathcal{X}),$$

*for some constant $C > 0$, (recall that $\Pi(\mu, \nu)$ and the relative entropy $H(\nu|\mu)$ are defined before Theorem 1.5).*

*Moreover the constants above are related as follows :*

*(1) $\Rightarrow$ (2) with $a = \kappa\sqrt{\lambda}$ and $b = 1$, for some universal constant $\kappa$.*

*(2) $\Rightarrow$ (3) with $C = a$.*

*(3) $\Rightarrow$ (1) with $\lambda = 2C^2$.*

Let us make some bibliographical comments about the different implications in Theorem 2.7. The implication (1) $\Rightarrow$ (2) is due to Bobkov and Ledoux [10], the implication (2) $\Rightarrow$ (3) is due to the first author [18, Theorem 5.1], and the implication (3) $\Rightarrow$ (1) is due to Maurey [32] or Otto and Villani [37]. The equivalence between (1) and (3) was first proved by Bobkov et al. [6].

*Remark 2.8* It is worth noting that the implication (2) $\Rightarrow$ (3) follows from Theorem 1.5 for $p = 2$ by a change of metric argument. Namely, suppose that $\mu$ satisfies the concentration property (2) of Theorem 2.7, for some $a > 0$, and define $\tilde{d}(x, y) = \sqrt{\theta(ad(x, y))}$ for all $x, y \in \mathcal{X}$. It is not difficult to check that the function $\theta^{1/2}$ is subadditive, and therefore $\tilde{d}$ defines a new distance on $\mathcal{X}$. The $\ell_2$ extension of $\tilde{d}$ to the product $\mathcal{X}^n$ is

$$\tilde{d}_2(x, y) = \left[\sum_{i=1}^{n} \theta(ad(x_i, y_i))\right]^{1/2}, \quad x, y \in \mathcal{X}^n,$$

and it holds

$$\widetilde{A}_{a,r} = \left\{x \in \mathcal{X}^n; \tilde{d}_2(x, A) \leq \sqrt{r}\right\}, \quad \forall A \subset \mathcal{X}^n.$$

Therefore, statement (2) can be restated by saying that $\mu$ satisfies $\mathbf{CI}_2^\infty(\alpha)$ (with respect to the distance $\tilde{d}$) with the Gaussian concentration profile $\alpha(r) = be^{-r^2}$. Applying Theorem 1.5, we conclude that $\mu$ satisfies the 2-Talagrand transport entropy inequality with the constant 1 with respect to the distance $\tilde{d}$, which is exactly (3) with $C = a$.

An immediate consequence of Theorem 1.2 and of Bobkov–Ledoux theorem (1) $\Rightarrow$ (2) above is the following result showing the equivalence between the two forms of dimension free exponential concentration.

**Theorem 2.9** *Let $\mu$ be a probability measure on $\mathcal{X}$. The following are equivalent.*

*(1) The probability measure $\mu$ satisfies the Talagrand exponential type dimension free concentration inequality (2.2) with constants a and b.*
*(2) The probability measure $\mu$ satisfies $\mathbf{CI}_2^\infty(\alpha)$ with a profile $\alpha(u) = b'e^{-a'u}$, $u \geq 0$.*

*Moreover, the constants are related as follows: (1) $\Rightarrow$ (2) with $a' = 2a$ and $b' = eb$, and (2) $\Rightarrow$ (1) with $a = \kappa a'/\sqrt{\log(2b')}$ (for some universal constant $\kappa$) and $b = 1$.*

We do not know if there exists a direct proof of the implication (2) $\Rightarrow$ (1).

*Proof* We have already proved that (1) implies (2) (see Remark 2.6). Let us prove the converse. According to Theorem 1.2 we conclude from (2) that $\mu$ satisfies the Poincaré inequality with a constant $\lambda \geq \left(\frac{\overline{\Phi}^{-1}(\alpha(u))}{u}\right)^2$, for all $u$ such that $\alpha(u) < 1/2$. A classical inequality gives $\overline{\Phi}(t) \leq \frac{1}{2}e^{-t^2/2}$, $t \geq 0$. Therefore, $\overline{\Phi}^{-1}(t) \geq 2\sqrt{-\log(2t)}$, for all $t \in (0, 1/2)$. Hence, taking $u = 2\log(2b')/a'$ [which guarantees that $\alpha(u) = 1/(4b')$] yields to $\lambda \geq \frac{a'^2}{\log(2b')}$. According to the implication (1) $\Rightarrow$ (2) in Theorem 2.7 we conclude that $\mu$ satisfies Talagrand concentration inequality (2.2) with $a = \kappa a'/\sqrt{\log(2b')}$. $\square$

## 3 Some properties of the inf-convolution operators

In this short (technical) section, we recall some properties of the inf-convolution operators related to Hamilton–Jacobi equations and to the concentration of measure, in the setting of metric spaces (recall that $(\mathcal{X}, d)$ is a complete separable metric space).

### 3.1 Inf-convolution operators and Hamilton–Jacobi equations

In this paragraph, we shall only consider the case $p = 2$. We do this restriction for simplicity and also since only that particular case will be used in the proof of Theorem 1.2 (in the next section). However we mention that most of the results of this section can be extended to any $p > 1$.

The following proposition collects two basic observations about the operators $Q_t$, $t > 0$.

**Proposition 3.1** *Let $h : \mathcal{X} \to \mathbb{R}$ be a Lipschitz function. Then*

*(i) for all $x \in \mathcal{X}$, $Q_t h(x) \to h(x)$, when $t \to 0^+$;*

*(ii) for all $\nu \in \mathcal{P}(\mathcal{X})$,*

$$\limsup_{t \to 0^+} \frac{1}{t} \int h(x) - Q_t h(x) \, \nu(dx) \leq \frac{1}{4} \int |\nabla^- h(x)|^2 \, \nu(dx). \qquad (3.1)$$

Before giving the proof of Proposition 3.1, let us complete the picture by recalling the following theorem of [2,23] (improving preceding results of [5,30]). This result will not be used in the sequel.

**Theorem 3.2** *Let $h$ be a bounded function on a polish metric space $\mathcal{X}$. Then $(t, x) \mapsto Q_t h(x)$ satisfies the following Hamilton–Jacobi (in)equation*

$$\frac{d}{dt_+} Q_t h(x) \leq -\frac{1}{4} |\nabla Q_t h|^2(x), \quad \forall t > 0, \ \forall x \in \mathcal{X} \qquad (3.2)$$

*where $d/dt_+$ stands for the right derivative, and $|\nabla h|(x) = \limsup_{y \to x} \frac{|h(y) - h(x)|}{d(y,x)}$. Moreover, if the space $\mathcal{X}$ is geodesic (i.e. for all $x, y \in \mathcal{X}$ there exists at least one curve $(z_t)_{t \in [0,1]}$ such that $z_0 = x, z_1 = y$ and $d(z_s, z_t) = |t - s| d(x, y)$) then (3.2) holds with equality.*

Observe that, strangely, the two inequalities (3.1) and (3.2) go in the opposite direction. This suggests that, at $t = 0$, there should be equality, at least for some class of functions.

*Proof of Proposition 3.1* Let $L > 0$ be a Lipschitz constant of $h$; since $Q_t h \leq h$ one has

$$Q_t h(x) = \inf_{y \in B(x, 2Lt)} \left\{ h(y) + \frac{1}{t} d^2(x, y) \right\}.$$

(Namely, if $d(x, y) > 2Lt$, it holds $h(y) - h(x) + \frac{1}{t} d^2(x, y) \geq \left( -L + \frac{1}{t} d(x, y) \right) d(x, y) > 0$.)

Hence

$$0 \leq \frac{h(x) - Q_t h(x)}{t} = \sup_{y \in B(x, 2Lt)} \left\{ \frac{h(x) - h(y)}{t} - \frac{d^2(x, y)}{t^2} \right\}$$

$$\leq \sup_{y \in B(x, 2Lt)} \left\{ \frac{[h(x) - h(y)]_+}{d(x, y)} \frac{d(x, y)}{t} - \frac{d^2(x, y)}{t^2} \right\}$$

$$\leq \sup_{r \in \mathbb{R}} \left\{ \sup_{y \in B(x, 2Lt)} \frac{[h(x) - h(y)]_+}{d(x, y)} r - r^2 \right\} = \frac{1}{4} \sup_{y \in B(x, 2Lt)} \frac{[h(x) - h(y)]_+^2}{d^2(x, y)}.$$

We conclude from this that $0 \leq (h - Q_t h)/t \leq L^2/4$. This implies in particular that $Q_t h \to h$ when $t \to 0$. Taking the lim sup when $t \to 0^+$ gives

$$\limsup_{t \to 0^+} \frac{h(x) - Q_t h(x)}{t} \leq \frac{1}{4} |\nabla^- h(x)|^2. \qquad (3.3)$$

Inequality (3.1) follows from (3.3) using Fatou's lemma in its lim sup version. The application of Fatou's lemma is justified by the fact that the family of functions $\{(h - Q_t h)/t\}_{t > 0}$ is uniformly bounded. $\qquad \square$

*Remark 3.3* The proof of (3.3) can also be found in [45, Theorem 22.46] (see also [23, Proposition A.3], [2,5,30]).

