Potential Analysis
https://doi.org/10.1007/s11118-018-9743-1

Multiple Sets Exponential Concentration and Higher Order Eigenvalues

Nathaël Gozlan[1] · Ronan Herry[2,3]

Received: 4 May 2018 / Accepted: 26 October 2018 / Published online: 06 November 2018

Abstract
On a generic metric measured space, we introduce a notion of improved concentration of measure that takes into account the parallel enlargement of k distinct sets. We show that the k-th eigenvalues of the metric Laplacian gives exponential improved concentration with k sets. On compact Riemannian manifolds, this allows us to recover estimates on the eigenvalues of the Laplace-Beltrami operator in the spirit of an inequality of [11].

Keywords Concentration of measure phenomenon · Eigenvalues of the Laplacian · Poincaré inequality

Mathematics Subject Classification (2010) 35P15 · 60E15 · 26D10

1 Introduction

Let (M, g) be a smooth compact connected Riemannian manifold with its normalized volume measure μ and its geodesic distance d. The Laplace-Beltrami operator Δ is then a non-positive operator whose spectrum is discrete. Let us denote by $\lambda^{(k)}$, $k = 0, 1, 2 \ldots$, the eigenvalues of $-\Delta$ written in increasing order. With these notations $\lambda^{(0)} = 0$ (achieved for constant functions) and (by connectedness) $\lambda^{(1)} > 0$ is the so-called spectral gap of M.

✉ Nathaël Gozlan
natael.gozlan@parisdescartes.fr

Ronan Herry
ronan.herry@uni.lu

[1] MAP5 (UMR CNRS 8145), Université Paris Descartes, 45 rue des Saints-Pères, 75270 Paris Cedex 6, France

[2] MRU, Université du Luxembourg, 6 Avenue de la Fonte, L-4363 Esch-sur-Alzette, Luxembourg

[3] LAMA (UMR CNRS 8050), Université Paris Est Marne la Vallée, 5 bd Descartes, 77454 Marne la Vallée Cedex 2, France

The study of the spectral gap of Riemannian manifolds is, by now, a very classical topic which has found important connections with numerous geometrical and analytical questions and properties. The spectral gap constant $\lambda^{(1)}$ is for instance related to Poincaré type inequalities and governs the speed of convergence of the heat flow to equilibrium. It is also related to Ricci curvature via the classical Lichnerowicz theorem [20] and to Cheeger isoperimetric constant via Buser's theorem [7]. We refer to [5, 8] and the references therein for a complete picture.

Another important property of the spectral gap constant, first observed by Gromov and Milman [16], is that it controls exponential concentration of measure phenomenon for the reference measure μ. The result states as follows. Define for all Borel sets $A \subset M$, its r-enlargement A_r as the (open) set of all $x \in E$ such that there exists $y \in A$ with $d(x, y) < r$. Then, for any $A \subset M$ such that $\mu(A) \geq 1/2$ it holds

$$\mu(A_r) \geq 1 - be^{-a\sqrt{\lambda^{(1)}}r}, \qquad \forall r > 0, \tag{1.1}$$

where $a, b > 0$ are some universal constants (according to [19, Theorem 3.1], one can take $b = 1$ and $a = 1/3$). Note that this implication is very general and holds on any metric space supporting a Poincaré inequality (see [19, Corollary 3.2]). See also [1, 6, 15, 26] for alternative derivations, generalizations or refinements of this result.

This note is devoted to a multiple sets extension of the above result. Roughly speaking, we will see that if $A_1, \ldots, A_k$ are sets which are pairwise separated in the sense that $d(A_i, A_j) := \inf\{d(x, y) : x \in A_i, y \in A_j\} > 0$ for any $i \neq j$ and A is their union then the probability of A_r goes exponentially fast to 1 at a rate given by $\sqrt{\lambda^{(k)}}$ as soon as r is such that the sets $A_{i,r}$, $i = 1, \ldots, k$ remain separated. More precisely, it follows from Theorem 1.1 (whose setting is actually more general) that, if $A_1, \ldots, A_k$ are such that $\mu(A_i) \geq \frac{1}{k+1}$ and $d(A_{i,r}, A_{j,r}) > 0$ for all $i \neq j$, then, denoting $A = A_1 \cup \ldots \cup A_k$, it holds

$$\mu(A_r) \geq 1 - \frac{1}{k+1} \exp\left(-c \min\left(r^2\lambda^{(k)}; r\sqrt{\lambda^{(k)}}\right)\right), \tag{1.2}$$

for some universal constant c. This kind of probability estimates first appeared, in a slightly different but essentially equivalent formulation in the work of Chung, Grigor'yan and Yau [10, 11] (see also the related paper [12] by Friedman and Tillich). Nevertheless, the method of proof we use to arrive at (1.2) (based on the Courant-Fischer min-max formula for the $\lambda^{(k)}$'s) is quite different from the one of [10, 11] and seems more elementary and general. This is discussed in details in Section 2.5.

The paper is organized as follows. In Section 2, we prove (1.2) in an abstract metric space framework. This framework contains, in particular, the compact Riemannian case equipped with the Laplace operator presented above. The Section 2.5 contains a detailed discussion of our result with the one of Chung, Grigor'yan & Yau. In Section 3, we recall various bounds on eigenvalues on several non-negatively curved manifolds. Section 4 gives an extension of (1.2) to discrete Markov chains on graphs. In Section 5, we give a functional formulation of the results of Sections 2 and 4. As a corollary of this functional formulation, we obtain a deviation inequality as well as an estimate for difference of two Lipschitz extensions of a Lipschitz function given on k subsets. Finally, Section 6 discusses open questions related to this type of concentration of measure phenomenon.

2 Multiple Sets Exponential Concentration in Abstract Spaces

2.1 Courant-Fischer Formula and Generalized Eigenvalues in Metric Spaces

Let us recall the classical Courant-Fischer min-max formula for the k-th eigenvalue ($k \in \mathbb{N}$) of $-\Delta$, noted $\lambda^{(k)}$, on a compact Riemannian manifold (M, g) equipped with its (normalized) volume measure μ:

$$\lambda^{(k)} = \inf_{\substack{V \subset \mathcal{C}^\infty(M) \\ \dim V = k+1}} \sup_{f \in V \setminus \{0\}} \frac{\int |\nabla f|^2 \, \mathrm{d}\mu}{\int f^2 \, \mathrm{d}\mu}, \tag{2.1}$$

where ∇f is the Riemannian gradient, defined through the Riemannian metric g (see *e.g* [8]) and $|\nabla f|^2 = g(\nabla f, \nabla f)$. The formula (2.1) above does not make explicitly reference to the differential operator Δ. It can be therefore easily generalized to a more abstract setting, as we shall see below.

In all what follows, (E, d) is a complete, separable metric space and μ a reference Borel probability measure on E. Following [9], for any function $f : E \to \mathbb{R}$ and $x \in E$, we denote by $|\nabla f|(x)$ the *local Lipschitz constant* of f at x, defined by

$$|\nabla f|(x) = \begin{cases} 0 \text{ if } x \text{ is isolated} \\ \limsup_{y \to x} \frac{|f(x) - f(y)|}{d(x,y)} \text{ otherwise.} \end{cases}$$

Note that when E is a smooth Riemannian manifold, equipped with its geodesic distance d, then, the local Lipschitz constant of a differentiable function f at x coincides with the norm of $\nabla f(x)$ in the tangent space $T_x E$. With this notion in hand, a natural generalization of (2.1) is as follows (we follow [23, Definition 3.1]):

$$\lambda^{(k)}_{d,\mu} := \inf_{\substack{V \subset H^1(\mu) \\ \dim V = k+1}} \sup_{f \in V \setminus \{0\}} \frac{\int |\nabla f|^2 \, \mathrm{d}\mu}{\int f^2 \, \mathrm{d}\mu}, \qquad k \geq 0, \tag{2.2}$$

where $H^1(\mu)$ denotes the space of functions $f \in L^2(\mu)$ such that $\int |\nabla f|^2 \, \mathrm{d}\mu < +\infty$. In order to avoid heavy notations, we drop the subscript and we simply write $\lambda^{(k)}$ instead of $\lambda^{(k)}_{d,\mu}$ within this section.

Remark 1 At our level of generality, the quantities $\{\lambda^{(k)}_{d,\mu};\ k \in \mathbb{N}\}$ do not always correspond to eigenvalues of some linear heat operator. However, these quantities are still relevant in order to derive our multi-set concentration estimates.

