Functional inequalities for discrete gradients and applications to the geometric distribution

Aldéric Joulin* and Nicolas Privault†
Université de La Rochelle
Avenue Michel Crépeau
17042 La Rochelle Cedex

February 3, 2004

Abstract

We present several functional inequalities for finite difference gradients, such as a Cheeger inequality, Poincaré and (modified) logarithmic Sobolev inequalities, associated deviation estimates, and an exponential integrability property. In the particular case of the geometric distribution on $\mathbb{N}$ we use an integration by parts formula to compute the optimal isoperimetric and Poincaré constants, and to obtain an improvement of our general logarithmic Sobolev inequality. By a limiting procedure we recover the corresponding inequalities for the exponential distribution. These results have applications to interacting spin systems under a geometric reference measure.

Key words: Geometric distribution, isoperimetry, logarithmic Sobolev inequalities, spectral gap, Herbst method, deviation inequalities, Gibbs measures.
Mathematics Subject Classification. 60E15, 60E07, 60K35.

1 Introduction

Isoperimetry consists in finding sets of minimal surface among sets of a given volume, i.e. to search for optimal constants c in inequalities of the form

$$cI\left(\mu(A)\right) \leq \mu_s(A), \tag{1.1}$$

*ajoulin@univ-lr.fr
†nprivaul@univ-lr.fr

where μ_s and μ are respectively surface and volume measures and I is a non-negative function on $[0,1]$. Isoperimetric constants are linked, via co-area formulas, to functional inequalities such as Poincaré or logarithmic Sobolev inequalities. Discrete isoperimetry has been studied in various contexts, such as reversible Markov chains [10], [15], graph theory [1, 18], statistical mechanics, cf. e.g. [9].

In this paper we consider the general discrete setting of a probability space $(E, \mathcal{E}, \mu)$, and a finite difference gradient d^+ defined as $\mathrm{d}^+ f = f \circ \tau - f$, where $\tau : E \to E$ is an absolutely continuous mapping. Typically $E = \mathbb{N}$ and $\mathrm{d}^+ f(k) = f(k+1) - f(k)$, in this case d^+ can be used to express the surface measure of a set as the expectation of a discrete gradient norm. However, E can be a more general, even uncountable, space. The abstract case of a metric space has been considered in [5], [2] for a gradient having the derivation property.

In Section 2 we prove a discrete generalization of Cheeger's inequality [7], i.e. a lower bound on the spectral gap λ_μ in terms of the isoperimetry constant h_μ, using the arguments of [1] and [18]. When μ is the geometric distribution π on $\mathbb{N}$ with parameter $p \in (0,1)$ we show in Section 3.1 that $h_\pi = (1-p)/p$ and $\lambda_\pi = (1-\sqrt{p})^2/p$. The lower bound for λ_π obtained from Cheeger's inequality turns out to be optimal for the geometric distribution.

A measure μ is said to satisfy a logarithmic Sobolev inequality [12] with gradient d and constant $C > 0$ when

$$\mathrm{Ent}_\mu[f^2] \leq C \mathbb{E}_\mu \left[|\mathrm{d}f|^2 \right], \tag{1.2}$$

where $\mathrm{Ent}_\mu[f] = \mathbb{E}_\mu[f \log f] - \mathbb{E}_\mu[f] \log \mathbb{E}_\mu[f]$ denotes the entropy of f under μ. If the gradient $\mathrm{d}f$ has the derivation property, (1.2) is equivalent to the following modified logarithmic Sobolev inequality

$$\mathrm{Ent}_\mu[e^f] \leq \frac{C}{4} \mathbb{E}_\mu \left[|\mathrm{d}f|^2 e^f \right]. \tag{1.3}$$

Such modified inequalities have been established for Poisson and Bernoulli measures on $\mathbb{N}$ in [3], using the finite difference gradient d^+. On the other hand, modified logarithmic Sobolev inequalities for the exponential distribution have been obtained in [2] under the additional hypothesis that f is $c-$Lipschitz, i.e. $|\mathrm{d}f| \leq c$, when d has the derivation property.

In Section 3.2 we adapt the method of [2] to the geometric distribution, which can be viewed as a discrete analog of the exponential distribution, since the inter-jump times of the Poisson (resp. binomial) process have exponential (resp. geometric) distributions. For this we use an integration by parts formula and replace the derivation rule used in [2] with bounds on the finite difference gradient d^+, deduced from the mean value theorem. As noted in [9], the logarithmic Sobolev inequality does not hold for d^+ as stated in (1.3) under the geometric distribution with parameter p (take e.g. $f_a(n) = n \log a$ and let $a \nearrow 1/p$). We will show that (1.3) does hold for the geometric distribution under the further assumption $|\mathrm{d}^+ f| \leq c$, with a constant depending on c. In Section 3.3, using the Herbst method we obtain a deviation result for the geometric distribution, which differs from the deviation inequality recently obtained in [13] from the covariance representation method for infinitely divisible distributions. Although the integral part $[X]$ of an exponential random variable with parameter λ has a geometric law of parameter $e^{-\lambda}$, it does not seem possible to apply existing results on the exponential distribution [2] in our setting. For example when $f : \mathbb{N} \to \mathbb{R}$ is Lipschitz, $f([X])$ is not the composition of a Lipschitz function with X. However, exponential random variables can be approximated in distribution by geometric random variables, and in this way we recover the functional inequalities proved in [2] for the exponential distribution.

In Section 4 we obtain a more general result, stating that any distribution μ that satisfies a Poincaré inequality with constant λ_μ for a finite difference gradient also satisfies a logarithmic Sobolev inequality of modified type for all function f such that $|\mathrm{d}^+ f| \leq c$, which implies deviation bounds.

In Section 4.2 we present an exponential integrability criterion.

Let us mention that the results of this paper can be applied to an interacting spin system under a geometric reference measure, for which a logarithmic Sobolev inequality and a deviation inequality can be proved, extending the results established in [9] under Poisson reference measures.

Notation

Given a probability space $(E, \mathcal{E}, \mu)$, let $\tau : E \to E$ denote a map absolutely continuous with respect to μ. We denote by d^+ the finite difference gradient operator defined as

$$\mathrm{d}^+ f = f \circ \tau - f = \tau f - f,$$

where $f \circ \tau$ will be denoted by τf for shortness of notation. If $x \in \mathbb{R}$ is such that $\mu(f \geq x) \geq 1/2$ and $\mu(f \leq x) \geq 1/2$, we say that x is a median of f under μ, and write $m(f) = x$. We recall that for every median of f we have

$$\mathbb{E}_\mu \left[|f - m(f)| \right] = \inf_{a \in \mathbb{R}} \mathbb{E}_\mu \left[|f - a| \right].$$

We will need the co-area formula

$$\mathbb{E}_\mu \left[|\mathrm{d}^+ f| \right] = \int_{-\infty}^{+\infty} \mathbb{E}_\mu \left[|\mathrm{d}^+ 1_{\{f > t\}}| \right] dt, \tag{1.4}$$

which follows easily from the relations

$$(b - a)^\pm = \int_{-\infty}^{\infty} (1_{\{a > t\}} - 1_{\{b > t\}})^\pm dt, \quad a, b \in \mathbb{R}.$$

If $E = \mathbb{N}$ then it is natural to consider the shift $\tau f(k) = f(k+1)$, $k \in \mathbb{N}$, and the associated gradient

$$\mathrm{d}^+ f(k) = f(k+1) - f(k), \quad k \in \mathbb{N}. \tag{1.5}$$

Note that given $A \subset \mathbb{N}$ we have

$$\{|\mathrm{d}^+ 1_A| > 0\} = \{|\mathrm{d}^+ 1_A| \geq 1\} = \{k \in A \ : \ k+1 \in A^c\} \cup \{k \in A^c \ : \ k+1 \in A\},$$

i.e. $\{|\mathrm{d}^+ 1_A| > 0\}$ represents a frontier ∂A of A, and $\mathbb{E}_\mu[|\mathrm{d}^+ 1_A|]$ represents the measure of ∂A. Throughout this paper, μ denotes an arbitrary probability measure on E, while π denotes the geometric distribution with parameter $p \in (0, 1)$ on $E = \mathbb{N}$.