3.2 Inf-convolution operators and concentration of measure

In this subsection we briefly recall the short proof of Proposition 1.9 for the sake of completeness [proposition that relates the infimum convolution operator and the concentration property $\mathbf{CI}_p^\infty(\alpha)$].

*Proof* Let $f : \mathcal{X}^n \to \mathbb{R} \cup \{+\infty\}$ be a function bounded from below with $\mu^n(f = +\infty) < 1/2$. By definition of $m(f)$, it holds $\mu^n(f \le m(f)) \ge 1/2$. Define $A = \{f \le m(f)\}$. If $\mu$ satisfies the dimension free concentration property $\mathbf{CI}_p^\infty(\alpha)$, since by definition of $m(f)$, $\mu^n(A) \ge 1/2$, one has $\mu^n(\mathcal{X}^n \setminus A_{u,p}) \le \alpha(u)$, for all $u \ge 0$. Then, observe that

$$Q_t f(x) \le m(f) + \frac{1}{t^{p-1}} d_p^p(x, A), \quad \forall x \in \mathcal{X}^n.$$

Hence $\{Q_t f > m(f) + r\} \subset \{d_p(\cdot, A) > r^{1/p} t^{1-1/p}\} = \mathcal{X} \setminus A_{r^{1/p} t^{1-1/p}, p}$, which proves (1.8).

To prove the converse, take a Borel set $A \subset \mathcal{X}^n$ such that $\mu^n(A) \ge 1/2$ and consider the function $i_A$ equals to 0 on $A$ and $+\infty$ on $A^c$. For this function, $Q_t i_A = d_p^p(x, A)/t^{p-1}$ and one can choose $m(i_A) = 0$. Applying (1.8) gives the concentration property $\mathbf{CI}_p^\infty(\alpha)$. $\square$

## 4 Poincaré inequality and concentration of measure

This section is dedicated to the proof of our main result Theorem 1.2, and to Corollary 1.8. Moreover, we shall explain how Theorem 1.2 can be used to improve any non-trivial dimension free concentration property to an exponential one.

4.1 From dimension free concentration to the Poincaré inequality: proof of Theorem 1.2

*Proof of Theorem 1.2* Let $h : \mathcal{X} \to \mathbb{R}$ be a bounded Lipschitz function such that $\int h \, d\mu = 0$. For all $n \in \mathbb{N}^*$, define $f_n : \mathcal{X}^n \to \mathbb{R}^+$ by

$$f_n(x) = h(x_1) + \cdots + h(x_n), \quad \forall x = (x_1, \ldots, x_n) \in \mathcal{X}^n.$$

Our aim is to apply the central limit theorem. Applying (1.8) to $f_n$ with $t = 1/\sqrt{n}$ and $r = \sqrt{n}u$, for some $u > 0$, we easily arrive at

$$\mu^n \Bigg( \frac{1}{\sqrt{n}\sigma_n} \sum_{i=1}^n \Big[ Q_{1/\sqrt{n}} h(x_i) - \mu(Q_{1/\sqrt{n}} h) \Big] > \frac{1}{\sigma_n \sqrt{n}} m(f_n)$$

$$+ \frac{\sqrt{n}}{\sigma_n} \mu \Big( h - Q_{1/\sqrt{n}} h \Big) + \frac{u}{\sigma_n} \Bigg) \le \alpha(\sqrt{u}), \tag{4.1}$$

where $\sigma_n^2 = \mathrm{Var}_\mu(Q_{1/\sqrt{n}} h)$ and $m(f_n)$ is a median of $f_n$ under $\mu^n$, that is to say any number $m \in \mathbb{R}$ such that $\mu^n(f \ge m) \ge 1/2$ and $\mu^n(f_n \le m) \ge 1/2$.

We deal with each term of (4.1) separately. According to point (i) of Proposition 3.1, we observe that $\sigma_n \to \sigma = \sqrt{\mathrm{Var}_\mu(h)}$, when $n$ goes to $\infty$ and according to Point (ii) of Proposition 3.1, that

$$\limsup_{n \to +\infty} \sqrt{n} \mu \Big( h - Q_{1/\sqrt{n}} h \Big) \le \frac{1}{4} \int |\nabla^- h|^2 \, d\mu.$$

On the other hand, let $m_n = m(f_n)/(\sqrt{n}\sigma)$. According to the central limit theorem (Theorem 1.10) the law of the random variables $T_n = f_n/(\sqrt{n}\sigma)$ under $\mu^n$ converges weakly to the

standard Gaussian. Since weak convergence implies the convergence of quantiles as soon as the limit distribution has a continuous repartition function (see for instance [44, Lemma 21.2]), we have in particular, $m_n \to 0$ as $n \to \infty$.

Now, fix $\varepsilon > 0$. According to the above observations, (4.1) yields, for any $u > 0$ and any $n$ sufficiently large,

$$\mu^n \left( \frac{1}{\sqrt{n}\sigma_n} \sum_{i=1}^n \left[ Q_{1/\sqrt{n}}h(x_i) - \mu(Q_{1/\sqrt{n}}h) \right] > \frac{1+\varepsilon}{\sigma} \left( \varepsilon + \frac{1}{4} \int |\nabla^- h|^2 \, d\mu + u \right) \right)$$
$$\leq \alpha(\sqrt{u}). \tag{4.2}$$

In order to apply (again) the central limit theorem (Theorem 1.10), introduce the following random variables

$$\widetilde{T}_n(x) = \sum_{i=1}^n \frac{1}{\sqrt{n}\sigma_n} \left[ Q_{1/\sqrt{n}}h(x_i) - \mu(Q_{1/\sqrt{n}}h) \right]$$

under $\mu^n$. Since $h$ is bounded $Q_{1/\sqrt{n}}h$ is uniformly bounded in $n$, and by Lebesgue theorem, as $n$ goes to $\infty$, we see that the Lindeberg condition (1.9) is verified:

$$\int \frac{1}{\sigma_n^2} \left( Q_{1/\sqrt{n}}h - \mu(Q_{1/\sqrt{n}}h) \right)^2 \mathbb{1}_{\left| Q_{1/\sqrt{n}}h - \mu(Q_{1/\sqrt{n}}h) \right| > t\sqrt{n}\sigma_n} \, d\mu \to 0 \quad \text{as } n \to +\infty.$$

Therefore, letting $n$ go to $+\infty$, and $\varepsilon$ to 0, by continuity of $\overline{\Phi}$, the inequality (4.2) provides

$$\overline{\Phi} \left( \frac{\int |\nabla^- h|^2 \, d\mu}{4\sigma} + \frac{u}{\sigma} \right) \leq \alpha(\sqrt{u}), \quad \forall u \geq 0.$$

Let $u \geq 0$ be such that $\alpha(\sqrt{u}) < 1/2$ and $k(u) := \overline{\Phi}^{-1}(\alpha(\sqrt{u})) > 0$. We easily get from the latter inequality that

$$k(u)\sqrt{\mathrm{Var}_\mu(h)} \leq u + \frac{1}{4} \int |\nabla^- h|^2 \, d\mu.$$

Replacing $h$ by $sh$, $s > 0$, and taking the infimum, we arrive at

$$k(u)\sqrt{\mathrm{Var}_\mu(h)} \leq \inf_{s>0} \left\{ \frac{u}{s} + \frac{s}{4} \int |\nabla^- h|^2 \, d\mu \right\} = \sqrt{u \int |\nabla^- h|^2 \, d\mu}.$$

Optimizing over $u$, one concludes that Poincaré inequality (1.2) is satisfied with the constant $\lambda$ announced in Theorem 1.2.