2.2 Statement of the Main Results

To state our first main result, we need further notations: for any $k \geq 1$, we denote by Δ_k the set of vectors $(a_1, \ldots, a_k) \in [0, 1]^k$ satisfying the following linear constraints

$$\sum_{j=1}^{k} a_j \leq 1 \qquad \text{and} \qquad a_i + \sum_{j=1}^{k} a_j \geq 1,\ \forall i \in \{1, \ldots, k\}.$$

Recall the classical notation $d(A, B) = \inf\{d(x, y) : x \in A, y \in B\}$ of the distance between two sets $A, B \subset E$.

The following theorem is the main result of the paper and is proved in Section 2.3.

Theorem 2.1 *There exists a universal constant $c > 0$ such that, for any $k \geq 1$ and for all sets $A_1, \dots, A_k \subset E$ such that $\min_{i \neq j} d(A_i, A_j) > 0$ and $(\mu(A_1), \dots, \mu(A_k)) \in \Delta_k$, the set $A = A_1 \cup A_2 \cup \dots \cup A_k$ satisfies*

$$\mu(A_r) \geq 1 - (1 - \mu(A)) \exp\left(-c \min(r^2 \lambda^{(k)}; r\sqrt{\lambda^{(k)}})\right),$$

for all $0 < r \leq \frac{1}{2} \min_{i \neq j} d(A_i, A_j)$, where $\lambda^{(k)} \geq 0$ is defined by (2.2).

Note that, since $(1/(k+1), \dots, 1/(k+1)) \in \Delta_k$, Theorem 2.1 immediately implies Inequality (1.2). Note also, that for $k = 1$, we obtain an estimate similar to (1.1), but in our bound, taking $r = 0$ gives an equality between the left-hand side and the right-hand side, while (1.1) is only meaningful for $r \geq \log 2/(a\sqrt{\lambda^{(1)}})$.

Inverting our concentration estimate, we obtain the following statement that provides a bound on the $\lambda^{(k)}$'s.

Proposition 2.2 *Let (E, d, μ) be a metric measured space and $\lambda^{(k)}$ be defined as in (2.2). Let $A_1, \dots, A_k$ be measurable sets such that $(\mu(A_1), \dots, \mu(A_k)) \in \Delta_k$, then, with $r = \frac{1}{2} \min_{i \neq j} d(A_i, A_j)$ and $A_0 = E \setminus (\cup A_i)_r$,*

$$\lambda^{(k)} \leq \frac{1}{r^2} \psi\left(\frac{1}{c} \min_i \ln \frac{\mu(A_i)}{\mu(A_0)}\right),$$

where $\psi(x) = \max(x, x^2)$.

Proof Let $A = \cup_i A_i$. Inverting the formula in Theorem 2.1, we obtain

$$\lambda^{(k)} \leq \frac{1}{r^2} \psi\left(\frac{1}{c} \ln \frac{1 - \mu(A)}{1 - \mu(A_r)}\right),$$

where $\psi(x) = \max(x, x^2)$. By definition of Δ_k,

$$1 - \mu(A) = 1 - \sum_i \mu(A_i) \leq \min_i \mu(A_i).$$

Therefore, letting $A_0 = E \setminus A_r$, we obtain the announced inequality by non-decreasing monotonicity of ψ and ln. □

The collection of sets Δ_k, $k \geq 1$ has the following useful stability property:

Lemma 2.3 *Let $I_1, I_2, \dots, I_n$ be a partition of $\{1, \dots, k\}$, $k \geq 1$. Let $a = (a_1, \dots, a_k) \in \mathbb{R}^k$ and define $b = (b_1, \dots, b_n) \in \mathbb{R}^n$ by setting $b_i = \sum_{j \in I_i} a_j$, $i \in \{1, \dots, n\}$. If $a \in \Delta_k$ then $b \in \Delta_n$.*

Proof The proof is obvious and left to the reader. □

Thanks to this lemma it is possible to iterate Theorem 2.1 and to obtain a general bound for $\mu(A_r)$ for all values of $r > 0$. This bound will depend on the way the sets $A_{1,r}, \dots, A_{k,r}$ coalesce as r increases. This is made precise in the following definition.

Definition 2.1 (Coalescence graph of a family of sets) Let $A_1, \dots, A_k$ be subsets of E. The *coalescence graph* of this family of sets is the family of graphs $G_r = (V, E_r)$, $r > 0$, where $V = \{1, 2, \dots, k\}$ and the set of edges E_r is defined as follows: $\{i, j\} \in E_r$ if $d(A_{i,r}, A_{j,r}) = 0$.

Corollary 2.4 *Let $A_1, \ldots, A_k$ be subsets of E such that $\min_{i\neq j} d(A_i, A_j) > 0$ and $(\mu(A_1), \ldots, \mu(A_k)) \in \Delta_k$. For any $r > 0$, let $N(r)$ be the number of connected components in the coalescence graph G_r associated to $A_1, \ldots, A_k$. The function $(0, \infty) \to \{1, \ldots, k\} : r \mapsto N(r)$ is non-increasing and right-continuous. Define $r_i = \sup\{r > 0 : N(r) \geq k - i + 1\}$, $i = 1, \ldots, k$ and $r_0 = 0$ then it holds*

$$\mu(A_r) \geq 1 - (1 - \mu(A)) \exp\left(-c \sum_{i=1}^{k} \phi\left([r \wedge r_i - r_{i-1}]_+ \sqrt{\lambda^{(k-i+1)}}\right)\right), \ \forall r > 0, \quad (2.3)$$

where $\phi(x) = \min(x; x^2)$, $x \geq 0$ and c is the universal constant appearing in Theorem 2.1.

Observe that, contrary to usual concentration results, the bound given above depends on the geometry of the set A.

2.3 Proofs

First, we prove Corollary 2.4. The main argument is to repeatedly apply Theorem 2.1 until two sets or more coalesce.

Proof of Corollary 2.4 We proceed by induction over the number of components k. For $k = 1$, (2.3) follows immediately from Theorem 2.1. Let $k > 1$ and let us assume that (2.3) is true for any collection of subsets $B_1, \ldots, B_l$ satisfying the assumptions of Corollary 2.4 for all $l \in \{1, \ldots, k-1\}$. Let $A_1, A_2, \ldots, A_k$ be a collection of sets satisfying the assumptions of Corollary 2.4. According to Theorem 2.1, it holds

$$\mu(A_r) \geq 1 - (1 - \mu(A)) \exp\left(-c\phi(r\sqrt{\lambda^{(k)}})\right),$$

for all $0 < r \leq \frac{1}{2} \min_{i\neq j} d(A_i, A_j)$.

Let $k_1 = N(\frac{1}{2} \min_{i\neq j} d(A_i, A_j))$ and let $i_1 = k - k_1$. Then, for all $i \in \{1, \ldots, i_1\}$, $r_i = \frac{1}{2} \min_{i\neq j} d(A_i, A_j)$. So that, for all $0 < r \leq r_{i_1}$, the preceding bound can be rewritten as follows (note that only the term of index $i = 1$ gives a non zero contribution)

$$\begin{aligned}\mu(A_r) &\geq 1 - (1 - \mu(A)) \exp\left(-c \sum_{i=1}^{i_1} \phi\left([r \wedge r_i - r_{i-1}]_+ \sqrt{\lambda^{(k-i+1)}}\right)\right) \\ &= 1 - (1 - \mu(A)) \exp\left(-c \sum_{i=1}^{k} \phi\left([r \wedge r_i - r_{i-1}]_+ \sqrt{\lambda^{(k-i+1)}}\right)\right) \qquad (2.4)\end{aligned}$$

which shows that (2.3) is true for $0 < r \leq r_{i_1}$. Now let $I_1, \ldots, I_{k_1}$ be the connected components of G_{r_1} and define, for all $i \in \{1, \ldots, k_1\}$, $B_i = \cup_{j\in I_i} A_{j,r_1}$. It follows easily from Lemma 2.3 that $(\mu(B_1), \ldots, \mu(B_{k_1})) \in \Delta_{k_1}$. Since $\min_{i\neq j} d(B_i, B_j) > 0$, the induction hypothesis implies that

$$\mu(B_s) \geq 1 - (1 - \mu(B)) \exp\left(-c \sum_{i=1}^{k_1} \phi\left([s \wedge s_i - s_{i-1}]_+ \sqrt{\lambda^{(k_1-i+1)}}\right)\right), \qquad \forall s > 0,$$

where $B = B_1 \cup \cdots \cup B_{k_1} = A_{r_1}$ and $s_i = \sup\{s > 0 : N'(s) \geq k_1 - i + 1\}$, $i \in \{1, \ldots, k_1\}$ ($s_0 = 0$) with $N'(s)$ the number of connected components in the graph G'_s associated to