2 Isoperimetric and Poincaré inequalities

Given a measure μ on E, let h_μ denote the optimal constant in the inequality

$$h_\mu \mathbb{E}_\mu \left[|f - m(f)| \right] \leq \mathbb{E}_\mu \left[|\mathrm{d}^+ f| \right], \tag{2.1}$$

i.e.

$$h_\mu = \inf_{f \neq \text{const}} \frac{\mathbb{E}_\mu [|\mathrm{d}^+ f|]}{\mathbb{E}_\mu [|f - m(f)|]}.$$

Several analogs of Proposition 2.1 and Proposition 2.2 below have already been proved in [1], [15], [18], for connected graphs and for Markov chains, under reversibility or ergodicity assumptions. The gradient used in our setting is different but the proofs are similar and stated for completeness.

Proposition 2.1 *We have*

$$h_\mu = \inf_{0 < \mu(A) \leq \frac{1}{2}} \frac{\mathbb{E}_\mu [|\mathrm{d}^+ 1_A|]}{\mu(A)}. \tag{2.2}$$

Proof. We will prove the equality

$$h_\mu = \inf_{\mu(A) > 0} \frac{\mathbb{E}_\mu [|\mathrm{d}^+ 1_A|]}{\min(\mu(A), 1 - \mu(A))},$$

which clearly implies (2.2). Recall that $m(1_A) = 0$ if $\mu(A) \leq 1/2$, and $m(1_A) = 1$ if $\mu(A) \geq 1/2$, and $\mathbb{E}_\mu [|1_A - m(1_A)|] = \min(\mu(A), 1 - \mu(A))$. Let

$$h = \inf_{0 < \mu(A) \leq \frac{1}{2}} \frac{\mathbb{E}_\mu [|\mathrm{d}^+ 1_A|]}{\mu(A)}.$$

We have

$$h \mathbb{E}_\mu [|1_A - m(1_A)|] \leq \mathbb{E}_\mu \left[|\mathrm{d}^+ 1_A|\right], \quad A \in \mathcal{E}, \quad \mu(A) \leq 1/2.$$

From the co-area formula (1.4) we have, since $m(1_{\{f > t\}}) = 0$, $t \geq m(f)$, and $m(1_{\{f \leq t\}}) = 0$, $t \leq m(f)$:

$$\begin{aligned}
\mathbb{E}_\mu \left[|\mathrm{d}^+ f|\right] &= \int_{-\infty}^{+\infty} \mathbb{E}_\mu \left[|\mathrm{d}^+ 1_{\{f > t\}}|\right] dt \\
&\geq \int_{m(f)}^{+\infty} \mathbb{E}_\mu \left[|\mathrm{d}^+ 1_{\{f > t\}}|\right] dt + \int_{-\infty}^{m(f)} \mathbb{E}_\mu \left[|\mathrm{d}^+ 1_{\{f \leq t\}}|\right] dt \\
&\geq h \int_{m(f)}^{+\infty} \mathbb{E}_\mu \left[1_{\{f > t\}}\right] dt + h \int_{-\infty}^{m(f)} \mathbb{E}_\mu \left[1_{\{f \leq t\}}\right] dt \\
&= h \mathbb{E}_\mu \left[(f - m(f))^+\right] + h \mathbb{E}_\mu \left[(m(f) - f)^+\right] \\
&= h \mathbb{E}_\mu [|f - m(f)|],
\end{aligned}$$

hence $h \leq h_\mu$. This concludes the proof, since the converse inequality is obvious. $\square$

Let λ_μ denote the optimal constant in the Poincaré inequality

$$\lambda_\mu \mathrm{Var}_\mu[f] \leq \mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2\right], \tag{2.3}$$

under μ, i.e.

$$\lambda_\mu = \inf_{f \neq \mathrm{const}} \frac{\mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2\right]}{\mathrm{Var}_\mu\left[f\right]}.$$

The next result is a Cheeger type inequality, i.e. a lower bound on λ_μ which shows that the strict positivity of h_μ implies a Poincaré inequality.

Proposition 2.2 *We have*

$$\left(\sqrt{1+h_\mu}-1\right)^2 \leq \lambda_\mu \leq 2h_\mu. \tag{2.4}$$

Proof. Given a function f, let $g = f - m(f)$. We have $m(g) = 0$, which implies $m(g^{+^2}) = m(g^{-^2}) = 0$. Applying (2.1) to g^{+^2} and g^{-^2} we get

$$\begin{aligned}
h_\mu \mathbb{E}_\mu\left[g^2\right] &= h_\mu \mathbb{E}_\mu\left[g^{+^2}\right] + h_\mu \mathbb{E}_\mu\left[g^{-^2}\right] \\
&\leq \mathbb{E}_\mu\left[|\mathrm{d}^+ g^{+^2}| + |\mathrm{d}^+ g^{-^2}|\right] \\
&= \mathbb{E}_\mu\left[|2g^+ \mathrm{d}^+ g^+ + |\mathrm{d}^+ g^+|^2| + |2g^- \mathrm{d}^+ g^- + |\mathrm{d}^+ g^-|^2|\right] \\
&\leq 2\mathbb{E}_\mu\left[g^+ |\mathrm{d}^+ g^+| + g^- |\mathrm{d}^+ g^-|\right] + \mathbb{E}_\mu\left[|\mathrm{d}^+ g^+|^2 + |\mathrm{d}^+ g^-|^2\right] \\
&\leq 2\mathbb{E}_\mu\left[|g|\left(|\mathrm{d}^+ g^+| + |\mathrm{d}^+ g^-|\right)\right] + \mathbb{E}_\mu\left[|\mathrm{d}^+ g^+|^2 + |\mathrm{d}^+ g^-|^2\right] \\
&\leq 2\mathbb{E}_\mu\left[|g||\mathrm{d}^+ g|\right] + \mathbb{E}_\mu\left[|\mathrm{d}^+ g|^2\right] \\
&\leq 2\|g\|_2 \|\mathrm{d}^+ g\|_2 + \|\mathrm{d}^+ g\|_2^2,
\end{aligned}$$

where we used the relations $|\mathrm{d}^+ g^+| + |\mathrm{d}^+ g^-| = |\mathrm{d}^+ g|$ and $|\mathrm{d}^+ g^+|^2 + |\mathrm{d}^+ g^-|^2 \leq |\mathrm{d}^+ g|^2$. This implies

$$(\sqrt{1+h_\mu}-1)\|g\|_2 \leq \|\mathrm{d}^+ g\|_2.$$

In the general case we have

$$\begin{aligned}
(\sqrt{1+h_\mu}-1)^2 \mathrm{Var}_\mu[f] &= (\sqrt{1+h_\mu}-1)^2 \mathrm{Var}_\mu[g] \\
&\leq (\sqrt{1+h_\mu}-1)^2 \|g\|_2^2 \\
&\leq \mathbb{E}_\mu[|\mathrm{d}^+ g|^2] = \mathbb{E}_\mu[|\mathrm{d}^+ f|^2],
\end{aligned}$$

therefore $\lambda_\mu \geq \left(\sqrt{1+h_\mu}-1\right)^2$. Moreover we have

$$\lambda_\mu = \inf_{f\neq\text{const}} \frac{\mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2\right]}{\mathrm{Var}_\mu\left[f\right]} \leq \inf_{\substack{\emptyset\neq A\in\mathcal{E}\\ \mu(A)\leq 1/2}} \frac{\mathbb{E}_\mu\left[|\mathrm{d}^+ 1_A|\right]}{\mu(A)(1-\mu(A))} \leq 2h_\mu.$$

□

Note that (2.4) also yields an upper bound on h_μ in terms of λ_μ:

$$h_\mu \leq \lambda_\mu + 2\sqrt{\lambda_\mu}. \tag{2.5}$$

3 The geometric distribution

3.1 Optimal isoperimetric and Poincaré constants

We take $E = \mathbb{N}$ and the gradient

$$\mathrm{d}^+ f(k) = f(k+1) - f(k), \quad k \in \mathbb{N}.$$

Under π the Laplacian $\mathscr{L} = -\mathrm{d}_\pi^{+*}\mathrm{d}^+$ is given by

$$-\mathrm{d}_\pi^{+*}\mathrm{d}^+ f(k) = f(k+1) - f(k) + \frac{1}{p}1_{\{k\geq 1\}}(f(k-1)-f(k)),$$

i.e. $\mathscr{L} = \mathrm{d}^+ + \frac{1}{p}\mathrm{d}^-$ with

$$\mathrm{d}^- f(k) = 1_{\{k\geq 1\}}(f(k-1)-f(k)), \quad k \in \mathbb{N}.$$

Poincaré inequalities for general discrete distributions have been proved in [4], [6], [8], [17]. Theorem 1.3 in [4] shows in particular that a discrete distribution μ on $\mathbb{N}$ satisfies (2.3) if and only if

$$\mu(\{n\}) \geq c\mu([0,n])(1-\mu([0,n])), \quad n \geq 0,$$

for some constant $c > 0$. It is easily seen that the geometric distribution π with parameter $p \in (0,1)$ given by:

$$\pi(\{k\}) = p^k(1-p), \quad k \in \mathbb{N},$$

does satisfy this hypothesis. We now prove an isoperimetric inequality for the geometric distribution, which will imply a Poincaré inequality from Cheeger's inequality (2.4). The proof relies as in [2] on an integration by parts formula under π.