Now let $h : \mathcal{X} \to \mathbb{R}$ be an unbounded Lipschitz function. Consider the sequence of functions $h_n = (h \vee -n) \wedge n$, $n \in \mathbb{N}^*$ converging pointwise to $h$. For all $n \in \mathbb{N}^*$, $h_n$ is bounded and Lipschitz, and it is not difficult to check that

$$|\nabla^- h_n|(x) = 0 \text{ if } x \in \{h \leq -n\} \cup \{h > n\} \quad \text{and} \quad |\nabla^- h_n|(x)$$
$$= |\nabla^- h_n|(x) \text{ if } x \in \{-n < h \leq n\}. \tag{4.3}$$

In particular, the sequence $|\nabla^- h_n|$ converges monotonically to $|\nabla^- h|$. Applying Fatou's lemma and the monotone convergence theorem, we obtain

$$\lambda \iint (h(x) - h(y))^2 \, \mu(dx)\mu(dy) \leq \lambda \liminf_{n \to \infty} \mathrm{Var}_\mu(h_n)$$
$$\leq \liminf_{n \to \infty} \int |\nabla^- h_n|^2 \, d\mu = \int |\nabla^- h|^2 \, d\mu,$$

which completes the proof of Theorem 1.2. □

### 4.2 Poincaré inequality and boundedness of observable diameters of product probability spaces

In this section we prove Corollary 1.8.

*Proof of Corollary 1.8* Assume first that $\mu$ satisfies the Poincaré inequality (1.2) with the optimal constant $\lambda$. Then according to Theorem 1.3, $\mu$ satisfies $\mathbf{CI}_2^\infty(\alpha)$ with the concentration profile $\alpha(r) = be^{-\sqrt{\lambda}r}$, where $a, b$ are universal constants ($b \geq 1/2$). According to the first part of Lemma 1.7 [applied to the metric probability space $(\mathcal{X}^n, d_2, \mu^n)$], it follows that for all $n \in \mathbb{N}^*$, $\mathrm{Obs\,Diam}(\mathcal{X}^n, d_2, \mu^n, t) \leq 2\frac{\log(4b/t)}{a\sqrt{\lambda}}$, for all $t \leq 1$ and thus

$$r_\infty(t)\sqrt{\lambda} \leq a' \log(b'/t), \quad \forall t \leq 1$$

for some universal constants $a', b'$.

Conversely, assume that $0 < r_\infty(t_o) < \infty$ for some $t_o \in (0, 1/2)$. According to the second part of Lemma 1.7, $\mu$ satisfies $\mathbf{CI}_2^\infty(\beta_{t_o, r_\infty(t_o)})$, where the minimal profiles $\beta$ are defined in (1.5). According to Theorem 1.2, it follows that $\mu$ satisfies the Poincaré inequality with an optimal constant $\lambda > 0$ such that

$$\sqrt{\lambda} r_\infty(t_o) \geq \overline{\Phi}^{-1}(t_o).$$

According to the first step, we conclude that $r_\infty(t) < \infty$ for all $t \leq 1$, and so the inequality above is true for all $t \in (0, 1/2)$. □

### 4.3 Self improvement of dimension free concentration inequalities

The next result shows that a non-trivial dimension free concentration inequality can always be upgraded into an inequality with an exponential decay. This observation goes back to Talagrand [41, Proposition 5.1].

**Corollary 4.1** *If $\mu$ satisfies $\mathbf{CI}_2^\infty(\alpha)$ with a profile $\alpha$ such that $\alpha(r_o) < 1/2$ for some $r_o$, then it satisfies $\mathbf{CI}_2^\infty$ with an exponential concentration. More explicitly, it satisfies the dimension free concentration property with the profile $\tilde{\alpha}(r) = be^{-a\sqrt{\lambda}r}$, where $a, b$ are universal constants and*

$$\sqrt{\lambda} = \sup\left\{ \frac{\overline{\Phi}^{-1}(\alpha(r))}{r}; \quad r > 0 \ s.t \ \alpha(r) < 1/2 \right\}.$$

This result is an immediate corollary of Theorems 1.2 and 1.3.

In [41] this result was stated and proved only for probability measures on $\mathbb{R}$. We thank E. Milman for mentioning to us that the argument was in fact more general. For the sake of completeness, we extend below Talagrand's argument in a very general abstract framework. For a future use, we only assume that the dimension free concentration property holds on a good subclass of sets. We refer to the proof of Proposition 6.4 where this refinement will be used (the subclass of sets being the class of convex sets).

**Proposition 4.2** *Let $(\mathcal{X}, d)$ be a complete separable metric space, $p \geq 1$ and for all $n \in \mathbb{N}^*$ let $\mathcal{A}_n$ be a class of Borel sets in $\mathcal{X}^n$ satisfying the following conditions:*

*(i) For all $n \in \mathbb{N}^*$ and $r \geq 0$, if $A \in \mathcal{A}_n$ then $A_{r,p} \in \mathcal{A}_n$.*
*(ii) If $A \in \mathcal{A}_m$, then $A^n \in \mathcal{A}_{nm}$.*

*Suppose that a Borel probability measure $\mu$ on $\mathcal{X}$ satisfies the following dimension free concentration property: there exists $r_o > 0$, $a_o \in [0, 1/2)$ such that for all $n \in \mathbb{N}^*$,*

$$\mu^n(A_{r_o,p}) \geq 1 - a_o, \quad \forall A \in \mathcal{A}_n \text{ s.t. } \mu^n(A) \geq 1/2.$$

*Then, for any $\gamma \in (-\log(1 - a_o)/\log(2), 1)$, there exists $c \in [1/2, 1)$ depending only on $\gamma$ and $a_o$ such that for all $n \in \mathbb{N}^*$,*

$$\mu^n(A_{r,p}) \geq 1 - \frac{1-c}{\gamma}\gamma^{r/r_o}, \quad \forall r \geq 0, \quad \forall A \in \mathcal{A}_n \text{ s.t. } \mu^n(A) \geq c.$$

Note in particular that in the case $p = 2$ (and $\mathcal{A}_n$ the class of all Borel sets), we recover the conclusion of Corollary 4.1 with slightly less accurate constants.

*Proof* Given $A \in \mathcal{A}_1$, it holds $(A^n)_{r_o,p} \subset (A_{r_o})^n$ and, according to (i) and (ii), both sets belong to $\mathcal{A}_n$. Therefore, if $\mu(A) \geq (1/2)^{1/n}$, it holds $\mu(A_{r_o}) \geq (1 - a_o)^{1/n}$. Now, let $A \in \mathcal{A}_1$ be such that $\mu(A) \geq 1/2$ and let $n_A$ be the greatest integer $n \in \mathbb{N}^*$ such that $\mu(A) \geq (1/2)^{1/n}$. By definition of $n_A$,

$$\frac{\log(2)}{\log(1/\mu(A))} - 1 < n_A \leq \frac{\log(2)}{\log(1/\mu(A))}.$$

According to what precedes,

$$\mu(A_{r_o}^c) \leq 1 - (1 - a_o)^{1/n_A} \leq 1 - \exp\left(\frac{\log(1 - a_o)\log(1/\mu(A))}{\log(2) - \log(1/\mu(A))}\right).$$

The function $\varphi(u) = \exp\left(\frac{\log(1-a_o)\log(1/u)}{\log(2)-\log(1/u)}\right)$ satisfies

$$\varphi(u) = 1 - \frac{\log(1 - a_o)}{\log(2)}(u - 1) + o(u - 1),$$

when $u \to 1$. So $\frac{1-\varphi(\mu(A))}{1-\mu(A)} \to -\frac{\log(1-a_o)}{\log(2)} \in (0, 1)$, when $\mu(A) \to 1$. Therefore, if $\gamma$ is any number in the interval $(-\frac{\log(1-a_o)}{\log(2)}, 1)$, there exists $c > 1/2$ (depending only on $\gamma$) such that for all $A \in \mathcal{A}_1$ with $\mu(A) \geq c$ it holds

$$\mu\left(A_{r_o}^c\right) \leq \gamma\mu(A^c).$$

Iterating (which is possible thanks to (i) and the easy to check property $(A_{r_1,p})_{r_2,p} \subset A_{r_1+r_2,p}$) yields

$$\mu\left(A_{kr_o}^c\right) \leq \gamma^k\mu(A^c), \quad \forall k \in \mathbb{N}^*.$$

It follows easily that for all $u \geq 0$,

$$\mu\left(A_u^c\right) \leq ((1 - c)/\gamma)\gamma^{u/r_o}.$$

Applying the argument above to the product measure $\mu^p$, $p \in \mathbb{N}^*$, gives the conclusion. $\square$

## 5 From Poincaré inequality to exponential concentration: proof of Theorem 1.3

In this section, we give a proof of Theorem 1.3. Its conclusion is very classical in, say, a Euclidean setting. But to deal with the general metric space framework requires some additional technical ingredients that we present now.