$B_1, \ldots, B_{k_1}$. It is easily seen that $r_{i_1+i} = r_{i_1} + s_i$, for all $i \in \{0, 1 \ldots, k_1\}$. Therefore, we have that, for $r > r_{i_1}$,

$$\begin{aligned}\mu(A_r) &\geq \mu\left(B_{r-r_{i_1}}\right)\\ &\geq 1-(1-\mu(A_{r_{i_1}}))\exp\left(-c\sum_{i=i_1+1}^{k}\phi\left([r\wedge r_i - r_{i-1}]_+\sqrt{\lambda^{(k-i+1)}}\right)\right)\\ &\geq 1-(1-\mu(A))\exp\left(-c\sum_{i=1}^{k}\phi\left([r\wedge r_i - r_{i-1}]_+\sqrt{\lambda^{(k-i+1)}}\right)\right),\end{aligned}$$

where the last line is true by (2.4). □

To prove Theorem 2.1, we need some preparatory lemmas. Given a subset $A \subset E$, and $x \in E$, the minimal distance from x to A is denoted by

$$d(x, A) = \inf_{y\in A} d(x, y).$$

Lemma 2.5 *Let $A \subset E$ and $\epsilon > 0$, then $(E \setminus A_\epsilon)_\epsilon \subset E \setminus A$.*

Proof Let $x \in (E \setminus A_\epsilon)_\epsilon$. Then, there exists $y \in E \setminus A_\epsilon$ (in particular $d(y, A) \geq \epsilon$) such that $d(x, y) < \epsilon$. Since the function $z \mapsto d(z, A)$ is 1-Lipschitz, one has

$$d(x, A) \geq d(y, A) - d(x, y) > 0$$

and so $x \in E \setminus A$. □

Remark 2 In fact, we proved that $(E \setminus A_\epsilon)_\epsilon \subset E \setminus \bar{A}$. The converse is, in general, not true.

Lemma 2.6 *Let $A_1, \ldots, A_k$ be a family of sets such that $(\mu(A_1), \ldots, \mu(A_k)) \in \Delta_k$ and $r := \frac{1}{2}\min_{i\neq j} d(A_i, A_j) > 0$. Let $0 < \epsilon \leq r$ and set $A = \cup_{1\leq i\leq k} A_i$ and $A_0 = E \setminus (A_\epsilon)$. Then,*

$$\max_{i=0,\ldots,k} \frac{\mu(A_{i,\epsilon})}{\mu(A_i)} \leq \frac{1-\mu(A)}{1-\mu(A_\epsilon)}. \tag{2.5}$$

Proof First, this is true for $i = 0$. Indeed, by definition $A_0 = E \setminus (A_\epsilon)$ and, according to Lemma 2.5, $(A_0)_\epsilon \subset A^c$ (the equality is not always true), which proves (2.5) in this case. Now, let us show (2.5) for the other values of i. Since $\epsilon \leq r$, the $A_{j,\epsilon}$'s are disjoint sets. Thence, (2.5) is equivalent to

$$\left(1-\sum_{j=1}^{k}\mu(A_{j,\epsilon})\right)\mu(A_{i,\epsilon}) \leq \left(1-\sum_{j=1}^{k}\mu(A_j)\right)\mu(A_i).$$

This inequality is true as soon as

$$\left(1-\mu(A_{i,\epsilon})-m_i\right)\mu(A_{i,\epsilon}) \leq (1-\mu(A_i)-m_i)\,\mu(A_i),$$

denoting $m_i = \sum_{j\neq i}^{k}\mu(A_j)$. The function $f_i(u) = (1-u-m_i)u$, $u \in [0, 1]$, is decreasing on the interval $[(1-m_i)/2, 1]$. We conclude from this that (2.5) is true for all $i \in \{1, \ldots, k\}$, as soon as $\mu(A_i) \geq (1-m_i)/2$ for all $i \in \{1, \ldots, k\}$ which amounts to $(\mu(A_1), \ldots, \mu(A_k)) \in \Delta_k$. □

For $p > 1$, we define the function $\chi_p \colon [0, \infty[\to [0, 1]$ by

$$\chi_p(x) = (1 - x^p)^p, \quad \text{for} \quad x \in [0, 1] \quad \text{and} \quad \chi_p(x) = 0 \quad \text{for} \quad x > 1.$$

It is easily seen that $\chi_p(0) = 1$, $\chi_p'(0) = \chi_p(1) = \chi_p'(1) = 0$, that χ_p takes values in $[0, 1]$ and that χ_p is continuously differentiable on $[0, \infty[$. We use the function χ_p to construct smooth approximations of indicator functions on E, as explained in the next statement.

Lemma 2.7 *Let $A \subset E$ and consider the function $f(x) = \chi_p(d(x, A)/\epsilon)$, $x \in E$, where $\epsilon > 0$ and $p > 1$. For all $x \in E$, it holds*

$$|\nabla f|(x) \le p^2 \epsilon^{-1} \mathbf{1}_{A_\epsilon \setminus A}$$

Proof Thanks to the chain rule for the local Lipschitz constant (see e.g. [2, Proposition 2.1]),

$$\left| \nabla \chi_p \left(\frac{d(\cdot, A)}{\epsilon} \right) \right| (x) \le \epsilon^{-1} \chi_p' \left(\frac{d(\cdot, A)}{\epsilon} \right) |\nabla d(\cdot, A)|(x).$$

The function $d(\cdot, A)$ being Lipschitz, its local Lipschitz constant is ≤ 1 and, thereby,

$$|\nabla f|(x) \le \chi_p' \left(\frac{d(x, A)}{\epsilon} \right).$$

In particular, thanks to the aforementioned properties of χ, $|\nabla f|$ vanishes on A (and even on $\overline{A}$) and on $\{x \in E : d(x, A) \ge \epsilon\} = E \setminus A_\epsilon$. On the other hand, a simple calculation shows that $|\chi_p'| \le p^2$ which proves the claim. □

Proof of Theorem 2.1 Take Borel sets $A_1, \ldots, A_k$ with $\frac{1}{2} \min_{i \ne j} d(A_i, A_j) \ge r > 0$ and $(\mu(A_1), \ldots, \mu(A_k)) \in \Delta_k$ and consider $A = A_1 \cup \cdots \cup A_k$. Let us show that, for any $0 < \epsilon \le r$, it holds

$$\left(1 + \lambda^{(k)} \epsilon^2\right) (1 - \mu(A_\epsilon)) \le (1 - \mu(A)). \tag{2.6}$$

Let $A_0 = E \setminus (A_\epsilon)$ and set $f_i(x) = \chi_p(d(x, A_i)/\epsilon)$, $x \in E$, $i \in \{0, \ldots, k\}$, where $p > 1$. According to Lemma 2.7 and the fact that $f_i = 1$ on A_i, we obtain

$$\int |\nabla f_i|^2 \, \mathrm{d}\mu = \frac{p^4}{\epsilon^2} \mu(A_{i,\epsilon} \setminus A_i) \quad \text{and} \quad \int f_i^2 \, \mathrm{d}\mu \ge \mu(A_i). \tag{2.7}$$

Since the f_i's have disjoint supports they are orthogonal in $L^2(\mu)$ and, in particular, they span a $k+1$ dimensional subspace of $H^1(\mu)$. Thus, by definition of $\lambda^{(k)}$,

$$\lambda^{(k)} \le \sup_{a \in \mathbb{R}^{k+1}} \frac{\int |\nabla \left(\sum_{i=0}^k a_i f_i \right)|^2 \, \mathrm{d}\mu}{\int \left(\sum_{i=0}^k a_i f_i \right)^2 \mathrm{d}\mu} \le \sup_{a \in \mathbb{R}^{k+1}} \frac{\int \left(\sum_{i=0}^k |a_i| |\nabla f_i| \right)^2 \mathrm{d}\mu}{\int \left(\sum_{i=0}^k a_i f_i \right)^2 \mathrm{d}\mu},$$

where the second inequality comes from the following easy to check sub-linearity property of the local Lipschitz constant:

$$|\nabla (af + bg)| \le |a||\nabla f| + |b||\nabla g|.$$

Since the $f_i's$ and the $|\nabla f_i|'s$ are two orthogonal families, we conclude using (2.7), that

$$\frac{\lambda^{(k)} \epsilon^2}{p^4} \le \sup_{a \in \mathbb{R}^{k+1}} \frac{\sum_{i=0}^k a_i^2 \left(\mu(A_{i,\epsilon}) - \mu(A_i) \right)}{\sum_{i=0}^k a_i^2 \mu(A_i)},$$

which amounts to

$$1+\frac{\lambda^{(k)}\epsilon^2}{p^4} \leq \max_{i=0,\ldots,k} \frac{\mu(A_{i,\epsilon})}{\mu(A_i)}. \tag{2.8}$$