Lemma 3.1 *Let $f : \mathbb{N} \to \mathbb{R}$. We have*

$$\mathbb{E}_\pi[f] = f(0) + \frac{p}{1-p}\mathbb{E}_\pi[\mathrm{d}^+ f]. \tag{3.1}$$

Proof. Letting $g = f - f(0)$ we have the Radon-Nikodym type relation

$$\mathbb{E}_\pi[\tau g] = \frac{1}{p}\mathbb{E}_\pi[g], \tag{3.2}$$

since $g(0) = 0$, and

$$\mathbb{E}_\pi[\mathrm{d}^+ f] = \mathbb{E}_\pi[\mathrm{d}^+ g] = \mathbb{E}_\pi[\tau g] - \mathbb{E}_\pi[g] = \left(\frac{1}{p} - 1\right)\mathbb{E}_\pi[g] = \frac{1-p}{p}(\mathbb{E}_\pi[f] - f(0)).$$

$\square$

Note that to some extent, (3.2) characterizes the values $\pi(k)$ of the geometric distribution, $k \geq 1$, except for $\pi(0)$. Instead of d^+ we may use the gradient d^-, since similarly to the integration by parts formula we have the isometry

$$\mathbb{E}_\pi[N(\mathrm{d}^+ f)] = \frac{1}{p}\mathbb{E}_\pi[N(-\mathrm{d}^- f)],$$

for e.g. $N(x) = x$, $N(x) = |x|$, $N(x) = |x|^2$, which is equivalent to the reversibility of the birth and death process with generator $\mathscr{L} = -\mathrm{d}_\pi^{+*}\mathrm{d}^+$. In particular the gradient norm expectations generally used in the context of graphs and Markov chains [1], [15], [18], are here of the form

$$\mathbb{E}_\pi[N(\mathrm{d}^+ f)] + \frac{1}{p}\mathbb{E}_\pi[N(\mathrm{d}^- f)] = 2\mathbb{E}_\pi[N(\mathrm{d}^+ f)]$$

for $N(x) = |x|$, $N(x) = |x|^2$, and coincide with $\mathbb{E}_\pi[N(\mathrm{d}^+ f)]$ up to a constant factor.

Proposition 3.2 *Under the geometric distribution π we have*

$$h_\pi = \frac{1-p}{p}. \tag{3.3}$$

Proof. From the integration by parts formula (3.1) we have

$$\mathbb{E}_\pi\left[|f - m(f)|\right] \leq \mathbb{E}_\pi\left[|f - f(0)|\right] = \frac{p}{1-p}\mathbb{E}_\pi\left[\mathrm{d}^+|f - f(0)|\right] \leq \frac{p}{1-p}\mathbb{E}_\pi\left[|\mathrm{d}^+ f|\right],$$

which shows $h_\pi \geq (1-p)/p$. On the other hand, letting $f_n = 1_{[n+1,\infty)}$, $n \in \mathbb{N}$, we have for any $n \in \mathbb{N}$ such that $\pi([n+1,\infty)) \leq 1/2$:

$$h_\pi \leq \frac{\mathbb{E}_\pi\left[|\mathrm{d}^+ f|\right]}{\mathbb{E}_\pi\left[|f - m(f)|\right]} = \frac{\pi(\{n\})}{\pi([n+1,\infty))} = \frac{1-p}{p}.$$

$\square$

In particular, the isoperimetric inequality becomes an inequality for functions of the form $f_n = 1_{[n+1,\infty)}$, with $n \geq -\log 2/\log p$.

Proposition 3.3 *Under the geometric distribution π we have*

$$\lambda_\pi = \frac{(1-\sqrt{p})^2}{p}. \tag{3.4}$$

Proof. Using Cheeger's inequality (2.4) and Relation (3.3) we get $(1-\sqrt{p})^2/p \leq \lambda_\pi$. On the other hand, with $f_a(k) = a^k$ we have:

$$\lambda_\pi \leq \frac{\mathbb{E}_\pi\left[|\mathrm{d}^+ f_a|^2\right]}{\mathrm{Var}_\pi[f_a]} = (a-1)^2 \frac{a^2p^2 + 1 - 2ap}{a^2p + p - 2ap}, \quad a < 1/\sqrt{p},$$

and taking the limit as $a \to 1/\sqrt{p}$ we get $\lambda_\pi \leq (1-\sqrt{p})^2/p$. $\square$

Here, the lower bound on λ_π obtained from Cheeger's inequality coincides with the optimal Poincaré constant. The Poincaré inequality under π is not an equality in the linear case $f(k) = a + bk$:

$$\frac{(1-p)^2}{p}\mathrm{Var}_\pi[f] = \mathbb{E}_\pi\left[|\mathrm{d}^+ f|^2\right].$$

In fact, from Corollary 5.1 of [8], equality in the linear case holds only under the Poisson distribution.

Remark 3.4 *The lower bound of λ_π can be directly obtained from the integration by parts formula* (3.1) *under π.*

Proof. Letting $g = f - f(0)$ we have $g(0) = 0$ and from (3.1) applied to g^2 we obtain:

$$\begin{aligned}
\|g\|_2^2 &= \frac{p}{1-p}\mathbb{E}_\pi\left[\mathrm{d}^+(g^2)\right] \\
&= \frac{p}{1-p}\mathbb{E}_\pi\left[g\mathrm{d}^+ g + \tau g \mathrm{d}^+ g\right] \\
&\leq \frac{p}{1-p}\left(\|g\|_2\|\mathrm{d}^+ g\|_2 + \|\tau g\|_2\|\mathrm{d}^+ g\|_2\right) \\
&= \frac{p}{1-p}\left(\|g\|_2\|\mathrm{d}^+ f\|_2 + \frac{1}{\sqrt{p}}\|g\|_2\|\mathrm{d}^+ f\|_2\right) \\
&= \frac{\sqrt{p}}{1-\sqrt{p}}\|g\|_2\|\mathrm{d}^+ f\|_2,
\end{aligned}$$

hence

$$\|g\|_2 \leq \frac{\sqrt{p}}{1-\sqrt{p}} \|\mathrm{d}^+ f\|_2,$$

and

$$\frac{(1-\sqrt{p})^2}{p} \mathrm{Var}_\pi[f] = \frac{(1-\sqrt{p})^2}{p} \mathrm{Var}_\pi[g] \leq \frac{(1-\sqrt{p})^2}{p} \|g\|_2^2 \leq \mathbb{E}_\pi \left[|\mathrm{d}^+ f|^2 \right].$$

□

Using an approximation in distribution of exponential random variables by geometric random variables, we recover the Poincaré inequality of Lemma 2.1 in [2] for the exponential distribution with its optimal constant, cf. [11].

Proposition 3.5 *Let Y be an exponentially distributed random variable with parameter $-\log p$. We have*

$$\mathrm{Var}\left[f(Y)\right] \leq \frac{4}{(\log p)^2} \mathbb{E}\left[|f'(Y)|^2 \right]$$

for all Lipschitz function f on $\mathbb{R}$.

Proof. Let X_ε be a geometric random variable with parameter p^ε. We have

$$\mathrm{Var}\left[f(\varepsilon X_\varepsilon)\right] \leq \frac{\varepsilon^2 p^\varepsilon}{(1-\sqrt{p^\varepsilon})^2} \mathbb{E}\left[\left(\frac{f(\varepsilon X_\varepsilon + \varepsilon) - f(\varepsilon X_\varepsilon)}{\varepsilon} \right)^2 \right].$$

It remains to let ε go to 0 and to use the convergence of $\varepsilon X_\varepsilon$ in distribution to the exponential random variable Y with parameter $-\log p$. □

In a similar way, Proposition 3.2 yields an isoperimetric inequality under the exponential distribution with parameter $-\log p$:

$$\mathbb{E}[|f(Y) - m(f(Y))|] \leq -\frac{1}{\log p} \mathbb{E}\left[|f'(Y)|\right].$$

The above constants $h = -\log p$ and $\lambda = (\log p)^2/4$ also satisfy the classical Cheeger inequality $\lambda \geq h^2/4$ which holds in the continuous case, cf. [7].