### 5.1 Technical preparation

In order to regularize Lipschitz functions, we shall introduce an approximate sup-convolution operator. More precisely, for all $\varepsilon > 0$, and for all function $f : \mathcal{X}^n \to \mathbb{R} \cup \{-\infty\}$, we define the (approximate) sup-convolution operator by

$$R_\varepsilon f(x) := \sup_{y \in \mathcal{X}^n} \left\{ f(y) - \sqrt{\varepsilon + d_2^2(x, y)} \right\}, \quad x \in \mathcal{X}^n. \tag{5.1}$$

The next lemma collects some useful properties about $R_\varepsilon$.

**Lemma 5.1** *Let $f : \mathcal{X}^n \to \mathbb{R} \cup \{-\infty\}$ be a function taking at least one finite value and $\varepsilon > 0$. Then*

*(i) If $R_\varepsilon f(x_o) < \infty$ for some $x_o \in \mathcal{X}^n$, then $R_\varepsilon f$ is finite everywhere and is 1-Lipschitz with respect to $d_2$. Moreover, it holds*

$$\sum_{i=1}^n |\nabla_i^- R_\varepsilon f(x)|^2 \leq 1, \quad \forall x \in \mathcal{X}^n.$$

*(ii) If $f$ is 1-Lipschitz, then for all $x \in \mathcal{X}^n$,*

$$f(x) - \sqrt{\varepsilon} \leq R_\varepsilon f(x) \leq f(x).$$

*(iii) If $(\mathcal{X}, d)$ is a Banach space and $f$ is convex, then $R_\varepsilon f$ is also convex.*

*Proof* We fix $\varepsilon > 0$.

Point (iii) follows easily from the fact that $R_\varepsilon f$ is a supremum of convex functions, since by a simple change of variable

$$R_\varepsilon f(x) = \sup_{z \in \mathcal{X}^n} \left\{ f(x - z) - \sqrt{\varepsilon + \|z\|_2^2} \right\}, \quad x \in \mathcal{X}^n.$$

Point (ii) is also easy. Indeed, the first inequality follows by choosing $y = x$. For the second inequality, observe that, since $f$ is 1-Lipschitz,

$$R_\varepsilon f(x) - f(x) = \sup_{y \in \mathcal{X}^n} \left\{ f(y) - f(x) - \sqrt{\varepsilon + d_2^2(x, y)} \right\} \leq \sup_{r \geq 0} \left\{ r - \sqrt{\varepsilon + r^2} \right\} = 0.$$

Now we turn to the proof of Point (i).

The first part of the statement follows from the fact that $R_\varepsilon f$ is a supremum of 1-Lipschitz functions. To prove the other part, we need to fix some notation. For $x = (x_1, \ldots, x_n), z = (z_1, \ldots, z_n) \in \mathcal{X}^n$ and $i \in \{1, 2, \ldots, n\}$, we set

$$\bar{x}^i z = (x_1, \ldots, x_{i-1}, z_i, x_{i+1}, \ldots, x_n).$$

Also, we set $\theta(u) := \sqrt{\varepsilon + u}$, $u \in \mathbb{R}$, so that $R_\varepsilon f(x) = \sup_y \{f(y) - \theta(d_2^2(x, y))\}$ (observe that $\theta$ is concave).

Fix $x = (x_1, \ldots, x_n) \in \mathcal{X}^n$ and a parameter $\eta \in (0, 1)$ that will be chosen later on and consider $z = (z_1, \ldots, z_n) \in \mathcal{X}^n$ such that $z_i \neq x_i$ for all $i \in \{1, \ldots, n\}$. We assume that $R_\varepsilon f$ is everywhere finite. Hence, since $\eta \min_{1 \leq i \leq n} d(x_i, z_i) > 0$, there exists $\hat{y} = \hat{y}(x, z, \eta)$ such that

$$R_\varepsilon f(x) \leq f(\hat{y}) - \sqrt{\varepsilon + d_2^2(x, \hat{y})} + \eta \min_{1 \leq i \leq n} d(x_i, z_i).$$

As a consequence, using that $\theta$ is concave, for all $1 \leq i \leq n$, we have

$$\left[ R_\varepsilon f(\bar{x}^i z_i) - R_\varepsilon f(x) \right]_- = \left[ R_\varepsilon f(x) - R_\varepsilon f(\bar{x}^i z_i) \right]_+ \leq \left[ \theta(d_2^2(x^i z_i, \hat{y})) - \theta(d_2^2(x, \hat{y})) \right]_+$$
$$+ \eta \min_{1 \leq i \leq n} d(x_i, z_i)$$
$$\leq \left[ d_2^2(x^i z_i, \hat{y}) - d_2^2(x, \hat{y}) \right]_+ \theta'(d_2^2(x, \hat{y})) + \eta \min_{1 \leq i \leq n} d(x_i, z_i)$$
$$= \left[ d(\hat{y}_i, z_i) - d(\hat{y}_i, x_i) \right]_+ (d(\hat{y}_i, z_i) + d(\hat{y}_i, x_i)) \theta'(d_2^2(x, \hat{y}))$$
$$+ \eta \min_{1 \leq i \leq n} d(x_i, z_i)$$
$$\leq 2d(z_i, x_i) d(\hat{y}_i, z_i) \theta'(d_2^2(x, \hat{y})) + \eta \min_{1 \leq i \leq n} d(x_i, z_i)$$

where, in the last line we used that the positive part $\left[ d(\hat{y}_i, z_i) - d(\hat{y}_i, x_i) \right]_+$ guarantees that $d(\hat{y}_i, x_i) \leq d(\hat{y}_i, z_i)$ and the triangular inequality.

Using the Cauchy-Schwarz inequality, it follows for any $\delta > 0$ that

$$\frac{\left[ R_\varepsilon f(\bar{x}^i z_i) - R_\varepsilon f(x) \right]_-^2}{d^2(x_i, z_i)} \leq \left( 2d(\hat{y}_i, z_i) \theta'(d_2^2(x, \hat{y})) + \eta \right)^2 \leq (1+\delta) 4 d^2(\hat{y}_i, z_i) \theta'(d_2^2(x, \hat{y}))^2$$
$$+ \left( 1 + \frac{1}{\delta} \right) \eta^2.$$

Therefore, using the triangular and the Cauchy-Schwarz inequalities, we get for any $\delta > 0$,

$$\sum_{i=1}^n \frac{\left[ R_\varepsilon f(\bar{x}^i z_i) - R_\varepsilon f(x) \right]_-^2}{d^2(x_i, z_i)} \leq (1+\delta) 4 d_2^2(\hat{y}, z) \theta' \left( d_2^2(x, \hat{y}) \right)^2 + n \left( 1 + \frac{1}{\delta} \right) \eta^2$$
$$\leq (1+\delta) 4 (d_2(\hat{y}, x) + d_2(x, z))^2 \theta' \left( d_2^2(x, \hat{y}) \right)^2 + n \left( 1 + \frac{1}{\delta} \right) \eta^2$$
$$\leq (1+\delta)^2 4 d_2^2(\hat{y}, x) \theta' \left( d_2^2(x, \hat{y}) \right)^2$$
$$+ (1+\delta) \left( 1 + \frac{1}{\delta} \right) 4 d_2^2(x, z) \theta' \left( d_2^2(x, \hat{y}) \right)^2 + n \left( 1 + \frac{1}{\delta} \right) \eta^2$$
$$\leq (1+\delta)^2 + (1+\delta) \left( 1 + \frac{1}{\delta} \right) \frac{d_2^2(x, z)^2}{\varepsilon} + n \left( 1 + \frac{1}{\delta} \right) \eta^2,$$

using that $4u^2 \theta'(u^2)^2 \leq 1$ and $\theta'(u^2)^2 \leq 1/(4\varepsilon)$. The expected result follows at once taking the limits $z_i \to x_i$, and then $\eta \to 0$ and $\delta \to 0$. This ends the proof of the lemma.    □

The next ingredient in the proof of Theorem 1.3 is the so-called Herbst argument that we explain now. We recall the following notation: for any $f : \mathcal{X}^n \to \mathbb{R}$, set

$$m(f) = \inf \left\{ m \in \mathbb{R}; \mu^n(f \leq m) \geq 1/2 \right\}.$$

**Proposition 5.2** *Assume that $\mu$ satisfies the Poincaré inequality (1.2) with a constant $\lambda > 0$. Then, there exist universal constants $a, b > 0$ such that then for all $n \in \mathbb{N}^*$, it holds*

$$\mu^n\left(f > m(f) + \sqrt{\frac{2}{\lambda}} + r\right) \leq b \exp\left(-a\sqrt{\lambda}r\right), \quad \forall r \geq 0$$