Applying Lemma 2.6 and sending p to 1 gives (2.6). Now, if $n \in \mathbb{N}$ and $0 < \epsilon$ are such that $n\epsilon \leq r$, then iterating (2.6) immediately gives

$$\left(1+\lambda^{(k)}\epsilon^2\right)^n (1-\mu(A_{n\epsilon})) \leq 1-\mu(A).$$

Optimizing this bound over n for a fixed ε gives

$$(1-\mu(A_r)) \leq (1-\mu(A)) \exp\left(-\sup\left\{\lfloor r/\epsilon\rfloor \log\left(1+\lambda^{(k)}\epsilon^2\right) : \epsilon \leq r\right\}\right).$$

Thus, letting

$$\Psi(x) = \sup\left\{\lfloor t\rfloor \log\left(1+\frac{x}{t^2}\right) : t \geq 1\right\}, \qquad x \geq 0, \tag{2.9}$$

it holds

$$(1-\mu(A_r)) \leq (1-\mu(A)) \exp\left(-\Psi\left(\lambda^{(k)}r^2\right)\right).$$

Using Lemma 2.8 below, we deduce that $\Psi\left(\lambda^{(k)}r^2\right) \geq c\min(r^2\lambda^{(k)}; r\sqrt{\lambda^{(k)}})$, with $c = \log(5)/4$, which completes the proof. □

Lemma 2.8 *The function Ψ defined by (2.9) satisfies*

$$\Psi(x) \geq \frac{\log(5)}{4}\min(x; \sqrt{x}), \qquad \forall x \geq 0.$$

Proof Taking $t = 1$, one concludes that $\Psi(x) \geq \log(1+x)$, for all $x \geq 0$. The function $x \mapsto \log(1+x)$ being concave, the function $x \mapsto \frac{\log(1+x)}{x}$ is non-increasing. Therefore, $\log(1+x) \geq \frac{\log(5)}{4}x$ for all $x \in [0,4]$. Now, let us consider the case where $x \geq 4$. Observe that $\lfloor t\rfloor \geq t/2$ for all $t \geq 1$ and so, for $x \geq 4$,

$$\Psi(x) \geq \frac{1}{2}\sup_{t\geq 1}\left\{t\log\left(1+\frac{x}{t^2}\right)\right\} \geq \frac{\log(5)}{4}\sqrt{x},$$

by choosing $t = \sqrt{x}/2 \geq 1$. Thereby,

$$\Psi(x) \geq \frac{\log(5)}{4}\left[x\mathbf{1}_{[0,4]}(x) + \sqrt{x}\mathbf{1}_{[4,\infty)}(x)\right] \geq \frac{\log(5)}{4}\min(x; \sqrt{x}),$$

which completes the proof. □

Remark 3 The conclusion of Lemma 2.8 can be improved. Namely, it can be shown that

$$\Psi(x) = \max\left(\left(1+\lfloor\frac{\sqrt{x}}{a}\rfloor\right)\log\left(1+\frac{x}{\left(1+\lfloor\frac{\sqrt{x}}{a}\rfloor\right)^2}\right); \left(\lfloor\frac{\sqrt{x}}{a}\rfloor\right)\log\left(1+\frac{x}{\left(\lfloor\frac{\sqrt{x}}{a}\rfloor\right)^2}\right)\right),$$

(the second term in the maximum being treated as 0 when $\sqrt{x} < a$) where $0 < a < 2$ is the unique point where the function $(0,\infty) \to \mathbb{R} : u \mapsto \log(1+u^2)/u$ achieves its supremum. Therefore,

$$\Psi(x) \sim \frac{\log(1+a^2)}{a}\sqrt{x}$$

when $x \to \infty$. The reader can easily check that $\frac{\log(1+a^2)}{a} \simeq 0.8$. In particular, it does not seem possible to reach the constant $c = 1$ in Theorem 2.1 using this method of proof.

2.4 Two More Multi-Set Concentration Bounds

The condition $(\mu(A_1), \ldots, \mu(A_k)) \in \Delta_k$ can be seen as the multi-set generalization of the condition, standard in concentration of measure, that the size of the enlarged set has to be bigger than $1/2$. Indeed, the reader can easily verify that $(\frac{1}{k+1}, \ldots, \frac{1}{k+1}) \in \Delta_k$. However, in practice, this condition can be difficult to check. We provide two more multi-set concentration inequalities that hold in full generality. The method of proof is the same as for Theorem 2.1 and is based on (2.8).

Proposition 2.9 *Let (E, d, μ) be a metric measured space and $\lambda^{(k)}$ be defined as in (2.2). Let $(A_1, \ldots, A_k)$ be k Borel sets, $A = \cup_i A_i$ and $A_0 = E \setminus A_r$. Then, with $a_{(1)} = \min_{1\le i\le k} \mu(A_i)$, the following two bounds hold:*

$$1 - \mu(A_r) \le (1 - \mu(A))\frac{1}{\prod_{i=1}^k \mu(A_i)} \exp\left(-c \min\left(r^2\lambda^{(k)}, r\sqrt{\lambda^{(k)}}\right)\right);$$
$$1 - \mu(A_r) \le (1 - \mu(A))\frac{1}{\mu(A)^{\mu(A)/a_{(1)}}} \exp\left(-c \min\left(r^2\lambda^{(k)}, r\sqrt{\lambda^{(k)}}\right)\right).$$

Proof Fix $N \in \mathbb{N}$ and $\epsilon > 0$ such that $N\epsilon \le r$. For $i = 1, \ldots, k$ and $n \le N$, we define

$$\begin{aligned}
\alpha_i(n) &= \frac{\mu(A_{i,n\epsilon})}{\mu(A_{i,(n-1)\epsilon})};\\
M_n &= \max_{1\le i\le k} \alpha_i(n) \vee \frac{1 - \mu(A_{(n-1)\epsilon})}{1 - \mu(A_{n\epsilon})};\\
L_n &= \{i \in \{1, \ldots, k\} | M_n = \alpha_i(n)\};\\
N_i &= \sharp\{n \in \{1, \ldots, N\} | i = \inf L_n\};\\
N_0 &= N - \sum_{i=1}^k N_i.
\end{aligned}$$

Roughly speaking, the number N_i $(0 \le i \le k)$ counts the number of time where the set A_i growths in iterating (2.8). Lemma 2.6 asserts that in the case where $(\mu(A_1), \ldots, \mu(A_k)) \in \Delta_k$, then $N_0 = N$. However, we still obtain from (2.8), for $1 \le i \le k$,

$$\frac{1}{\mu(A_i)} \ge \prod_{n=1}^N \alpha_i(n) \ge \left(1 + \lambda^{(k)}\epsilon^2\right)^{N_i}. \tag{2.10}$$

The first inequality is true because $\mu(A_{i,N\epsilon}) \le 1$ and a telescoping argument. The second inequality is true because, as n ranges from 1 to N, by definition of the number N_i and (2.8), there are, at least N_i terms appearing in the product that can be bounded by $(1 + \lambda^{(k)}\epsilon^2)$. The other terms are bounded above by 1. The case of $i = 0$ is handled in a similar fashion and we obtain:

$$\begin{aligned}
1 - \mu(A_{N\epsilon}) &\le (1 - \mu(A))\left(1 + \lambda^{(k)}\epsilon^2\right)^{-N_0}\\
&= (1 - \mu(A))\left(1 + \lambda^{(k)}\epsilon^2\right)^{-N} \prod_{i=1}^k \left(1 + \lambda^{(k)}\epsilon^2\right)^{N_i}.
\end{aligned} \tag{2.11}$$

The announced bounds will be obtain by bounding the product appearing in the right-hand side and an argument similar to the end of the proof of Theorem 2.1. From (2.10), we have that,

$$\prod_{i=1}^{k}\left(1+\lambda^{(k)}\epsilon^2\right)^{N_i} \leq \frac{1}{\prod_{i=1}^{k}\mu(A_i)}. \tag{2.12}$$

Also, from (2.10),

$$\mu(A_{i,N\epsilon}) \geq \left(1+\lambda^{(k)}\epsilon^2\right)^{N_i}\mu(A_i).$$

Because $N\epsilon \leq r$, the sets $A_{1,N\epsilon},\dots,A_{k,N\epsilon}$ are pairwise disjoint and, thereby,

$$1 \geq \sum \mu(A_{i,N\epsilon}) \geq \sum_{i=1}^{k}\left(1+\lambda^{(k)}\epsilon^2\right)^{N_i}\mu(A_i).$$

Fix $\theta > 0$ to be chosen later. By convexity of exp,

$$\begin{aligned}1+(1-\mu(A))\left(1+\lambda^{(k)}\epsilon^2\right)^{\theta} &\geq \exp\left(\left(\sum_{i=1}^{k}\mu(A_i)N_i+(1-\mu(A))\theta\right)\log\left(1+\lambda^{(k)}\epsilon^2\right)\right)\\ &\geq \exp\left(\left(a_{(1)}\sum_{i=1}^{k}N_i+(1-\mu(A))\theta\right)\log\left(1+\lambda^{(k)}\epsilon^2\right)\right).\end{aligned}$$