3.2 Modified logarithmic Sobolev inequality

In this section we obtain a modified logarithmic Sobolev inequality for the geometric distribution π on $E = \mathbb{N}$, with $\mathrm{d}^+ f(k) = f(k+1) - f(k)$, $k \in \mathbb{N}$.

Lemma 3.6 *Let $c < -\log p$ and let $f : \mathbb{N} \to \mathbb{R}$ be such that $\mathrm{d}^+ f \leq c$ and $f(0) = 0$. We have*

$$\mathbb{E}_\pi \left[f^2 e^f \right] \leq \frac{pe^c}{(1-\sqrt{pe^c})^2} \mathbb{E}_\pi \left[e^f |\mathrm{d}^+ f|^2 \right]. \tag{3.5}$$

Proof. From the integration by parts formula (3.1) we have

$$\begin{aligned}
\mathbb{E}_\pi \left[f^2 e^f \right] &= \frac{p}{1-p} \mathbb{E}_\pi \left[\mathrm{d}^+ (f^2 e^f) \right] \\
&= \frac{p}{1-p} \mathbb{E}_\pi \left[e^f \left(e^{\mathrm{d}^+ f} \left(|\mathrm{d}^+ f|^2 + 2f \mathrm{d}^+ f \right) + f^2 (e^{\mathrm{d}^+ f} - 1) \right) \right] \\
&\leq \frac{pe^c}{1-p} \mathbb{E}_\pi \left[e^f \left(|\mathrm{d}^+ f|^2 + 2|f||\mathrm{d}^+ f| \right) \right] + \frac{p(e^c - 1)}{1-p} \mathbb{E}_\pi \left[f^2 e^f \right],
\end{aligned}$$

hence

$$\begin{aligned}
\mathbb{E}_\pi \left[f^2 e^f \right] &\leq \frac{pe^c}{1-pe^c} \mathbb{E}_\pi \left[e^f \left(|\mathrm{d}^+ f|^2 + 2|f||\mathrm{d}^+ f| \right) \right] \\
&\leq \frac{pe^c}{1-pe^c} \mathbb{E}_\pi \left[|\mathrm{d}^+ f|^2 e^f \right] + 2 \frac{pe^c}{1-pe^c} \mathbb{E}_\pi \left[f^2 e^f \right]^{1/2} \mathbb{E}_\pi \left[e^f |\mathrm{d}^+ f|^2 \right]^{1/2},
\end{aligned}$$

which implies (3.5). □

Theorem 3.7 *Let $0 < c < -\log p$ and let $f : \mathbb{N} \to \mathbb{R}$ such that $|\mathrm{d}^+ f| \leq c$. We have*

$$\mathrm{Ent}_\pi \left[e^f \right] \leq \frac{pe^c}{(1-p)(1-\sqrt{pe^c})} \mathbb{E}_\pi \left[|\mathrm{d}^+ f|^2 e^f \right]. \tag{3.6}$$

Proof. From the inequality $-u \ln u \leq 1 - u$, $u > 0$, we have:

$$\mathrm{Ent}_\pi \left[e^f \right] = \mathbb{E}_\pi \left[f e^f \right] - \mathbb{E}_\pi \left[e^f \right] \ln \mathbb{E}_\pi \left[e^f \right] \leq \mathbb{E}_\pi \left[f e^f - e^f + 1 \right]. \tag{3.7}$$

Let again $g = f - f(0)$, and let $h(v) = ve^v - e^v + 1$. We have $h \circ g(0) = 0$, and applying (3.1) to $h \circ g$ we get :

$$\begin{aligned}
\mathrm{Ent}_\pi \left[e^g \right] &\leq \mathbb{E}_\pi [h \circ g] \\
&= \frac{p}{1-p} \mathbb{E}_\pi [\mathrm{d}^+ (h \circ g)]
\end{aligned}$$

$$
\begin{aligned}
&= \frac{p}{1-p}\mathbb{E}_\pi[h\circ(g+\mathrm{d}^+g)-h\circ g] \\
&\leq \frac{p}{1-p}\mathbb{E}_\pi\left[\left(|\mathrm{d}^+g|^2+|g||\mathrm{d}^+g|\right)e^{g+|\mathrm{d}^+g|}\right],
\end{aligned}
$$

where the inequality

$$
h(a+b)-h(a)\leq(b^2+|ab|)e^{a+|b|}, \qquad a,b\in\mathbb{R},
$$

follows from the mean value theorem. From Lemma 3.6 and the Schwarz inequality we obtain:

$$
\begin{aligned}
\mathrm{Ent}_\pi\left[e^f\right] &= e^{f(0)}\mathrm{Ent}_\pi\left[e^g\right] \\
&\leq \frac{pe^{f(0)}}{1-p}\mathbb{E}_\pi\left[\left(|\mathrm{d}^+g|^2+|g||\mathrm{d}^+g|\right)e^{g+|\mathrm{d}^+g|}\right]. \\
&\leq \frac{pe^{c+f(0)}}{1-p}\left(\mathbb{E}_\pi\left[|\mathrm{d}^+g|^2e^g\right]+\mathbb{E}_\pi\left[g^2e^g\right]^{1/2}\mathbb{E}_\pi\left[e^g|\mathrm{d}^+g|^2\right]^{1/2}\right) \\
&\leq \frac{pe^{c+f(0)}}{1-p}\left(1+\frac{\sqrt{pe^c}}{1-\sqrt{pe^c}}\right)\mathbb{E}_\pi\left[|\mathrm{d}^+g|^2e^g\right] \\
&= \frac{pe^c}{(1-p)(1-\sqrt{pe^c})}\mathbb{E}_\pi\left[|\mathrm{d}^+f|^2e^f\right].
\end{aligned}
$$

□

In higher dimensions, consider the multi-dimensional gradient defined as

$$
\mathrm{d}_i^+f(k)=f(k+e_i)-f(k), \quad i=1,\ldots,n,
$$

where f is a function on $\mathbb{N}^n$, $k=(k_1,\ldots,k_n)\in\mathbb{N}^n$, $(e_1,\ldots,e_n)$ is the canonical basis of $\mathbb{R}^n$, and the gradient norm

$$
\|\mathrm{d}^+f(k)\|^2=\sum_{i=1}^n|\mathrm{d}_i^+f(k)|^2=\sum_{i=1}^n|f(k+e_i)-f(k)|^2. \tag{3.8}
$$

From the tensorization property of entropy, (3.6) still holds with respect to $\pi^{\otimes n}$ in any finite dimension n:

$$
\mathrm{Ent}_{\pi^{\otimes n}}\left[e^f\right]\leq\frac{pe^c}{(1-p)(1-\sqrt{pe^c})}\mathbb{E}_{\pi^{\otimes n}}\left[\|\mathrm{d}^+f\|^2e^f\right], \tag{3.9}
$$

provided $|d_if|\leq c$, $i=1,\ldots,n$ (we may also take $(1-p)^{-1}(1-\sqrt{pe^c})^{-1}$ as logarithmic Sobolev constant). Using an approximation in law of the exponential distribution by renormalized geometric random variables, we recover the logarithmic Sobolev inequality of Proposition 2.2 in [2].

Proposition 3.8 *Let Y be an exponentially distributed random variable with parameter $-\log p$. We have*

$$\mathrm{Ent}\left[e^{f(Y)}\right] \leq \frac{2}{(\log p)(\log(p)+c)} \mathbb{E}\left[e^{f(Y)}|f'(Y)|^2\right],$$

for every c-Lipschitz function f on $\mathbb{R}$.

Proof. We apply (3.6) to X_ε, and get for every c-Lipschitz function f:

$$\mathrm{Ent}_\pi\left[e^{f(\varepsilon X_\varepsilon)}\right] \leq \frac{\varepsilon^2 p^\varepsilon e^{\varepsilon c}}{(1-p^\varepsilon)(1-\sqrt{p^\varepsilon e^{\varepsilon c}})} \mathbb{E}_\pi\left[e^{f(\varepsilon X_\varepsilon)}\left(\frac{f(\varepsilon X_\varepsilon+\varepsilon)-f(\varepsilon X_\varepsilon)}{\varepsilon}\right)^2\right].$$

It remains to let ε go to 0. $\square$

3.3 Deviation inequality

In this section we prove a deviation inequality for functions of several variables under $\pi^{\otimes n}$ using the Herbst method and the above modified logarithmic Sobolev inequality.