*for all function $f : \mathcal{X}^n \to \mathbb{R}$ such that*

$$\sum_{i=1}^{n} |\nabla_i^- f|^2(x) \leq 1, \quad \forall x \in \mathcal{X}^n. \tag{5.2}$$

*The same conclusion holds if the function $f$ satisfies*

$$\sum_{i=1}^{n} |\nabla_i^+ f|^2(x) \leq 1, \quad \forall x \in \mathcal{X}^n. \tag{5.3}$$

*Proof* It is well known that the Poincaré inequality tensorizes properly (see e.g. [28, proposition 5.6]). Indeed, recall that for all $n \in \mathbb{N}^*$ and for all product probability measure $\mu^n$

$$\lambda \mathrm{Var}_{\mu^n}(g) \leq \int \sum_{i=1}^{n} \mathrm{Var}_\mu(g_i)\, \mu^n(dx),$$

where $g_i(x_i) := g(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n)$, with $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ fixed. Therefore, if $\mu$ satisfies (1.2), then the product probability measure $\mu^n$ satisfies

$$\lambda \mathrm{Var}_{\mu^n}(g) \leq \int \sum_{i=1}^{n} |\nabla_i^- g|^2(x)\, \mu^n(dx), \tag{5.4}$$

for all function $g : \mathcal{X}^n \to \mathbb{R}$ that is Lipschitz in each coordinate.

Let $f : \mathcal{X}^n \to \mathbb{R}$ be bounded and such that (5.2) holds, and define $Z(s) = \log \int e^{sf}\, d\mu^n$, for all $s \geq 0$. Applying (5.4) to $g = e^{sf}$ (which is still bounded and Lipschitz) and using (5.2) yields easily to

$$\lambda \left[\int e^{2sf} d\mu^n - \left(\int e^{sf}\, d\mu^n\right)^2\right] \leq s^2 \int e^{2sf}\, d\mu^n.$$

Thus

$$\log(1 - s^2/\lambda) + Z(2s) \leq 2Z(s), \quad \forall 0 \leq s \leq \sqrt{\lambda}.$$

According to Hölder's inequality, the function $Z$ is convex. Therefore,

$$Z(2s) \geq Z(s) + Z'(s)s.$$

As a result,

$$\log(1 - s^2/\lambda) + Z'(s)s \leq Z(s), \quad \forall 0 \leq s \leq \sqrt{\lambda},$$

and so

$$\frac{d}{ds}\left(\frac{Z(s)}{s}\right) \leq -\frac{\log(1 - s^2/\lambda)}{s^2}, \quad \forall 0 < s \leq \sqrt{\lambda}.$$

Since $Z(s)/s \to \int f \, d\mu^n$ when $s \to 0$, we conclude that

$$\int e^{s(f - \int f \, d\mu^n)} d\mu^n \leq \exp\left(\frac{s}{\sqrt{\lambda}} \int_0^{s/\sqrt{\lambda}} \frac{-\log(1-v^2)}{v^2} \, dv\right), \quad \forall s \leq \sqrt{\lambda}.$$

Taking $s = \sqrt{\lambda}/2$, we easily get, by the (exponential) Chebychev Inequality,

$$\mu^n\left(f - \int f \, d\mu^n > r\right) \leq b e^{-\frac{\sqrt{\lambda}}{2}r}, \quad \forall r \geq 0, \tag{5.5}$$

with $b = \exp\left(\frac{1}{2} \int_0^{1/2} \frac{-\log(1-v^2)}{v^2} \, dv\right)$.

Now, since the probability measure $\mu^n$ satisfies the Poincaré inequality with constant $\lambda > 0$, (5.2) implies that

$$\lambda \operatorname{Var}_{\mu^n}(f) \leq 1.$$

Therefore, by Markov's inequality, it holds, for all $\varepsilon > 0$

$$\mu^n\left(f \leq \int f \, d\mu^n - \sqrt{\frac{2+\varepsilon}{\lambda}}\right) \leq \frac{\lambda}{2+\varepsilon} \operatorname{Var}_{\mu^n}(f) < \frac{1}{2}.$$

Hence, according to this definition of $m(f)$, it follows easily that

$$m(f) \geq \int f \, d\mu^n - \sqrt{\frac{2}{\lambda}}.$$

This inequality together with (5.5) provides the expected deviation inequality

$$\mu^n\left(f > m(f) + \sqrt{\frac{2}{\lambda}} + r\right) \leq b e^{-\frac{\sqrt{\lambda}}{2}r}, \quad \forall r \geq 0. \tag{5.6}$$

Now suppose that $f : \mathcal{X}^n \to \mathbb{R}$ is Lipschitz but not bounded. Consider the function $f_n = (f \vee -n) \wedge n, n \in \mathbb{N}^*$. Applying (4.3) componentwise, we see that $|\nabla_i^- f_n| \leq |\nabla_i^- f|$ for all $i \in \{1, 2, \ldots, n\}$. Therefore, $f_n$ satisfies (5.2). Applying (5.6) to $f_n$, and letting $n$ go to $\infty$, one sees easily that (5.6) still holds for the unbounded Lipschitz function $f$.

Exactly the same proof works under the condition (5.3) [using the equivalent inequality (1.3)]                                                                                                              □

## 5.2 Proof of Theorem 1.3

Thanks to the previous section (Sect. 5.1), we are now in position to prove Theorem 1.3.

*Proof of Theorem 1.3* Let $A$ be a measurable subset of $\mathcal{X}^n$ of measure $\mu^n(A) \geq 1/2$ and define for all $\varepsilon > 0$, $f_{A,\varepsilon}(x) := \sqrt{\varepsilon + d_2^2(x, A)}, x \in \mathcal{X}^n$. By definition of $R_\varepsilon$ given in (5.1), it holds $f_{A,\varepsilon} = -R_\varepsilon i_A$, where $i_A : \mathcal{X}^n \to \{-\infty; 0\}$ is the function defined by $i_A(x) = 0$, if $x \in A$ and $i_A(x) = -\infty$ otherwise. According to Point $(ii)$ of Lemma 5.1, we see that $f_{A,\varepsilon}$ satisfies Condition (5.3) of Proposition 5.2. Hence, observing that $m(f_{A,\varepsilon}) = \sqrt{\varepsilon}$, it follows from Proposition 5.2 that

$$\mu^n\left(\sqrt{\varepsilon + d_2^2(x, A)} > \sqrt{\varepsilon} + \sqrt{\frac{2}{\lambda}} + r\right) \leq b \exp\left(-a\sqrt{\lambda}r\right), \quad \forall r \geq 0.$$

Letting $\varepsilon$ go to 0 yields (after a change of variable) to,

$$1 - \mu^n(A_{r,2}) \leq b' \exp(-a\sqrt{\lambda}r), \quad \forall r \geq 0,$$

with $b' = \max(b, 1/2)e^{\sqrt{2a}}$. The proof of Theorem 1.3 is complete. $\qquad\square$

## 6 Poincaré inequality for convex functions and dimension free convex concentration property.

In this final section, we investigate the links between the dimension free concentration property $\mathbf{CI}_2^\infty(\alpha)$ restricted to convex sets, and the Poincaré inequality restricted to the class of convex functions (see below for a precise definition)

### 6.1 Convexity and convex concentration on geodesic spaces

To deal with convexity properties, we shall assume throughout the section that the metric space $(\mathcal{X}, d)$ is geodesic. This means that any two points $x$ and $y$ of $\mathcal{X}$ can be connected by at least one constant-speed continuous curve in $\mathcal{X}$: i.e. for all $x, y \in \mathcal{X}$, there exists $(x_t)_{t\in[0,1]}$ in $\mathcal{X}$ satisfying $x_0 = x$, $x_1 = y$, and for all $s, t \in [0, 1]$, $d(x_s, x_t) = |t - s|d(x_0, x_1)$. Such a curve is called a (constant-speed) geodesic joining $x$ to $y$.