Finally, with $p = 1-\mu(A)$ and $t = \theta\log(1+\lambda^{(k)}\epsilon^2)$, we obtain

$$\prod_{i=1}^{k}\left(1+\lambda^{(k)}\epsilon^2\right)^{N_i} \leq \left(\mathrm{e}^{-pt}+p\mathrm{e}^{(1-p)t}\right)^{1/a_{(1)}}.$$

We easily check that, the quantity in the right-hand side is minimal for $t = \log\frac{1}{1-p}$ at which it takes the value $(1-p)^{p-1} = \mu(A)^{-\mu(A)/a_{(1)}}$. Thus,

$$\prod_{i=1}^{k}(1+\lambda^{(k)}\epsilon^2)^{N_i} \leq \frac{1}{\mu(A)^{\mu(A)/a_{(1)}}}. \tag{2.13}$$

Combining (2.12) and (2.13) with (2.11) and the same argument as for (2.9), we obtain the two announced bounds. □

From (2.9), we can derive bounds on the $\lambda^{(k)}$'s. The proof is the same as the one of (2.2) and is omitted.

Proposition 2.10 *Let (E,d,μ) be a metric measured space and $\lambda^{(k)}$ be defined as in (2.2). Let $A_1,\dots,A_k$ be measurable sets, then, with $r = \frac{1}{2}\min_{i\neq j} d(A_i,A_j)$ and $A_0 = E\setminus(\cup A_i)_r$,*

$$\begin{aligned}\lambda^{(k)} &\leq \frac{1}{r^2}\psi\left(\frac{1}{c}\ln\frac{a_{(1)}}{\mu(A_0)}+\frac{1}{c}k\ln\frac{1}{a_{(1)}}\right);\\ \lambda^{(k)} &\leq \frac{1}{r^2}\psi\left(\frac{1}{c}\ln\frac{a_{(1)}}{\mu(A_0)}+\frac{1}{c}\frac{\mu(A)}{a_{(1)}}\ln\frac{1}{\mu(A)}\right),\end{aligned}$$

where $\psi(x) = \max(x,x^2)$ and $a_{(1)} = \min_{1\leq i\leq k}\mu(A_i)$.

2.5 Comparison with the Result of Chung-Grigor'yan-Yau

In [11], the authors obtained the following result:

Theorem 2.11 (Chung-Grigoryan-Yau [11]) *Let M be a compact connected smooth Riemannian manifold equipped with its geodesic distance d and normalized Riemannian volume μ. For any $k \geq 1$ and any family of sets $A_0, \ldots, A_k$, it holds*

$$\lambda^{(k)} \leq \frac{1}{\min_{i \neq j} d^2(A_i, A_j)} \max_{i \neq j} \log \left(\frac{4}{\mu(A_i)\mu(A_j)}\right)^2, \tag{2.14}$$

where $1 = \lambda^{(0)} \leq \lambda^{(1)} \leq \cdots \lambda^{(k)} \leq \cdots$ denotes the discrete spectrum of $-\Delta$.

Let us translate this result in terms of concentration of measure. Let $A_1, \ldots, A_k$ be sets such that $r = \frac{1}{2} \min_{1 \leq i < j \leq k} d(A_i, A_j) > 0$ and define $A = A_1 \cup \cdots \cup A_k$ and $A_0 = M \setminus A_s$, for some $0 < s \leq r$. Then, applying (2.14) to this family of $k + 1$ sets gives the following inequality

$$\min\left(a_{(2)}; 1 - \mu(A_s)\right) \leq \frac{4}{a_{(1)}} \exp(-\sqrt{\lambda^{(k)}} s), \tag{2.15}$$

with $a_{(1)}$ and $a_{(2)}$ being respectively the smallest number and the second smallest number among $(\mu(A_1), \ldots, \mu(A_k))$ (counted with multiplicity). Note that the right hand side is less than or equal to $a_{(2)}$ if and only if $s \geq s_o := \frac{1}{\sqrt{\lambda_k}} \log\left(\frac{4}{a_{(1)}a_{(2)}}\right)$, so that (2.15) is equivalent to the following statement:

$$\mu(A_s) \geq 1 - \frac{4}{a_{(1)}} \exp(-\sqrt{\lambda^{(k)}} s), \qquad \forall s \in [\min(s_o, r); r]. \tag{2.16}$$

We note that (2.16) holds for any family of sets, whereas the inequality given in Theorem 2.1 is only true when $(\mu(A_1), \ldots, \mu(A_k)) \in \Delta_k$. Also due to the fact that the constant c appearing in Theorem 2.1 is less than 1, (2.16) is asymptotically better than ours (see also Remark 3 above). On the other hand, one sees that (2.16) is only valid for s large enough (and its domain of validity can thus be empty when $s_o > r$) whereas our inequality is true on the whole interval $(0, r]$. It does not seem also possible to iterate (2.16) as we did in Corollary 2.4. Finally, observe that the method of proof used in [11] and [10] is based on heat kernel bounds and is very different from ours.

Let us translate Theorem 2.11 in a form closer to our Proposition 2.2. Fix k sets $A_1, \ldots, A_k$ such that $(\mu(A_1), \ldots, \mu(A_k)) \in \Delta_k$. Let $2r = \min d(A_i, A_j)$, where the infimum runs on $i, j = 1, \ldots, k$ with $i \neq j$. We have to choose a $(k + 1)$-th set. In view of Theorem 2.11, the most optimal choice is to choose $A_0 = E \setminus (\cup A_i)_r$. Indeed, it is the biggest set (in the sense of inclusion) such that $\min d(A_i, A_j) = r$ where this time the infimum runs on $i, j = 0, \ldots, k$ and $i \neq j$. We let $a_{(0)} = \mu(A_0)$ and we remark that if $(\mu(A_1), \ldots, \mu(A_k)) \in \Delta_k$ then $a_{(0)} \leq a_{(1)}$. The bound (2.14) can be read: for all $r > 0$,

$$\lambda^{(k)} \leq \frac{1}{r^2}\left(\log \frac{4}{a_{(1)}a_{(0)}}\right)^2.$$

Therefore, to compare it to our bound, we need to solve

$$\phi^{-1}\left(\frac{1}{c} \log \frac{a_{(1)}}{a_{(0)}}\right)^2 \leq \left(\log \frac{4}{a_{(1)}a_{(0)}}\right)^2.$$

Because the right-hand side is always ≥ 1, taking the square root and composing with the non-decreasing function ϕ yields

$$\frac{1}{c}\log\frac{a_{(1)}}{a_{(0)}} \leq \log\frac{4}{a_{(1)}a_{(0)}}.$$

That is

$$a_{(1)}^{1+c} \leq 4^c a_{(0)}^{1-c}.$$

In other words, on some range our bound is better and in some other range their bound is better. However, if the constant $c = 1$ could be attained in Theorem 2.1, this would show that our bound is always better. Note that comparing the bounds obtained in Proposition 2.10 and the one of [11] is not so clear as, without the assumption that $(\mu(A_1), \dots, \mu(A_k)) \in \Delta_k$ it is not necessary that $a_{(0)} \leq a_{(1)}$ and in that case we would have to compare different sets.

3 Eigenvalue Estimates for Non-Negatively Curved Spaces

We recall the values of the $\lambda^{(k)}$'s that appear in Theorem 2.1 in the case of two important models of positively curved spaces in geometry. Namely:

(i) The n-dimensional sphere of radius $\sqrt{\frac{n-1}{\rho}}$, $\mathbb{S}^{n,\rho}$ endowed with the natural geodesic distance $d_{n,\rho}$ arising from its canonical Riemannian metric and its normalized volume measure $\mu_{n,\rho}$ which has constant Ricci curvature equals to ρ and dimension n.

(ii) The n-dimensional Euclidean space $\mathbb{R}^n$ endowed with the n-dimensional Gaussian measure of covariance $\rho^{-1}\mathrm{Id}$,

$$\gamma_{n,\rho}(\mathrm{d}x) = \frac{\rho^{n/2}\mathrm{e}^{-\rho|x|^2/2}}{(2\pi)^{n/2}}\mathrm{d}x.$$

This space has dimension ∞ and curvature bounded below by ρ in the sense of [4].

These models arise as weighted Riemannian manifolds without boundary having a purely discrete spectrum. In that case, it was proved in [23, Proposition 3.2] that the λ_k's of (2.2) are exactly the eigenvalues (counted with multiplicity) of a self-adjoint operator that we give explicitly in the following. Using a comparison between eigenvalues of [23], we obtain an estimates for eigenvalues in the case of log-concave probability measure over the Euclidean $\mathbb{R}^n$.