Corollary 3.9 *Let $0 < c < -\log p$ and let f such that $|\mathrm{d}_i^+ f| \leq \beta$, $i = 1, \ldots, n$, and $\|\mathrm{d}^+ f\|^2 \leq \alpha^2$ for some $\alpha, \beta > 0$. Then for all $r > 0$,*

$$\pi^{\otimes n}(f - \mathbb{E}_{\pi^{\otimes n}}[f] \geq r) \leq \exp\left(-\min\left(\frac{c^2 r^2}{4a_{p,c}\alpha^2\beta^2}, \frac{rc}{\beta} - \alpha^2 a_{p,c}\right)\right), \tag{3.10}$$

where

$$a_{p,c} = \frac{pe^c}{(1-p)(1-\sqrt{pe^c})}$$

denotes the logarithmic Sobolev constant in (3.6).

Proof. Assume that $|\mathrm{d}_i^+ f| \leq c$, $i = 1, \ldots, n$. For $0 < t \leq 1$, let

$$H(t) = \frac{1}{t}\log \mathbb{E}_{\pi^{\otimes n}}[e^{tf}]$$

with $H(0^+) = \mathbb{E}_{\pi^{\otimes n}}[f]$. In order for $H(t)$ to be finite we may first assume that f is bounded, and then remove this assumption via a limiting argument once (3.10) is obtained. From (3.6) we have:

$$H'(t) = \frac{1}{t^2}\frac{\mathrm{Ent}_{\pi^{\otimes n}}[e^{tf}]}{\mathbb{E}_{\pi^{\otimes n}}[e^{tf}]} \leq \alpha^2 a_{p,c},$$

so that

$$H(t) \le \mathbb{E}_{\pi^{\otimes n}}[f] + t\alpha^2 a_{p,c},$$

hence

$$\mathbb{E}_{\pi^{\otimes n}}[e^{tf}] \le \exp\left(t\mathbb{E}_{\pi^{\otimes n}}[f] + t^2\alpha^2 a_{p,c}\right), \quad 0 < t \le 1. \tag{3.11}$$

Finally, using Chebychev's inequality we obtain from (3.11):

$$\begin{aligned}
\pi^{\otimes n}(f - \mathbb{E}_{\pi^{\otimes n}}[f] \ge r) &\le \inf_{t\in(0,1]} e^{-tr}\mathbb{E}_{\pi^{\otimes n}}[\exp(t(f - \mathbb{E}_{\pi^{\otimes n}}[f]))] \\
&\le \exp\left(\inf_{t\in(0,1]}(-tr + t^2\alpha^2 a_{p,c})\right) \\
&= \exp\left(-\min\left(\frac{r^2}{4\alpha^2 a_{p,c}}, r - \alpha^2 a_{p,c}\right)\right), \quad r > 0,
\end{aligned}$$

where we used the fact (see e.g. Corollary 2.11 in [16]) that the above minimum is attained at $t = \min(1, \frac{r}{2\alpha^2 a_{p,c}})$. Assume now that f satisfies $|\mathrm{d}_i^+ f| \le \beta$, $i = 1, \ldots, n$, for some $\beta > 0$. Then cf/β satisfies the above hypothesis and we get

$$\pi^{\otimes n}(f - \mathbb{E}_{\pi^{\otimes n}}[f] \ge r) \le \exp\left(-\min\left(\frac{c^2 r^2}{4a_{p,c}\alpha^2\beta^2}, \frac{rc}{\beta} - \alpha^2 a_{p,c}\right)\right).$$

□

Corollary 3.9 implies in particular $\mathbb{E}_\pi[e^{\alpha f}] < \infty$ for all $\alpha < c/\beta$ and $|\mathrm{d}^+ f| < c$. The condition $c < -\log p$ in Corollary 3.9 is necessary, since $f(k) = ck$ is not exponentially integrable under the geometric distribution π when $c \ge -\log p$. When $n = 1$, $\alpha = \beta$ and $r \ge 2c\beta a_{p,c}$, we have

$$\pi(f - \mathbb{E}_\pi[f] \ge r) \le \exp\left(-\frac{rc}{\beta} + c^2 a_{p,c}\right) \le \exp\left(-\frac{rc}{2\beta}\right), \tag{3.12}$$

and if $r \le 2c\beta a_{p,c}$:

$$\pi(f - \mathbb{E}_\pi[f] \ge r) \le \exp\left(-\frac{r^2}{4a_{p,c}\beta^2}\right).$$

These bounds can be compared to the result of [13] :

$$\pi(f - \mathbb{E}_\pi[f] \ge r) \le \left(1 + (1-p)\frac{r}{\beta}\right)\exp\left(-\left(\frac{r}{\beta} + \frac{p}{1-p}\right)\log\frac{p + p(1-p)r/\beta}{p + (1-p)r/\beta}\right),$$

$r > 0$, and to the exact deviation

$$\pi(X - \mathbb{E}_\pi[X] \ge r) = \exp\left(\left(\left[r + \frac{1}{1-p}\right] - 1\right)\log p\right)$$

for X a geometric random variable with parameter p, where $[x]$ denotes the integral part of $x \in \mathbb{R}$. Applying the inequality (3.10) to $-f$, we obtain the concentration inequality

$$\pi^{\otimes n}(|f - \mathbb{E}_\pi[f]| \geq r) \leq 2 \exp\left(-\min\left(\frac{c^2 r^2}{4a_{p,c}\alpha^2\beta^2}, \frac{rc}{\beta} - \alpha^2 a_{p,c}\right)\right). \qquad (3.13)$$

Consider the negative binomial distribution ν with parameters $n \geq 1$ and $p \in (0,1)$, defined as

$$\nu(\{k\}) = \binom{n+k-1}{n-1}(1-p)^n p^k, \quad k \in \mathbb{N}.$$

Negative binomial random variables can be constructed as sums of n independent and identically distributed geometric variables with parameter p. Therefore, if we apply (3.10) to

$$f(k_1, \ldots, k_n) = \phi(k_1 + \cdots + k_n), \quad (k_1, \ldots, k_n) \in \mathbb{N}^n,$$

we obtain the modified logarithmic Sobolev inequality

$$\mathrm{Ent}_\nu\left[e^\phi\right] \leq n a_{p,c} \mathbb{E}_\nu\left[|\mathrm{d}^+\phi|^2 e^\phi\right],$$

and the deviation inequality

$$\nu(\phi - \mathbb{E}_\nu[\phi] \geq r) \leq \exp\left(-\min\left(\frac{r^2}{4n a_{p,c}\beta^2}, \frac{rc}{\beta} - c^2 n a_{p,c}\right)\right),$$

where $\phi : \mathbb{N} \to \mathbb{R}$ satisfies $|d\phi| \leq \beta$, for the negative binomial distribution μ. Similar results can be obtained for the tensor product of negative binomial laws with parameters $n_1, \ldots, n_d$, namely by replacing $a_{p,c} = \frac{pe^c}{(1-p)(1-\sqrt{pe^c})}$ with $(n_1 + \cdots + n_d)a_{p,c}$ in (3.6) and (3.10).

Geometric and negative binomial random variables can be constructed as hitting times of the binomial process, thus they can be viewed as random variables on Bernoulli space. However, applying to them the Poincaré and logarithmic Sobolev inequalities on Bernoulli space (see e.g. [14]) yields results that are weaker than the above inequalities.

4 The abstract case

In this section, we turn again to the general case of a probability space $(E, \mathcal{E}, \mu)$ with an absolutely continuous mapping $\tau : E \to E$. We show that modified logarithmic

Sobolev and deviation inequalities hold for every measure μ on E which satisfies a Poincaré inequality

$$\lambda_\mu \mathrm{Var}_\mu[f] \leq \mathbb{E}_\mu \left[|\mathrm{d}^+ f|^2 \right] \tag{4.1}$$

with respect to d^+, i.e. for every measure μ such that $\lambda_\mu > 0$. The application of the general results of this section to the geometric distribution using the spectral gap value (3.4) of λ_π allow to recover the inequalities of Section 3.2. However, explicit calculations show that the results are recovered with worse constants for all $p \in (0, e^{-c})$, especially as p approaches e^{-c}.