**Definition 6.1** Let $(\mathcal{X}, d)$ be a geodesic space.

(i) A set $A \subset \mathcal{X}$ is said to be convex if for all $x_0, x_1 \in A$ and all geodesics $(x_t)_{t\in[0,1]}$ joining $x_0$ to $x_1$, it holds $x_t \in A$, $\forall t \in [0, 1]$.
(ii) A function $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ is said to be convex if for all geodesics $(x_t)_{t\in[0,1]}$ in $\mathcal{X}$,

$$f(x_t) \leq (1 - t)f(x_0) + tf(x_1).$$

Accordingly, one will say that $\mu \in \mathcal{P}(\mathcal{X})$ satisfies the dimension free convex concentration property with the concentration profile $\alpha$ and with respect to the $\ell_2$ product structure (in short $\mathbf{CCI}_2^\infty(\alpha)$), if for all convex subset $A \subset \mathcal{X}^n$, with $\mu^n(A) \geq 1/2$,

$$\mu^n(A_{r,2}) \geq 1 - \alpha(r), \quad \forall r \geq 0. \tag{6.1}$$

As for $\mathbf{CI}_2^\infty(\alpha)$ in Proposition 1.9, the convex concentration property can be characterized using the inf-convolution operator $Q_t$.

**Proposition 6.2** *Let $\mu \in \mathcal{P}(\mathcal{X})$; $\mu$ satisfies $\mathbf{CCI}_2^\infty(\alpha)$ if and only if for all $n \in \mathbb{N}^*$ and for all convex function $f : \mathcal{X}^n \to \mathbb{R} \cup \{+\infty\}$ bounded from below and such that $\mu^n(f = +\infty) < 1/2$, it holds*

$$\mu^n(Q_t f > m(f) + r) \leq \alpha(\sqrt{tr}), \quad \forall r, t > 0, \tag{6.2}$$

*where $m(f) = \inf\{m \in \mathbb{R}; \mu^n(f \leq m(f)) \geq 1/2\}$.*

*Proof* By the following two observations the proof of the above proposition (that we omit) becomes essentially identical to the one of Proposition 1.9 in Sect. 3.2: (a) By definition, if $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ is convex then any level set $\{f \leq r\}$, $r \in \mathbb{R}$, is convex; (b) for any convex set $A$, the function $i_A$, that equals 0 on $A$ and $+\infty$ on $\mathcal{X} \setminus A$, is convex. $\qquad\square$

For technical reasons, we will have to distinguish between the following two versions of the Poincaré inequality in restriction to convex functions (in short *convex Poincaré*):

$$\lambda \mathrm{Var}_\mu(f) \leq \int |\nabla^- f|^2 \, d\mu, \quad \forall f \text{ Lipschitz and convex} \tag{6.3}$$

and

$$\lambda \mathrm{Var}_{\mu}(f) \leq \int |\nabla^+ f|^2 \, d\mu, \quad \forall f \text{ Lipschitz and convex.} \tag{6.4}$$

The argument used in Point (3) of Remark 1.1 to prove the equivalence between (1.2) and (1.3) in the usual setting does not work anymore since if $f$ is convex, $-f$ is no more convex. However, if $(\mathcal{X}, d)$ is finite dimensional Banach space or a smooth Riemannian manifold and if one assumes that $\mu$ is absolutely continuous with respect to the Lebesgue (or the volume) measure, by Rademacher's theorem, both gradients in the definition of (6.3) and (6.4) are equal, except on a set of $\mu$-measure 0, which, in turn, guarantees that (6.3) and (6.4) are equivalent.

### 6.2 Dimension free convex concentration implies the convex Poincaré inequality (6.3)

Starting from the functional characterization of $\mathbf{CCI}_2^{\infty}(\alpha)$ stated in Proposition 6.2 and following the lines of the proof of Theorem 1.2, we shall obtain the first main theorem of this section (a counterpart of Theorem 1.2 in the convex situation).

**Theorem 6.3** *If $\mu$ satisfies the dimension free convex concentration property $\mathbf{CCI}_2^{\infty}(\alpha)$, then $\mu$ satisfies the convex Poincaré inequality*

$$\lambda \mathrm{Var}_{\mu}(f) \leq \int |\nabla^- f|^2 \, d\mu,$$

*for all locally Lipschitz[1] and convex function with finite variance with the constant $\lambda$ defined by*

$$\sqrt{\lambda} = \sup \left\{ \frac{\overline{\Phi}^{-1}(\alpha(r))}{r}; \quad r > 0 \text{ s.t } \alpha(r) \leq 1/2 \right\}.$$

*If moreover $\int_0^{\infty} r\alpha(r) \, dr < \infty$ then the variance of all convex and Lipschitz functions is finite and thus $\mu$ verifies (6.3).*

*Proof* The proof is very similar to the proof of Theorem 1.2, but one needs to take care of the technical difficulties coming from the restriction to convex sets/functions. We give here only the main lines and omit most of the details. Also, we may use the notation of the proof of Theorem 1.2.

Let $h : \mathcal{X} \to \mathbb{R}$ be a locally Lipschitz convex function such that $\int h \, d\mu = 0$ and $\int h^2 \, d\mu < +\infty$. For all $n \in \mathbb{N}^*$, define $f_n : \mathcal{X}^n \to \mathbb{R}^+$ by

$$f_n(x) = h(x_1) + \cdots + h(x_n), \quad \forall x = (x_1, \ldots, x_n) \in \mathcal{X}^n.$$

We first observe that $f_n$ is convex. Indeed, this point is an easy consequence of the fact that $(x_t)_{t \in [0,1]}$ is a geodesic in $(\mathcal{X}^n, d_2)$ if and only if for all $i \in \{1 \ldots, n\}$, $(x_{i,t})_{t \in [0,1]}$ are geodesics in $(\mathcal{X}, d)$ (where $x_t = (x_{1,t}, \ldots, x_{n,t})$). Therefore we may apply Proposition 6.2 to the function $f_n$.

For the next step of the proof, we need some analogue of Proposition 3.1 for convex locally Lipschitz functions $h \colon \mathcal{X} \to \mathbb{R}$ (not necessarily globally Lipschitz). From the definition of the convexity property of $h$, one may easily check that for all $x, y \in \mathcal{X}$

$$h(x) - h(y) \leq d(x, y)|\nabla^- h(x)|.$$

---

[1] By *locally Lipschitz*, we mean Lipschitz on every ball of finite radius.

It follows that for all $t > 0$,

$$0 \leq \frac{h(x) - Q_t h(x)}{t} = \sup_{y \in \mathcal{X}} \left\{ \frac{h(x) - h(y)}{t} - \frac{d^2(x, y)}{t^2} \right\}$$

$$\leq \sup_{y \in \mathcal{X}} \left\{ |\nabla^- h(x)| \frac{d(x, y)}{t} - \frac{d^2(x, y)}{t^2} \right\}$$

$$= \sup_{r \in \mathbb{R}} \left\{ |\nabla^- h(x)| r - r^2 \right\} \leq \frac{1}{4} |\nabla^- h(x)|^2 < \infty.$$

This implies in particular that $Q_t h(x) \to h(x)$ for all $x \in \mathcal{X}$. Moreover, it holds

$$\sqrt{n} \mu \left( h - Q_{1/\sqrt{n}} h \right) \leq \frac{1}{4} \int |\nabla^- h|^2 \, d\mu, \quad \forall n \in \mathbb{N}^*.$$

Let us first assume that $h$ is bounded from below. Then it holds, $|Q_{1/\sqrt{n}} h| \leq |h| + |\inf h|$ for all $n \in \mathbb{N}^*$. Since $\int h^2 \, d\mu < \infty$, applying the dominated convergence theorem yields to

$$\lim_{n \to +\infty} \sigma_n^2 = \lim_{n \to +\infty} \text{Var}_\mu(Q_{1/\sqrt{n}} h) = \text{Var}_\mu(h) = \sigma^2.$$

Since $\int h^2 \, d\mu < +\infty$, we may show that any median of $m_n$ of $f_n/(\sqrt{n}\sigma)$ tends to 0 as $n$ goes to $+\infty$. As a consequence, (4.2) holds for $n$ sufficiently large. The rest of the proof is identical to the one of Theorem 1.2: we apply the Central Limit Theorem (Theorem 1.10), observing that the Lindeberg condition holds since $|Q_{1/\sqrt{n}} h| \leq |\inf h| + |h|$ for all $n \in \mathbb{N}^*$, and $\int h^2 \, d\mu < \infty$. This proves the claim in the case where $h$ is assumed bounded from below. To show that the Poincaré inequality still holds if $h$ is not bounded from below, one considers the approximation sequence $h_n = h \vee -n$, $n \in \mathbb{N}^*$ (note that $h_n$ is convex) and one follows the same line of reasoning as in the end of the proof of Theorem 1.2.