Example 1 (Spheres) On $\mathbb{S}^{n,\rho}$, the eigenvalues of minus the Laplace-Beltrami operator (see for instance [3, Chapter 3]) are of the form $\rho^{-2}(n-1)^2 l(l+n-1)$ for $l \in \mathbb{N}$ and the dimension of the corresponding eigenspace $H_{l,n}$ is

$$\dim H_{l,n} = \frac{2l+n-1}{l}\binom{l+n-2}{l-1}, \text{ if } l > 0 \;\; \dim H_{l,n} = 1, \text{ if } l = 0.$$

Consequently,

$$D_{l,n} := \dim \bigoplus_{l'=0}^{l} H_{l',n} = \binom{n+l}{l} + \binom{n+l-1}{l-1},$$

and $\lambda^{(k)} = \rho^{-2}(n-1)^2 l(l+n-1)$ if and only if $D_{l-1,n} < k \leq D_{l,n}$ where $\lambda^{(k)}$ is the k-th eigenvalues of $-\Delta_{\mathbb{S}^{n,\rho}}$ and coincides with the variational definition given in (2.2).

Example 2 (Gaussian spaces) On the Euclidean space $\mathbb{R}^n$, equipped with the Gaussian measure $\gamma_{n,\rho}$, the corresponding weighted Laplacian is $\Delta_{\gamma_{n,\rho}} = \Delta_{\mathbb{R}^n} - \rho x \cdot \nabla$. The eigenvalues of $-\Delta_{\gamma_{n,\rho}}$ are exactly of the form $\rho^2 q$ and the dimension of the associated eigenspace $H_{q,n}$ is

$$\dim H_{q,n} = \binom{n+q-1}{q}.$$

Consequently,

$$D_{q,n} := \dim \bigoplus_{q'=0}^{q} H_{q',n} = \binom{n+q}{q},$$

and $\lambda^{(k)} = \rho^{-2} q$ if and only if $D_{q-1,n} < k \leq D_{q,n}$ where $\lambda^{(k)}$ is the k-th eigenvalues of $-\Delta_{\gamma_{n,\rho}}$ and coincides with the variational definition given in (2.2).

Example 3 (Log-concave Euclidean spaces) We study the case where $E = \mathbb{R}^n$, d is the Euclidean distance and μ is a strictly log-concave probability measure. By this we mean that $\mu(\mathrm{d}x) = \mathrm{e}^{-V(x)}\mathrm{d}x$, where $V : \mathbb{R}^n \to \mathbb{R}$ such that V is $\mathcal{C}^2$ and satisfying $\nabla^2 V \geq \rho$ for some $\rho > 0$. It is a consequence of [4, Proposition 4] that such a condition on V implies that the semigroup generated by the solution of the stochastic differential equation

$$\mathrm{d}X_t = \sqrt{2}\mathrm{d}B_t - \nabla V(X_t)\mathrm{d}t,$$

where B is a Brownian motion on $\mathbb{R}^n$, satisfies the curvature-dimension $CD(\infty, \rho)$ of Bakry-Emery and, therefore, holds the log-Sobolev inequality, for all $f \in \mathcal{C}_c^\infty(\mathbb{R}^n)$,

$$\mathbf{Ent}_\mu f^2 \leq \frac{2}{\rho} \int |\nabla f(x)|^2 \mu(\mathrm{d}x).$$

Such an inequality implies the super-Poincaré of [27, Theorem 2.1] that in turns implies that the self-adjoint operator $L = -\Delta + \nabla V \cdot \nabla$ has a purely discrete spectrum. In that case, the $\lambda^{(k)}$ of (2.2) corresponds to these eigenvalues and [23] showed that

$$\lambda^{(k)} \geq \lambda^{(k)}_{\gamma_{n,\rho}},$$

where $\lambda^{(k)}_{\gamma_{n,\rho}}$ is the eigenvalues of $-\Delta_{\gamma_{n,\rho}}$ of the previous example.

4 Extension to Markov Chains

As in the classical case (see [19, Theorem 3.3]), our continuous result admits a generalization on finite graphs or more broadly in the setting of Markov chains on a finite state space. We consider a finite set E and $X = (X_n)$ be a irreducible time-homogeneous Markov chain with state space E. We write $p(x, y) = \mathbb{P}(X_1 = y | X_0 = x)$ and we regard p as a matrix. We assume that X admits a reversible probability measure μ on E such that $p(x, y)\mu(x) = p(y, x)\mu(y)$ and $\mu(y) = \sum_x p(x, y)\mu(x)$. This induces a graph structure on E by the following procedure. Set the elements of E as the vertex of the graph and for $x, y \in E$ connect them with an edge if $p(x, y) > 0$. As the chain is irreducible, this graph is connected. We equip E with the induced graph distance d. We write $L = p - I$, where I stands for the identity. The operator $-L$ is a symmetric positive operator on $\mathcal{L}^2(\mu)$. We let $\lambda^{(k)}$ be the eigenvalues of this operator. Then, our Theorem 2.1 extends as follows:

Theorem 4.1 *For any $k \geq 1$ and for all sets $A_1, \dots, A_k \subset E$ such that $\min_{i \neq j} d(A_i, A_j) \geq 1$ and $(\mu(A_1), \dots, \mu(A_k)) \in \Delta_k$ the set $B = A_1 \cup A_2 \cup \cdots \cup A_k$ satisfies*

$$\mu(B_n) \geq 1 - (1 - \mu(B))\left(1 + \lambda^{(k)}\right)^{-n},$$

for all $1 \leq n \leq \frac{1}{2} \min_{i \neq j} d(A_i, A_j)$ where $\lambda^{(k)}$ is the k-th eigenvalue of the operator $-L$ acting on $\mathcal{L}^2(\mu)$.

Proof We let $\Pi(x, y) = p(x, y)\mu(x)$ and

$$\mathcal{E}(f, g) = \frac{1}{2} \sum (f(y) - f(x))(g(y) - g(x))\Pi(x, y) = \langle f, -Lg \rangle_\mu.$$

For any set A, we define the discrete boundary of A as $\partial A = A_1 \setminus A \cup (A^C)_1 \setminus A^C$. Let (X_n) be the Markov chain with transition kernel p and initial distribution μ. By reversibility of μ, (X_0, X_1) is an exchangeable pair of law Π whose the marginals are given by μ. Then, for a set U, we have

$$\mathcal{E}(1_U) = \mathbb{E}1_U(X_0)(1_U(X_0) - 1_U(X_1)) = \mathbb{P}(X_0 \in U, X_1 \notin U) \leq \mathbb{P}(X_1 \in \partial U) = \mu(\partial U).$$

Observe that if $d(U, V) \geq 1$, U and V are disjoint and $U \times V \notin \operatorname{supp} \Pi$ so that $\mathcal{E}(1_U, 1_V) = 0$. By Courant-Fischer's min-max theorem

$$\lambda^{(k)} = \min_{\dim V = k+1} \max_{f \in V} \frac{\mathcal{E}(f, f)}{\mu(f^2)}.$$

Choose sets $A_1, \dots, A_k$ with $d(A_i, A_j) \geq 2n$ $(i \neq j)$ and $(\mu(A_1), \dots, \mu(A_k)) \in \Delta_k$. Set $f_i = 1_{A_i}$. The f_i's have disjoint support and so they are orthogonal in $L^2(\mu)$. By the previous variational representation of $\lambda^{(k)}$, we have

$$\lambda^{(k)} \leq \sup_{a_i} \frac{\mathcal{E}\left(\sum_{i=0}^k a_i f_i\right)}{\int \left(\sum_{i=0}^k a_i f_i\right)^2 \mathrm{d}\mu} = \sup_{a_i} \frac{\sum a_i a_{i'} \mathcal{E}(f_i, f_{i'})}{\sum a_i a_{i'} \int f_i f_{i'} \mathrm{d}\mu} = \sup_{a_i} \frac{\sum_{i=0}^k a_i^2 \mathcal{E}(f_i)}{\sum_{i=0}^k a_i \int f_i^2 \mathrm{d}\mu}.$$

In other words,

$$\lambda^{(k)} \leq \max_{i=0,\dots,k} \frac{\mu((A_i)_1) + \mu((A_i^C)_1) - 1}{\mu(A_i)} \leq \frac{\mu((A_i)_1) - \mu(A_i)}{\mu(A_i)},$$

where the last inequality comes from the fact that, by Lemma 2.5, $\mu(E \setminus (E \setminus A)_1) \geq \mu(A)$. Consider the set $B = \cup_{i=1}^k A_i$ and choose $A_0 = E \setminus B_1$. In that case, by Lemma 2.6 with $\epsilon = 1$, we have

$$\max_{i=0,\dots,k} \frac{\mu((A_i)_1)}{\mu(A_i)} \leq \frac{1 - \mu(B)}{1 - \mu(B_1)}.$$

Thus, we proved that

$$(1 + \lambda^{(k)})(1 - \mu(B_1)) \leq (1 - \mu(B)).$$

We derive the announced result by an immediate recursion. □

5 Functional Forms of the Multiple Sets Concentration Property

We investigate the functional form of the multi-sets concentration of measure phenomenon results obtained in Sections 2 and 4.