4.1 Logarithmic Sobolev inequality and deviation inequality

Before turning to the main result of this section, we need the two following propositions whose proofs are adapted from [2], replacing the chain rule of derivation by the mean value theorem, and postponed to the end of this section. The next proposition is a generalization of Lemma 3.6.

Proposition 4.1 *Let $c > 0$. For any f on E such that $|\mathrm{d}^+ f| \leq c$ with $c^2 e^c \leq 4\lambda_\mu$ and $\mathbb{E}_\mu[f] = 0$,*

$$\mathbb{E}_\mu \left[f^2 e^f \right] \leq \alpha_{\mu,c} \mathbb{E}_\mu \left[|\mathrm{d}^+ f|^2 e^f \right], \tag{4.2}$$

where $\alpha_{\mu,c} = \frac{e^c\left((2+c)\sqrt{\lambda_\mu}+c\right)^2}{\lambda_\mu\left(2\sqrt{\lambda_\mu}-ce^{c/2}\right)^2}$.

The next statement is a modification of Proposition 3.4 in [2].

Proposition 4.2 *For any $f : E \to \mathbb{R}$ such that $\mathbb{E}_\mu[f] = 0$ and $|\mathrm{d}^+ f| \leq c$ we have*

$$\mathbb{E}_\mu \left[f^2 + \tau f^2 \right] \leq e^{c\left(1+\sqrt{\frac{5}{\lambda_\mu}}\right)} \mathbb{E}_\mu \left[\left(f^2 + \tau f^2 \right) e^{-|f|} \right]. \tag{4.3}$$

The following is a modified logarithmic Sobolev inequality which holds whenever $\lambda_\mu > 0$.

Theorem 4.3 *Assume that $f : E \to \mathbb{R}$ satisfies $|\mathrm{d}^+ f| \leq c$ with $c^2 e^c \leq 4\lambda_\mu$,*

$$\mathrm{Ent}_\mu \left[e^f \right] \leq \frac{1}{2} e^{c\left(1+\sqrt{\frac{5}{\lambda_\mu}}\right)} \mathbb{E}_\mu \left[(\alpha_{\mu,c} |\mathrm{d}^+ f|^2 + 2e^{2c} \alpha_{\mu,c} |\mathrm{d}^+ \tau f|^2 + 2e^{2c} \|\mathrm{d}^+ f\|^2_{L^2(\mu)}) e^f \right]. \tag{4.4}$$

Proof. It suffices to suppose $\mathbb{E}_\mu[f]=0$. From the inequality $x\log x \geq x-1$, $x>0$, we have

$$\mathrm{Ent}_\mu\left[e^f\right] \leq \mathbb{E}_\mu\left[fe^f - e^f + 1\right] = \mathbb{E}_\mu\left[\int_0^1 tf^2 e^{tf} dt\right],$$

hence

$$\mathrm{Ent}_\mu\left[e^f\right] \leq \int_0^1 t\varphi(t)dt,$$

where $\varphi(t) = \mathbb{E}_\mu\left[\left(f^2+\tau f^2\right)e^{tf}\right]$, $0 \leq t \leq 1$, is a convex function which satisfies $\varphi(t) \leq \max(\varphi(0), \varphi(1))$, $0 \leq t \leq 1$. Moreover, by Proposition 4.2 we have

$$\varphi(0) \leq e^{c\left(1+\sqrt{\frac{5}{\lambda_\mu}}\right)}\varphi(1),$$

hence

$$\mathrm{Ent}_\mu\left[e^f\right] \leq \int_0^1 te^{c\left(1+\sqrt{\frac{5}{\lambda_\mu}}\right)}\varphi(1)dt = \frac{1}{2}e^{c\left(1+\sqrt{\frac{5}{\lambda_\mu}}\right)}\mathbb{E}_\mu\left[\left(f^2+\tau f^2\right)e^f\right].$$

Since $|\mathrm{d}^+\left(\tau f - \mathbb{E}_\mu[\tau f]\right)| = |\mathrm{d}^+\tau f| \leq c$, Proposition 4.1 applied to $\tau f - \mathbb{E}_\mu[\tau f]$ implies:

$$\begin{aligned}
\mathbb{E}_\mu\left[\tau f^2 e^f\right] &\leq e^c\mathbb{E}_\mu\left[\tau f^2 e^{\tau f}\right] \\
&\leq 2e^{c+\mathbb{E}_\mu[\tau f]}\mathbb{E}_\mu\left[\left(\tau f - \mathbb{E}_\mu[\tau f]\right)^2 e^{\tau f - \mathbb{E}_\mu[\tau f]}\right] + 2e^c(\mathbb{E}_\mu[\tau f])^2\mathbb{E}_\mu\left[e^{\tau f}\right] \\
&\leq 2e^c\alpha_{\mu,c}\mathbb{E}_\mu\left[|\mathrm{d}^+\tau f|^2 e^{\tau f}\right] + 2e^{2c}(\mathbb{E}_\mu[\tau f])^2\mathbb{E}_\mu\left[e^f\right] \\
&= 2e^{2c}\alpha_{\mu,c}\mathbb{E}_\mu\left[|\mathrm{d}^+\tau f|^2 e^f\right] + 2e^{2c}(\mathbb{E}_\mu[\mathrm{d}^+ f])^2\mathbb{E}_\mu\left[e^f\right].
\end{aligned}$$

Hence

$$\begin{aligned}
\mathrm{Ent}_\mu\left[e^f\right] &\leq \frac{1}{2}e^{c\left(1+\sqrt{\frac{5}{\lambda_\mu}}\right)}\mathbb{E}_\mu\left[\left(f^2+\tau f^2\right)e^f\right] \\
&\leq \frac{1}{2}e^{c\left(1+\sqrt{\frac{5}{\lambda_\mu}}\right)}\mathbb{E}_\mu\left[(\alpha_{\mu,c}|\mathrm{d}^+ f|^2 + 2e^{2c}\alpha_{\mu,c}|\mathrm{d}^+\tau f|^2 + 2e^{2c}\mathbb{E}_\mu[|\mathrm{d}^+ f|^2])e^f\right].
\end{aligned}$$

□

We also have

$$\mathrm{Ent}_\mu\left[e^f\right] \leq \frac{1}{2}e^{c\left(1+\sqrt{\frac{5}{\lambda_\mu}}\right)}(\alpha_{\mu,c} + 2e^{2c}\alpha_{\mu,c} + 2e^{2c})|\mathrm{d}^+ f|_\infty\mathbb{E}_\mu[e^f], \qquad |\mathrm{d}^+ f| \leq c.$$

By tensorization, Theorem 4.3 implies in higher dimensions

$$\mathrm{Ent}_{\mu^{\otimes n}}\left[e^f\right] \leq \frac{1}{2}e^{c\left(1+\sqrt{\frac{5}{\lambda_\mu}}\right)}\mathbb{E}_\mu\left[(\alpha_{\mu,c}\|\mathrm{d}^+ f\|^2 + 2e^{2c}\alpha_{\mu,c}\|\mathrm{d}^+\tau f\|^2 + 2e^{2c}\|\mathrm{d}^+ f\|^2_{L^2(E;\mathbb{R}^n)})e^f\right]$$

$$\leq \quad m_{\mu,c} \|\mathrm{d}^+ f\|^2_{L^\infty(E^n,\mathbb{R}^n)} \mathbb{E}_\mu \left[e^f \right],$$

where

$$m_{\mu,c} = \frac{1}{2} e^{c\left(1+\sqrt{\frac{5}{\lambda_\mu}}\right)} \left(\alpha_{\mu,c} + 2e^{2c}\alpha_{\mu,c} + 2e^{2c}\right)$$

and

$$\begin{aligned} \mathrm{d}_i^+ f(x_1,\ldots,x_n) &= \tau_i f(x_1,\ldots,x_n) - f(x_1,\ldots,x_n), \\ &= f(x_1,\ldots,x_{i-1},\tau_i(x_i),x_{i+1},\ldots,x_n) - f(x_1,\ldots,x_n), \end{aligned}$$

provided $|\mathrm{d}_i^+ f| \leq c$, $i = 1,\ldots,n$. As in Section 3.2, we obtain as a corollary a deviation inequality for the product measure $\mu^{\otimes n}$ on E^n:

Corollary 4.4 *Assume that μ satisfies a Poincaré inequality (4.1). Let $c > 0$ such that $c^2 e^c \leq 4\lambda_\mu$, and let f such that $|\mathrm{d}_i^+ f| \leq \beta$, $i = 1,\ldots,d$, and $\|\mathrm{d}^+ f\|^2 \leq \alpha^2$, for some $\alpha, \beta > 0$. Then for all $r > 0$,*

$$\mu^{\otimes n}\left(f - \mathbb{E}_\mu^{\otimes n}[f] \geq r\right) \leq \exp\left(-\min\left(\frac{c^2 r^2}{4 m_{\mu,c}\alpha^2\beta^2}, \frac{rc}{\beta} - \alpha^2 m_{\mu,c}\right)\right). \qquad (4.5)$$

Next we provide the proofs of Proposition 4.1 and Proposition 4.2.