Finally, to complete the proof, observe that if $f : \mathcal{X} \to \mathbb{R}$ is convex and 1-Lipschitz, then $A = \{ f \leq m(f) \}$ is a convex set with $\mu(A) \geq 1/2$, and since $A_{r,2} \subset \{ f \leq m + r \}$, we conclude from $\mathbf{CCI}_2(\alpha)$ that

$$\mu(f > m(f) + r) \leq \alpha(r), \quad \forall r \geq 0.$$

Since $\int_0^\infty r\alpha(r) \, dr < \infty$, an integration by part shows that $\int [f - m(f)]_+^2 \, d\mu < \infty$. Therefore, if $f$ is convex, 1-Lipschitz and bounded from below, then $f$ has a finite variance and so applying the first part of the proof, we conclude that $\text{Var}_\mu(f) \leq 1/\lambda$. Using the same truncation as above, we see that this inequality extends to all convex and 1-Lipschitz functions. This complete the proof. $\square$

6.3 Convex Poincaré inequality implies exponential dimension free convex concentration

To get the converse, namely to get a counterpart of Gromov–Milman theorem for convex sets [that the convex Poincaré inequalities (6.3) or (6.4) imply the dimension free concentration property $\mathbf{CCI}_2^\infty(\alpha)$] with an exponential profile, we need to add the structural assumption that the underlying metric space is of Busemann type.

Recall that $(\mathcal{X}, d)$ is said to be a *Busemann's space* if the distance $d : \mathcal{X}^2 \to \mathbb{R}^+$ is a convex function (of both variables). Banach spaces are obvious examples of such spaces. Another important class of Busemann's spaces are complete connected Riemannian manifolds with *non positive* sectional curvature. We refer to [4,8,38] for more informations on the topic and a proof of this statement.

Let us mention some elementary properties of Busemann's spaces:

(B1) if $\mathcal{X}$ is a Busemann's space, any two points of $\mathcal{X}$ are joined by a unique constant speed geodesic;

(B2) for any convex subset $A \subset \mathcal{X}$, the function $\mathcal{X} \ni x \mapsto d(x, A)$ is convex;

(B3) if $(\mathcal{X}, d)$ is a Busemann's space, then $(\mathcal{X}^n, d_2)$ is a Busemann's space.

We are now in position to state our second main theorem (a counterpart of Gromov–Milman's Theorem 1.3).

**Theorem 6.4** *Let $(\mathcal{X}, d)$ be a geodesic space and $\mu$ be a probability measure. Assume one of the following hypotheses: either*

(a) *$\mathcal{X}$ is a Busemann's space and $\mu$ satisfies the convex Poincaré inequality (6.4) with a constant $\lambda > 0$.*

*or*

(b) *$(\mathcal{X}, \|\cdot\|)$ is a Banach space (with $\|x - y\| = d(x, y)$, $x, y \in \mathcal{X}$) and $\mu$ satisfies the convex Poincaré inequality (6.3) with a constant $\lambda > 0$;*

*Then $\mu$ satisfies the dimension free convex concentration property $\mathbf{CCI}_2^\infty(\alpha)$ with the profile*

$$\alpha(r) = b \exp(-a\sqrt{\lambda} r), \quad r \geq 0,$$

*where $a, b$ are universal constants.*

As a direct corollary of Theorems 6.4 and 6.3, when $\mathcal{X}$ is a Banach space, we conclude that, similarly to the general case, the set of probability measures $\mu$ satisfying the dimension free convex concentration property $\mathbf{CCI}_2^\infty(\alpha)$ with exponential profile coincides with the set of probability measures verifying the convex Poincaré inequality (6.3).

The remaining of the section is dedicated to the proof of Theorem 6.4. There is a technical obstacle for applying Herbst argument as in Proposition 5.2. Namely, since the class of convex functions is not stable under truncation from above (if $f$ is convex and $a \in \mathbb{R}$, $\min(f; a)$ is in general not convex) it is delicate to deal safely with $\int e^{sf} d\mu$, $s \geq 0$. Although a method based on $p$-th moments could perhaps be considered in replacement,

we use instead an argument based on Corollary 4.2: we first show that under (6.3) and (6.4), $\mu$ verifies the dimension free convex concentration property $\mathbf{CCI}_2^\infty(\alpha)$ with some polynomial concentration profile and then we upgrade it into an exponential one using Corollary 4.2.

*Proof* We start as in the proof of Proposition 1.3 by noticing that the convex Poincaré inequalities (6.3) and (6.4) tensorize properly. Therefore, if $\mu$ verifies one of the convex Poincaré inequalities, then for all $n \in \mathbb{N}^*$,

$$\lambda \mathrm{Var}_{\mu^n}(f) \leq \int \sum_{i=1}^n \left| \nabla_i^{+/-} f \right|^2 d\mu^n,$$

for all Lipschitz and convex function $f : \mathcal{X}^n \to \mathbb{R}$. In particular, if $f$ is 1-Lipschitz and such that $\sum_{i=1}^n |\nabla_i^{+/-} f|^2 \leq 1$, then $\mathrm{Var}_{\mu^n}(f) \leq 1/\lambda$. Applying Jensen's inequality, we see that

$$1/\lambda \geq \mathrm{Var}_{\mu^n}(f) = \inf_{a \in \mathbb{R}} \int (f - a)^2 d\mu^n \geq \left( \inf_{a \in \mathbb{R}} \int |f - a| d\mu^n \right)^2 = \left( \int |f - m(f)| d\mu^n \right)^2$$

Thus, it follows immediately from Markov's inequality that

$$\mu^n (f > m(f) + r) \leq \frac{1}{\sqrt{\lambda} r}, \quad \forall r > 0. \tag{6.5}$$

Let us first assume that Assumption (a) holds. Observe that the function $\mathcal{X}^n \to \mathbb{R}$ : $x \mapsto f_A(x) := d_2(x, A)$ is convex and 1-Lipschitz when $A \subset \mathcal{X}^n$ is a convex set [this follows from properties (B2) and (B3) above]. Moreover, using Point (i) of Lemma 5.1 and arguing as in the proof of Theorem 1.3, we see that $f_{A,\varepsilon} := \sqrt{\varepsilon + f_A^2}$ satisfies the condition $\sum_{i=1}^n |\nabla_i^+ f_A|^2 \le 1$. Since the function $u \mapsto \sqrt{\varepsilon + u^2}$ is convex and increasing, the function $f_{A,\varepsilon}$ is itself convex. Applying (6.5) to $f_{A,\varepsilon}$ and letting $\varepsilon \to 0$, we get

$$\mu^n(A_{r,2}) \ge 1 - \frac{1}{\sqrt{\lambda} r}, \tag{6.6}$$

for all convex set $A$.

In particular, taking $r_o = 4/\sqrt{\lambda}$ we are in the framework of Corollary 4.2, with $a_o = 1/4$ and $\mathcal{A}_n$ being the class of convex sets of $\mathcal{X}^n$. (Note that Assumptions (i) and (ii) are well verified by $\mathcal{A}_n$: a product of convex sets is always convex and properties (B2) and (B3) above show that the enlargement of a convex set remains convex). We thus conclude that there are universal constants $0 < \gamma < 1$ and $1/2 \le c < 1$ such that, for all $n \in \mathbb{N}^*$, for all convex set $A$ such that $\mu^n(A) \ge c$

$$\mu^n(A_{r,2}) \ge 1 - be^{-a\sqrt{\lambda} r}, \quad \forall r \ge 0,$$

with $b = (1 - c)/\gamma$ and $a = -\log(\gamma)/4$. Now, if $\mu^n(A) \ge 1/2$, then applying (6.6) we see that $\mu^n(A_{r_1,2}) \ge c$ for $r_1 = \frac{1}{\sqrt{\lambda}(1-c)}$. We easily conclude from this that $\mu$ verifies $\mathbf{CCI}_2^\infty(\alpha)$ with $\alpha(r) = b'e^{-a\sqrt{\lambda} r}$, for some other universal constant $b'$.