Proposition 5.1 *Let (E, d) be a metric space equipped with a Borel probability measure μ. Let $\alpha_k : [0, \infty) \to [0, \infty)$. The following properties are equivalent:*

1. *For all Borel sets $A_1, \dots, A_k \subset E$ such that $(\mu(A_1), \dots, \mu(A_k)) \in \Delta_k$, the set $A = A_1 \cup \dots \cup A_k$ satisfies*
$$\mu(A_r) \geq 1 - (1 - \mu(A))\alpha_k(r), \qquad \forall 0 < r \leq \frac{1}{2} \min_{i \neq j} d(A_i, A_j). \tag{5.1}$$
2. *For all 1-Lipschitz functions $f_1, \dots, f_k : E \to \mathbb{R}$ such that the sublevel sets $A_i = \{f_i \leq 0\}$ are such that $(\mu(A_1), \dots, \mu(A_k)) \in \Delta_k$, the function $f^* = \min(f_1, \dots, f_k)$ satisfies*
$$\mu(f^* < r) \geq 1 - \mu(f^* \leq 0)\alpha_k(r), \qquad \forall 0 < r \leq \frac{1}{2} \min_{i \neq j} d(A_i, A_j).$$

Together with Theorem 2.1 or Theorem 4.1, one thus sees that the presence of multiple wells can improve the concentration properties of a Lipschitz function.

Proof It is clear that (2) implies (1) when applied to $f_i(x) = d(x, A_i)$, in which case $A_i = \{f_i \leq 0\}$ and $f^*(x) = d(x, A)$. The converse is also very classical. First, observe that $\{f^* < r\} = \cup_{i=1}^k \{f_i < r\}$. Then, since f_i is 1-Lipschitz, it holds $A_{i,r} \subset \{f_i < r\}$ with $A_i = \{f_i \leq 0\}$ and so letting $A = A_1 \cup \dots \cup A_k$, it holds $A_r \subset \{f^* < r\}$. Therefore, applying (1) to this set A gives (2). □

When (5.1) holds, we will say that the probability metric space (E, d, μ) satisfies the multi-set concentration of measure property of order k with the concentration profile α_k.

In the usual setting ($k = 1$), the concentration of measure phenomenon implies deviation inequalities for Lipschitz functions around their median. The next result generalizes this well known fact to $k > 1$.

Proposition 5.2 *Let (E, d, μ) be a probability metric space satisfying the multi-set concentration of measure property of order k with the concentration profile α_k and $f : E \to \mathbb{R}$ be a 1-Lipschitz function. If $I_1, \dots, I_k \subset \mathbb{R}$ are k disjoint Borel sets such that $(\mu(f \in I_1), \dots, \mu(f \in I_k)) \in \Delta_k$, then it holds*
$$\mu\left(f \in \cup_{i=1}^k I_{i,r}\right) \geq 1 - (1 - \mu(f \in \cup_{i=1}^k I_i))\alpha_k(r), \qquad \forall 0 < r \leq \frac{1}{2} \min_{i \neq j} d(I_i, I_j)$$

Proof Let ν be the image of μ under the map f. Since f is 1-Lipschitz, the metric space $(\mathbb{R}, |\cdot|, \nu)$ satisfies the multi-set concentration of measure property of order k with the same concentration profile α_k as μ. Details are left to the reader. □

Let us conclude this section by detailling some application of potential interest in approximation theory.

Suppose that $f : E \to \mathbb{R}$ is some 1-Lipschitz function and $A_1, \dots A_k$ are (pairwise disjoint) subsets of E such that $(\mu(A_1), \dots, \mu(A_k)) \in \Delta_k$. Let us assume that the restrictions $f_{|A_i}$, $i \in \{1, \dots, k\}$ are known and that one wishes to estimate or reconstruct f outside $A = \cup_{i=1}^k A_i$. To that aim, one can consider an explicit 1-Lipschitz extension of $f_{|A}$, that is to say a 1-Lipschitz function $g : E \to \mathbb{R}$ (constructed based on our knowledge of f on A exclusively) such that $f = g$ on A. There are several canonical ways to perform the extension of a Lipschitz function defined on a sub domain (known as

Kirszbraun-McShane-Whitney extension [18, 22, 28]). One can consider for instance the functions

$$g_+(x) = \inf_{y\in A}\{f(y) + d(x, y)\} \qquad \text{or} \qquad g_-(x) = \sup_{y\in A}\{f(y) - d(x, y)\}, \qquad x \in E.$$

It is a very classical fact that functions g_- and g_+ are 1-Lipschitz extensions of $f_{|A}$ and moreover that any extension g of $f_{|A}$ satisfies $g_- \le g \le g_+$ (see e.g [17]).

The following simple result shows that, for any 1-Lipschitz extension g of $f_{|A}$, the probability of error $\mu(|f - g| > r)$ is controlled by the multi-set concentration profile α_k. In particular, in the framework of our Theorem 2.1, this probability of error is controled by $\lambda^{(k)}$.

Proposition 5.3 *Let (E, d, μ) be a probability metric space satisfying the multi-set concentration of measure property of order k with the concentration profile α_k and $f : E \to \mathbb{R}$ be a 1-Lipschitz function. Let $A_1, \dots A_k$ be subsets of E such that $(\mu(A_1), \dots, \mu(A_k)) \in \Delta_k$; then for any 1-Lipschitz extension g of $f_{|A}$, it holds*

$$\mu(|f - g| \ge r) \le (1 - \mu(A))\alpha_k(r/2), \qquad \forall 0 < r \le \min_{i\neq j} d(A_i, A_j).$$

Proof The function $h : E \to \mathbb{R}$ defined by $h(x) = |f - g|(x)$, $x \in E$, is 2-Lipschitz and vanishes on A. Therefore, for any $x \in E$ and $y \in A$, it holds $h(x) \le h(y) + 2d(x, y) = 2d(x, y)$. Optimizing over $y \in A$ gives that $h(x) \le 2d(x, A)$. Therefore $\{h \ge r\} \subset \{x : d(x, A) \ge r/2\} = \left(A_{r/2}\right)^c$ and so, if $0 < r \le \min_{i\neq j} d(A_i, A_j)$, it holds

$$\mu(|f - g| \ge r) \le (1 - \mu(A))\alpha_k(r/2). \qquad \square$$

Remark 4 Let us remark that Propositions 5.1 to 5.3 can be immediately extended under the following more general (but notationally heavier) multi-set concentration of measure assumption: there exists functions $\alpha_k : [0, \infty) \to [0, \infty)$ and $\beta_k : [0, \infty)^k \to [0, \infty]$ such that for all Borel sets $A_1, \dots, A_k \subset E$, the set $A = A_1 \cup \dots \cup A_k$ satisfies

$$\mu(A_r) \ge 1 - \beta_k(\mu(A_1), \cdots, \mu(A_k))\alpha_k(r), \qquad \forall 0 < r \le \frac{1}{2}\min_{i\neq j} d(A_i, A_j).$$

This framework contains the preceding one, by choosing $\beta_k(a) = 1 - \sum_{i=1}^k a_i$ if $a = (a_1, \dots, a_k) \in \Delta_k$ and $+\infty$ otherwise. It also contains the concentration bounds obtained in Proposition 2.9, corresponding respectively to

$$\beta_k(a) = \frac{1 - \sum_{i=1}^k a_i}{\prod_{i=1}^k a_i}, \text{ and } \beta_k(a) = \frac{1 - \sum_{i=1}^k a_i}{\left(\sum_{i=1}^k a_i\right)^{\sum_{i=1}^k a_i / \min(a_1, \cdots, a_k)}}, \quad a = (a_1, \dots, a_k).$$

6 Open Questions

We list open questions related to the multi-set concentration of measure phenomenon.