Proof of Proposition 4.1. Set $a^2 = \mathbb{E}_\mu\left[f^2 e^f\right]$ and $b^2 = \mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2 e^f\right]$. Since $\mathbb{E}_\mu[f] = 0$, the Poincaré inequality (4.1) implies

$$\begin{aligned} \lambda_\mu^2 \mathbb{E}_\mu\left[f e^{f/2}\right]^2 &\leq \mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2\right] \mathbb{E}_\mu\left[|\mathrm{d}^+\left(e^{f/2}\right)|^2\right] \\ &\leq \frac{1}{4}\mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2\right] \mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2 e^{f+|\mathrm{d}^+ f|}\right] \\ &\leq \frac{1}{4} e^c\, c^2\, b^2. \end{aligned} \qquad (4.6)$$

Applying again the Poincaré inequality to $f e^{f/2}$ and using the mean value theorem we have

$$\begin{aligned} \lambda_\mu \mathrm{Var}\left[f e^{f/2}\right] &\leq \mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2 \left(1 + \frac{|f| + |\mathrm{d}^+ f|}{2}\right)^2 e^{f+|\mathrm{d}^+ f|}\right] \\ &\leq e^c\, \mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2 \left(1 + \frac{|f| + c}{2}\right)^2 e^f\right] \\ &\leq \left(1 + \frac{c}{2}\right)^2 e^c b^2 + \frac{c^2 e^c a^2}{4} + \left(1 + \frac{c}{2}\right) e^c \mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2 |f| e^f\right] \end{aligned}$$

$$\begin{aligned}
&\le \left(1+\frac{c}{2}\right)^2 e^c b^2 + \frac{c^2 e^c a^2}{4} + \left(1+\frac{c}{2}\right) e^c abc \\
&\le e^c \left(\left(1+\frac{c}{2}\right) b + \frac{ac}{2}\right)^2.
\end{aligned}$$

Hence

$$a^2 = \mathbb{E}_\mu \left[f e^{f/2}\right]^2 + \mathrm{Var}\left[f e^{f/2}\right] \le \frac{e^c\, c^2\, b^2}{4\lambda_\mu^2} + \frac{e^c}{\lambda_\mu}\left(\left(1+\frac{c}{2}\right) b + \frac{ac}{2}\right)^2,$$

which leads to

$$a \le \frac{e^{c/2}\left((2+c)\sqrt{\lambda_\mu} + c\right)}{\sqrt{\lambda_\mu}\left(2\sqrt{\lambda_\mu} - c e^{c/2}\right)},$$

from which the conclusion follows. □

With $\lambda_\pi = (1-\sqrt{p})^2/p$, the condition $c^2 e^c \le 4\lambda_\pi$ implies $c < -\log p$, $p \in (0,1)$, hence Theorem 4.3 and Corollary 4.4 are weaker than Theorem 3.7 and Corollary 3.9 respectively, when $\mu = \pi$ is the geometric distribution.

Proof of Proposition 4.2. We have from the Poincaré inequality (4.1):

$$\begin{aligned}
\lambda_\mu \mathbb{E}_\mu\left[f^4\right] &= \lambda_\mu \mathrm{Var}_\mu\left[f^2\right] + \lambda_\mu (\mathbb{E}_\mu\left[f^2\right])^2 \\
&\le \mathbb{E}_\mu\left[|\mathrm{d}^+ f^2|^2\right] + \lambda_\mu (\mathbb{E}_\mu\left[f^2\right])^2 \\
&= \mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2 (f + \tau f)^2\right] + \lambda_\mu (\mathbb{E}_\mu\left[f^2\right])^2 \\
&\le 2c^2 \mathbb{E}_\mu\left[f^2 + \tau f^2\right] + \mathbb{E}_\mu\left[|\mathrm{d}^+ f|^2\right] \mathbb{E}_\mu\left[f^2\right] \\
&\le 3c^2 \mathbb{E}_\mu\left[f^2\right] + 2c^2 \mathbb{E}_\mu\left[\tau f^2\right].
\end{aligned}$$

Hence for all $u > 0$,

$$\mathbb{E}_\mu\left[|f|^3\right] \le \frac{u}{2}\mathbb{E}_\mu\left[f^2\right] + \frac{1}{2u}\mathbb{E}_\mu\left[f^4\right] \le c_1 \mathbb{E}_\mu\left[f^2\right] + c_2 \mathbb{E}_\mu\left[\tau f^2\right] \tag{4.7}$$

with $c_1 = \frac{3c^2}{2u\lambda_\mu} + \frac{u}{2}$ and $c_2 = \frac{c^2}{u\lambda_\mu}$. Let us consider the probability measure

$$d\rho = \frac{c_1 f^2 + c_2 \tau f^2}{c_1 \mathbb{E}_\mu\left[f^2\right] + c_2 \mathbb{E}_\mu\left[\tau f^2\right]} d\mu.$$

By Jensen's inequality,

$$\mathbb{E}_\mu\left[\left(c_1 f^2 + c_2 \tau f^2\right) e^{-|f|}\right] = \mathbb{E}_\rho\left[e^{-|f|}\right] \mathbb{E}_\mu\left[c_1 f^2 + c_2 \tau f^2\right]$$

$$\geq \quad e^{-\mathbb{E}_\rho[|f|]}\mathbb{E}_\mu\left[c_1 f^2 + c_2 \tau f^2\right].$$

From the inequality $ab^2 \leq a^3 + |b-a|(a^2+b^2)$, $a, b \geq 0$, we have

$$|f|\tau f^2 \leq |f|^3 + c\left(f^2 + \tau f^2\right),$$

hence

$$\begin{aligned}
\mathbb{E}_\rho\left[|f|\right]\mathbb{E}_\mu\left[c_1 f^2 + c_2\tau f^2\right] &= \mathbb{E}_\mu\left[c_1|f|^3 + c_2|f|\tau f^2\right] \\
&\leq \mathbb{E}_\mu\left[(c_1+c_2)|f|^3 + c_2 c\left(f^2 + \tau f^2\right)\right] \\
&\leq (c_1+c_2)\mathbb{E}_\mu\left[c_1 f^2 + c_2\tau f^2\right] + c_2 c\mathbb{E}_\mu\left[f^2 + \tau f^2\right] \\
&\leq (c_1+c_2+c)\mathbb{E}_\mu\left[c_1 f^2 + c_2\tau f^2\right],
\end{aligned}$$

where we used the fact that $c_2 \leq c_1$. Therefore,

$$\mathbb{E}_\rho\left[|f|\right] \leq c_1 + c_2 + c = \frac{5c^2 + u^2\lambda_\mu}{2u\lambda_\mu} + c.$$

Optimizing in u we obtain for $u = c\sqrt{\frac{5}{\lambda_\mu}}$:

$$\mathbb{E}_\rho\left[|f|\right] \leq c\left(1 + \sqrt{\frac{5}{\lambda_\mu}}\right).$$

□

As in [2] and references therein, we can obtain the following bound.

Proposition 4.5 *Let A, B be disjoint subsets of E. We have*

$$\mu(A)\mu(B) \leq 3\exp(-\sqrt{\lambda_\mu}e^{-\gamma_1/2}d(A,B)), \tag{4.8}$$

with $\gamma_1^2 e^{\gamma_1} = 2\lambda_\mu$.