Assume now that Assumption (b) holds. Let $A \subset \mathcal{X}^n$ be a convex subset. The function $f_A$ defined above is 1-Lipschitz and convex. Therefore, according to point (i) and (iii) Lemma 5.1, the function $R_\varepsilon f_A$, where $R_\varepsilon$ is the operator defined in (5.1), is 1 Lipschitz, *convex* and satisfies $\sum_{i=1}^n |\nabla_i^- R_\varepsilon f_A|^2 \le 1$. Applying (6.5), we conclude that

$$\mu^n (R_\varepsilon f_A > m(R_\varepsilon f_A) + r) \le \frac{1}{\sqrt{\lambda} r}, \quad \forall r > 0.$$

Moreover, according to point (ii) of Lemma 5.1, it holds $f_A - \sqrt{\varepsilon} \le R_\varepsilon f_A \le f_A$. Inserting this into the deviation inequality above and letting $\varepsilon \to 0$ yields to the conclusion that (6.6) holds for all convex $A$. The rest of the proof is identical. $\square$

## References

1. Aida, S., Stroock, D.: Moment estimates derived from poincaré and logarithmic sobolev inequalities. Math. Res. Lett. **1**, 75–86 (1994)
2. Ambrosio, L., Gigli, N., Savaré, G.: Calculus and heat flow in metric measure spaces and applications to spaces with ricci bounds from below. Invent. Math. **195**(2), 289–391 (2014)
3. Ané, C., Blachère, S., Chafaï, D., Fougères, P., Gentil, I., Malrieu, F., Roberto, C., Scheffer, G.: Sur les inégalités de Sobolev logarithmiques. In: Panoramas et Synthèses [Panoramas and Syntheses], vol. 10. Société Mathématique de France, Paris. With a preface by Dominique Bakry and Michel Ledoux (2000)
4. Burago, D., Burago, Y., Ivanov, I.: A Course in Metric Geometry, vol. 33. American Mathematical Society, Providence (2001)
5. Balogh, Z.M., Engulatov, A., Hunziker, L., Maasalo, O.E.: Functional inequalities and Hamilton–Jacobi equations in geodesic spaces. Potential Anal. **36**(2), 317–337 (2012)

6. Bobkov, S.G., Gentil, I., Ledoux, M.: Hypercontractivity of Hamilton–Jacobi equations. J. Math. Pures Appl. (9) **80**(7), 669–696 (2001)
7. Bobkov, S.G., Götze, F.: Discrete isoperimetric and poincaré-type inequalities. Probab. Theory Relat. Fields **114**(2), 245–277 (1999)
8. Bridson, M., Haefliger, A.: Metric Spaces of Non-Positive Curvature. In: Grundlehren der Mathematischen Wissenschaften, vol. 319. Springer, Berlin (1999)
9. Bobkov, S.G., Houdré, C.: Weak dimension-free concentration of measure. Bernoulli **6**(4), 621–632 (2000)
10. Bobkov, S.G., Ledoux, M.: Poincaré's inequalities and talagrand's concentration phenomenon for the exponential distribution. Probab. Theory Relat. Fields **107**(3), 383–400 (1997)
11. Borell, C.: The brunn–minkowski inequality in gauss space. Invent. Math. **30**(2), 207–216 (1975)
12. Buser, P.: A note on the isoperimetric constant. Ann. Sci. École Norm. Sup. (4) **15**(2), 213–230 (1982)
13. Cheeger, J.: A lower bound for the smallest eigenvalue of the Laplacian. In: Problems in Analysis (Papers Dedicated to Salomon Bochner, 1969). Princeton University Press, Princeton, pp 195–199 (1970)
14. Evans, L.C.: Partial differential equations. In: Graduate Studies in Mathematics, vol. 19, 2nd edn. American Mathematical Society, Providence (2010)
15. Feller, W.: An Introduction to Probability Theory and its Applications, vol. II, 2nd edn. Wiley, New York (1971)
16. Funano, K., Shioya, T.: Concentration, ricci curvature and laplacian. Geom. Funct. Anal. **23**(3), 888–936 (2013)
17. Gigli, N., Ledoux, M.: From log sobolev to talagrand: a quick proof. Discret. Contin. Dyn. Syst. **33**(5), 1927–1935 (2013)
18. Gozlan, N.: A characterization of dimension free concentration in terms of transport inequalities. Ann. Probab. **37**(6), 2480–2498 (2009)
19. Gozlan, N.: Poincaré inequalities and dimension free concentration of measure. Ann. Inst. Henri Poincaré Probab. Stat. **46**(3), 703–739 (2010)
20. Gozlan, N., Léonard, C.: Transport inequalities. a survey. Markov Process. Relat. Fields **16**(4), 635–736 (2010)
21. Gozlan, N., Roberto, C., Samson, P.M.: From concentration to logarithmic Sobolev and Poincaré inequalities. J. Funct. Anal. **260**(5), 1491–1522 (2011)
22. Gozlan, N., Roberto, C., Samson, P.M.: Characterization of Talagrand's transport-entropy inequalities on metric spaces. Ann. Probab. **41**(5), 3112–3139 (2013)
23. Gozlan, N., Roberto, C., Samson, P.M.: Hamilton–Jacobi equations on metric spaces and transport-entropy inequalities. Rev. Mat. Iberoam. **30**(1), 133–163 (2014)
24. Gromov, M.: Metric structures for Riemannian and non-Riemannian spaces. In: Progress in Mathematics, vol. 152. Birkhäuser Boston Inc., Boston (1999). Based on the 1981 French original [MR0682063 (85e:53051)]. With appendices by Katz, M., Pansu, P., Semmes, S. Translated from the French by Sean Michael Bates
25. Gromov, M., Milman, V.D.: A topological application of the isoperimetric inequality. Am. J. Math. **105**(4), 843–854 (1983)
26. Kannan, R., Lovász, L., Simonovits, M.: Isoperimetric problems for convex bodies and a localization lemma. Discret. Comput. Geom. **13**(3–4), 541–559 (1995)
27. Ledoux, M.: A simple analytic proof of an inequality by P. Buser. Proc. Am. Math. Soc. **121**(3), 951–959 (1994)
28. Ledoux, M.: The concentration of measure phenomenon. In: Mathematical Surveys and Monographs, vol. 89. American Mathematical Society, Providence (2001)
29. Ledoux, M.: Spectral gap, logarithmic Sobolev constant, and geometric bounds. In: Surveys in Differential Geometry. Vol. IX, pp. 219–240. Int. Press, Somerville (2004)
30. Lott, J., Villani., C.: Hamilton–Jacobi semigroup on length spaces and applications. J. Math. Pures Appl. **88**((3)), 219–229 (2007.)
31. Marton, K.: A simple proof of the blowing-up lemma. IEEE Trans. Inf. Theory **32**(3), 445–446 (1986)
32. Maurey, B.: Some deviation inequalities. Geom. Funct. Anal. **1**(2), 188–197 (1991)
33. Maz'ja, V.G.: Sobolev spaces. In: Springer Series in Soviet Mathematics. Springer, New York (1985)
34. Milman, E.: On the role of convexity in isoperimetry, spectral gap and concentration. Invent. Math. **177**(1), 1–43 (2009)
35. Milman, E.: Isoperimetric and concentration inequalities: equivalence under curvature lower bound. Duke Math. J. **154**(2), 207–239 (2010)
36. Milman, E.: Isoperimetric bounds on convex manifolds. In: Concentration, Functional Inequalities and Isoperimetry, vol. 545 of Contemp. Math., pp. 195–208. American Mathematical Society, Providence (2011)

37. Otto, F., Villani, C.: Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. J. Funct. Anal. **173**(2), 361–400 (2000)
38. Papadopoulos, A.: Metric spaces, convexity and nonpositive curvature. IRMA Lectures in Mathematics and Theoretical Physics, vol. 6. EMS (2005)
39. Schmuckenschläger, M.: Martingales, Poincaré type inequalities, and deviation inequalities. J. Funct. Anal. **155**(2), 303–323 (1998)
40. Sudakov, V.N., Cirel′son, B.S.: Extremal properties of half-spaces for spherically invariant measures. Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **41**, 14–24, 165 (1974). Problems in the theory of probability distributions, II
41. Talagrand, M.: A new isoperimetric inequality and the concentration of measure phenomenon. In: Geometric Aspects of Functional Analysis (1989–1990). Lecture Notes in Math., vol. 1469, pp. 94–124. Springer, Berlin (1991)
42. Talagrand, M.: Concentration of measure and isoperimetric inequalities in product spaces. Publ. Mathématiques de l'IHES **81**, 73–203 (1995)
43. Talagrand, M.: Transportation cost for Gaussian and other product measures. Geom. Funct. Anal. **6**(3), 587–600 (1996)
44. Van der Vaart, A.W.: Asymptotic statistics. In: Cambridge Series in Statistical and Probabilistic Mathematics, vol. 3. Cambridge University Press, Cambridge (1998)
45. Villani, C.: Optimal transport: old and new. In: Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338. Springer, Berlin (2009)