6.1 Gaussian Multi-Set Concentration

Using the terminology introduced in Section 5, Theorem 2.1 and the material exposed in Section 3 tell us that, if μ has a density of the form e^{-V} with respect to Lebesgue measure on $\mathbb{R}^n$ with a smooth function V such that Hess $V \ge \rho > 0$, then the probability metric space $(\mathbb{R}^n, |\cdot|, \mu)$ satisfies the multi-set concentration of measure property of order k with the concentration profile

$$\alpha_k(r) = \exp\left(-c \min(r^2 \lambda^{(k)}_{\gamma_n,\rho}; r\sqrt{\lambda^{(k)}_{\gamma_n,\rho}})\right), \qquad r \geq 0,$$

where $\lambda^{(k)}_{\gamma_n,\rho}$ denotes the kth eigenvalue of the n-dimensional centered Gaussian measure with covariance matrix $\rho^{-1}\mathrm{Id}$. Since the measure μ satisfies the log-Sobolev inequality, it is well known that it satisfies a (classical) Gaussian concentration of measure inequality. Therefore, it is natural to conjecture that μ satisfies a multi-set concentration of measure property of order $k \geq 1$ with a profile of the form

$$\beta_k(r) = \exp\left(-C_{k,\rho,n} r^2\right), \qquad r \geq 0,$$

for some constant $C_{k,\rho,n}$ depending solely on its arguments. In addition, it would be interesting to see how usual functional inequalities (Log-Sobolev, transport-entropy, ...) can be modified to catch such a concentration of measure phenomenon.

6.2 Equivalence Between Multi-Set Concentration and Lower Bounds on Eigenvalues in Non-Negative Curvature

Let us quickly recall the main finding of E. Milman [24, 25], that is, under non-negative curvature assumptions, a concentration of measure estimate implies a bound on the spectral gap. Let μ be a probability measure with a density of the form e^{-V} on a smooth connected Riemannian manifold M with V a smooth function such that

$$\mathrm{Ric} + \mathrm{Hess}\, V \geq 0. \tag{6.1}$$

Assume that μ satisfies a concentration inequality of the form: for all $A \subset M$ such that $\mu(A) \geq 1/2$

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

where α is a function such that $\alpha(r_o) < 1/2$ for at least one value $r_o > 0$. Then, letting λ_1 be the first non zero eigenvalue of the operator $-\Delta + \nabla V \cdot \nabla$, it holds $\lambda_1 \geq \frac{1}{4}\left(\frac{1-2\alpha(r_o)}{r_o}\right)^2$. It would be very interesting to extend Milman's result to a multi-set concentration setting. More precisely, if μ satisfies the curvature condition (6.1) and the multi-set concentration of measure property of order k with a profile of the form $\alpha_k(r) = \exp(-\min(ar^2, \sqrt{a}r))$, $r \geq 0$, can we find a universal function φ_k such that $\lambda_k \geq \varphi_k(a)$?

This question already received some attention in recent works by Funano and Shioya [13, 14]. In particular, let us mention the following improvement of the Chung-Grigor'yan-Yau inequality obtained in [13]. There exists a universal constant $c > 1$ such that if μ is a probability measure satisfying the non-negative curvature assumption (6.1), it holds: for any family of sets $A_0, A_1, \ldots, A_l$ with $1 \leq l \leq k$

$$\lambda^{(k)} \leq c^{k-l+1} \frac{1}{\min_{i\neq j} d^2\left(A_i, A_j\right)} \max_{i\neq j} \log\left(\frac{4}{\mu(A_i)\mu(A_j)}\right)^2. \tag{6.2}$$

Note that the difference with (2.14) is that $\lambda^{(k)}$ is estimated by a reduced number of sets. Using (6.2) (with $l = 1$) together with Milman's result recalled above, Funano showed that there exists some constant C_k depending only on k such that under the curvature condition (6.1), it holds $\lambda_k \leq C_k \lambda_0$ (recovering the main result of [14]). The constant C_k is explicit (contrary to the constant of [14]) and grows exponentially when $k \to \infty$. This result has been then improved by Liu [21], where a constant $C_k = O(k^2)$ has been obtained. As observed by Funano [13], a positive answer to the open question stated above would yield that under (6.1) the ratios λ_{k+1}/λ_k are bounded from above by a universal constant.

References

1. Aida, S., Stroock, D.: Moment estimates derived from Poincaré and logarithmic Sobolev inequalities. Math. Res. Lett. **1**(1), 75–86 (1994)
2. Ambrosio, L., Ghezzi, R.: Sobolev and bounded variation functions on metric measure spaces. In: Geometry, Analysis and Dynamics on Sub-Riemannian Manifolds. Vol. II, EMS Ser. Lect. Math., pp. 211–273. Eur. Math. Soc., Zürich (2016)
3. Atkinson, K., Han, W.: Spherical Harmonics and Approximations on the Unit Sphere: An Introduction, Volume 2044 of Lecture notes in Mathematics. Springer, Heidelberg (2012)
4. Bakry, D., Émery, M.: Diffusions hypercontractives. In: Séminaire de probabilités, XIX, 1983/84, volume 1123 of Lecture Notes in Math, pp. 177–206. Springer, Berlin (1985)
5. Bakry, D., Gentil, I., Ledoux, M.: Analysis and geometry of Markov diffusion operators, volume 348 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Cham (2014)
6. Bobkov, S., Ledoux, M.: Poincaré's inequalities and Talagrand's concentration phenomenon for the exponential distribution. Probab. Theory Related Fields **107**(3), 383–400 (1997)
7. Buser, P.: A note on the isoperimetric constant. Ann. Sci. É,cole Norm. Sup. (4) **15**(2), 213–230 (1982)
8. Chavel, I.: Eigenvalues in Riemannian Geometry, volume 115 of Pure and Applied Mathematics. Academic Press, Inc., Orlando (1984). Including a chapter by Burton Randol, With an appendix by Jozef Dodziuk
9. Cheeger, J.: Differentiability of Lipschitz functions on metric measure spaces. Geom. Funct. Anal. **9**(3), 428–517 (1999)
10. Chung, F.R.K., Grigor'yan, A., Yau, S.-T.: Eigenvalues and diameters for manifolds and graphs. In: Tsing Hua Lectures on Geometry & Analysis (Hsinchu, 1990–1991), pp. 79–105. Int. Press, Cambridge (1997)
11. Chung, F.R.K., Grigor'yan, A., Yau, S.-T.: Upper bounds for eigenvalues of the discrete and continuous Laplace operators. Adv. Math. **117**(2), 165–178 (1996)
12. Friedman, J., Tillich, J.-P.: Laplacian eigenvalues and distances between subsets of a manifold. J. Differential Geom. **56**(2), 285–299 (2000)
13. Funano, K.: Estimates of Eigenvalues of the Laplacian by a reduced number of subsets. Israel J. Math. **217**(1), 413–433 (2017)
14. Funano, K., Shioya, T.: Concentration, Ricci curvature, and Eigenvalues of Laplacian. Geom. Funct. Anal. **23**(3), 888–936 (2013)
15. Gozlan, N., Roberto, C., Samson, P.-M.: From dimension free concentration to the P,oincaré inequality. Calc. Var Partial Differential Equations **52**(3-4), 899–925 (2015)
16. Gromov, M., Milman, V.D.: A topological application of the isoperimetric inequality. Amer. J. Math. **105**(4), 843–854 (1983)
17. Heinonen, J.: Lectures on Lipschitz Analysis, volume 100 of *Report.* University of Jyväskylä Department of Mathematics and Statistics. University of Jyväskylä, Jyväskylä (2005)
18. Kirszbraun, M.: Uber die zusammenziehende und lipschitzsche transformationen. Fundam. Math. **22**, 77–108 (1934)
19. Ledoux, M.: The Concentration of Measure Phenomenon, Volume 89 of Mathematical Surveys and Monographs. American Mathematical Society, Providence (2001)
20. Lichnerowicz, A.: Géométrie des groupes de transformations. Travaux et recherches mathématiques, III. Dunod, Paris (1958)
21. Liu, S.: An optimal dimension-free upper bound for eigenvalue ratios. Arxiv e-prints, May (2014)
22. McShane, E.J.: Extension of range of functions. Bull. Amer. Math. Soc. **40**(12), 837–842 (1934)
23. Milman, E.: Spectral estimates, contractions and hypercontractivity. J. Spectr. Theory **8**(2), 669–714 (2018)
24. Milman, E.: On the role of convexity in isoperimetry, spectral gap and concentration. Invent. Math. **177**(1), 1–43 (2009)
25. Milman, E.: Isoperimetric and concentration inequalities: equivalence under curvature lower bound. Duke Math. J. **154**(2), 207–239 (2010)
26. Schmuckenschläger, M.: Martingales, Poincaré type inequalities, and deviation inequalities. J. Funct. Anal. **155**(2), 303–323 (1998)

27. Wang, F.-Y.: Functional inequalities for empty essential spectrum. J. Funct. Anal. **170**(1), 219–245 (2000)
28. Whitney, H.: Analytic extensions of differentiable functions defined in closed sets. Trans. Amer. Math. Soc. **36**(1), 63–89 (1934)