Proof. From the Poincaré inequality on $(E^2, \mu^{\otimes 2})$ we have:

$$\lambda_\mu \mathbb{E}_{\mu^{\otimes 2}}[f^2] \leq \mathbb{E}_{\mu^{\otimes 2}}[|\mathrm{d}_1^+ f|^2 + |\mathrm{d}_2^+ f|^2],$$

provided $\mathbb{E}_{\mu^{\otimes 2}}[f] = 0$. Applying this inequality to $f(x,y) = \sinh(tg(x,y)/2)$, $0 \leq t < \gamma_1$ with $g(x,y) = h(x) - h(y)$ and $|\mathrm{d}h| \leq 1$, and using the bound

$$|\sinh(x+y) - \sinh(x)| \leq |y|e^{|y|}\cosh x, \quad x, y \in \mathbb{R},$$

we have:

$$\begin{aligned}\lambda_\mu \mathbb{E}_{\mu^{\otimes 2}}[\cosh^2(tg/2)] - \lambda_\mu &= \lambda_\mu \mathbb{E}_{\mu^{\otimes 2}}[\sinh^2(tg/2)] \\ &\leq \frac{t^2}{4}\mathbb{E}_{\mu^{\otimes 2}}\left[\left(|\mathrm{d}_1^+ g|^2 e^{t|\mathrm{d}_1^+ g|} + |\mathrm{d}_2^+ g|^2 e^{t|\mathrm{d}_2^+ g|}\right)\cosh^2(tg/2)\right] \\ &\leq \frac{t^2}{2} e^{\gamma_1}\mathbb{E}_{\mu^{\otimes 2}}\left[\cosh^2(tg/2)\right].\end{aligned}$$

Hence

$$\mathbb{E}_{\mu^{\otimes 2}}[\cosh^2(tg/2)] = \frac{1}{2}\left(\mathbb{E}_{\mu^{\otimes 2}}[e^{tg}] + 1\right) \leq \frac{2\lambda_\mu}{2\lambda_\mu - t^2 e^{\gamma_1}},$$

and for all $t < \gamma_1$, if $h(x) = d(x, B)$ then

$$e^{td(A,B)}\mu(A)\mu(B) \leq \mathbb{E}_{\mu^{\otimes 2}}[1_{A\times B} e^{tg}] \leq \mathbb{E}_{\mu^{\otimes 2}}[e^{tg}] \leq \frac{2\lambda_\mu + t^2 e^{\gamma_1}}{2\lambda_\mu - t^2 e^{\gamma_1}}.$$

and it remains to take $t = \sqrt{\lambda_\mu} e^{-\gamma_1/2}$. □

4.2 Exponential integrability

The Herbst method used in the preceding sections relies on exponential integrability. Following [2], we obtain a bound of the Laplace transform with respect to any measure μ on E, provided it follows a Poincaré inequality (4.1).

Proposition 4.6 *Let $f : E \to \mathbb{R}$ such that $\mathbb{E}_\mu[f] = 0$, with $|\mathrm{d}^+ f| \leq \beta$ for some $\beta > 0$, and let c such that $c^2 e^c \leq 4\lambda_\mu$. Then, for every $0 \leq t < c/\beta$ we have*

$$\mathbb{E}_\mu[e^{tf}] \leq \frac{2\sqrt{\lambda_\mu} + t\beta e^{c/2}}{2\sqrt{\lambda_\mu} - t\beta e^{c/2}}. \tag{4.9}$$

Proof. We adapt the proof of Proposition 4.1 in [2]. It is sufficient to assume $\beta = 1$. We have

$$\begin{aligned}|\mathrm{d}^+ e^{\frac{t}{2}f}(x)| &= |e^{\frac{t}{2}\tau f(x)} - e^{\frac{t}{2}f(x)}| \\ &= \frac{t}{2}\left|\int_{f(x)}^{f(\tau(x))} e^{\frac{t}{2}} dt\right| \\ &\leq \frac{t}{2} e^{\frac{t}{2}(f(x) + |\mathrm{d}^+ f(x)|)} |\mathrm{d}^+ f(x)| \\ &\leq \frac{t}{2} e^{\frac{c}{2} + \frac{t}{2}f(x)} |\mathrm{d}^+ f(x)|, \qquad x \in E,\end{aligned}$$

and applying (4.1) to $e^{\frac{t}{2}f}$ we get, with $u(t) = \mathbb{E}_\mu[e^{tf}]$:

$$\lambda_\mu \left(u(t) - u(t/2)^2\right) \leq e^c \frac{t^2}{4} u(t),$$

i.e.

$$u(t) \leq \frac{4\lambda_\mu}{4\lambda_\mu - t^2 e^c}\, u(t/2)^2.$$

Applying the same inequality for $t/2$ and iterating, we have

$$u(t) \leq \prod_{k=0}^{\infty} \left(\frac{4\lambda_\mu}{4\lambda_\mu - e^c t^2/4^k} \right)^{2^k} \leq \frac{4\lambda_\mu}{4\lambda_\mu - e^c t^2}\, V(t),$$

with

$$V(t) = \prod_{k=1}^{\infty} \left(\frac{4\lambda_\mu}{4\lambda_\mu - e^c t^2/4^k} \right)^{2^k},$$

where the product converges whenever $t < c$. It can be shown as in [2] that $\sqrt{V}$ is convex. Moreover $V(0) = 1$ and $V\left(\frac{2\sqrt{\lambda_\mu}}{e^{c/2}}\right) \leq 4$, hence

$$\sqrt{V(t)} \leq \frac{2\sqrt{\lambda_\mu} + t e^{c/2}}{2\sqrt{\lambda_\mu}}.$$

□

It is easily checked that the assumption of Corollary 4.4 is consistent with that of Proposition 4.6.

Acknowledgements

We thank Djalil Chafaï and Christian Houdré for useful discussions.

References

[1] S. Bobkov, C. Houdré, and P. Tetali. λ_∞, vertex isoperimetry and concentration. *Combinatorica*, 20(2):153–172, 2000.

[2] S. Bobkov and M. Ledoux. Poincaré's inequalities and Talagrand's concentration phenomenon for the exponential distribution. *Probab. Theory Related Fields*, 107(3):383–400, 1997.

[3] S. G. Bobkov and M. Ledoux. On modified logarithmic Sobolev inequalities for Bernoulli and Poisson measures. *J. Funct. Anal.*, 156(2):347–365, 1998.

[4] S.G. Bobkov and F. Götze. Discrete isoperimetric and Poincaré-type inequalities. *Probab. Theory Related Fields*, 114(2):245–277, 1999.

[5] S.G. Bobkov and C. Houdré. Isoperimetric constants for product probability measures. *Ann. Probab.*, 25(1):184–205, 1997.

[6] T. Cacoullos and V. Papathanasiou. Characterizations of distributions by generalizations of variance bounds and simple proofs of the CLT. *J. Statist. Plann. Inference*, 63(2):157–171, 1997.

[7] J. Cheeger. A lower bound for the smallest eigenvalue of the Laplacian. In *Problems in analysis (Papers dedicated to Salomon Bochner, 1969)*, pages 195–199. Princeton Univ. Press, Princeton, N. J., 1970.

[8] L. H. Y. Chen and J. H. Lou. Characterization of probability distributions by Poincaré-type inequalities. *Ann. Inst. H. Poincaré Probab. Statist.*, 23(1):91–110, 1987.

[9] P. Dai Pra, A.M. Paganoni, and G. Posta. Entropy inequalities for unbounded spin systems. *Ann. Probab.*, 30(4):1959–1976, 2002.

[10] P. Diaconis and D. Stroock. Geometric bounds for eigenvalues of Markov chains. *Ann. Appl. Probab.*, 1(1):36–61, 1991.

[11] P. Fougères. Spectral gap for log-concave probability measures on the real line. Preprint, 2002.

[12] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4):1061–1083, 1975.

[13] C. Houdré. Remarks on deviation inequalities for functions of infinitely divisible random vectors. *Ann. Probab.*, 30(3):1223–1237, 2002.

[14] C. Houdré and N. Privault. Concentration and deviation inequalities in infinite dimensions via covariance representations. *Bernoulli*, 8(6):697–720, 2002.

[15] C. Houdré and P. Tetali. Isoperimetric invariants for product Markov chains and graph products. Combinatorica, to appear, 2003.

[16] M. Ledoux. Concentration of measure and logarithmic Sobolev inequalities. In *Séminaire de Probabilités XXXIII*, volume 1709 of *Lecture Notes in Math.*, pages 120–216. Springer, 1999.

[17] L. Miclo. An example of application of discrete Hardy's inequalities. *Markov Process. Related Fields*, 5(3):319–330, 1999.

[18] T. Stoyanov. Isoperimetric and related constants for graphs and Markov chains. PhD thesis, Georgia Institute of Technology, 2001.