

On Modified Logarithmic Sobolev Inequalities for Bernoulli and Poisson Measures

S. G. Bobkov*

*Department of Mathematics, Syktyvkar University,
167001 Syktyvkar, Russia*

and

M. Ledoux

*Département de Mathématiques, Laboratoire de Statistique et Probabilités associé au C.N.R.S.,
Université Paul-Sabatier, 31062 Toulouse, France*

Received July 7, 1997; revised November 6, 1997; accepted November 6, 1997

We show that for any positive function f on the discrete cube $\{0, 1\}^n$,

$$\text{Ent}_{\mu_p^n}(f) \leq pq E_{\mu_p^n} \left(\frac{1}{f} |Df|^2 \right)$$

where μ_p^n is the product measure of the Bernoulli measure with probability of success p , as well as related inequalities, which may be shown to imply in the limit the classical Gaussian logarithmic Sobolev inequality as well as a logarithmic Sobolev inequality for Poisson measure. We further investigate modified logarithmic Sobolev inequalities to analyze integrability properties of Lipschitz functions on discrete spaces. In particular, we obtain, under modified logarithmic Sobolev inequalities, some concentration results for product measures that extend the classical exponential inequalities for sums of independent random variables.

© 1998 Academic Press

1. INTRODUCTION

In his seminal 1975 paper, L. Gross [G] proved a logarithmic Sobolev inequality on the two-point space. Namely, let μ be the uniform measure on $\{0, 1\}$. Then, for any f on $\{0, 1\}$,

$$\int f^2 \log f^2 d\mu - \int f^2 d\mu \log \int f^2 d\mu \leq \frac{1}{2} \int |Df|^2 d\mu, \quad (1)$$

* Partially supported by the Alexander von Humboldt Foundation.

where $Df(x) = f(x+1) - f(x)$ (modulo 2). It is easily seen that the constant $\frac{1}{2}$ is optimal. In its hypercontractive form, this inequality was first established by A. Bonami [Bo].

The question of the best constant in the previous logarithmic Sobolev inequality for non-symmetric Bernoulli measure was settled seemingly only quite recently. Let μ_p be the Bernoulli measure on $\{0, 1\}$ with $\mu_p(\{1\}) = p$ and $\mu_p(\{0\}) = q = 1 - p$. Then, for any f on $\{0, 1\}$,

$$\int f^2 \log f^2 d\mu_p - \int f^2 d\mu_p \log \int f^2 d\mu_p \leq pq \frac{\log p - \log q}{p - q} \int |Df|^2 d\mu_p. \quad (2)$$

The constant is optimal, and is equal to $\frac{1}{2}$, when $p = q = \frac{1}{2}$. This result is mentioned in [H-Y] without proof, and worked out in [D-SC]. A simple proof, due to the first named author, is presented in the notes [SC]. O. Rothaus mentioned to the authors of [D-SC] that he computed this constant several years back from now. The main feature of this constant is that, when $p \neq q$, it significantly differs from the spectral gap given by the inequality

$$\int f^2 d\mu_p - \left(\int f d\mu_p \right)^2 \leq pq \int |Df|^2 d\mu_p.$$

Although inequality (2) is optimal, it presents a number of weak points. First of all, the product property of entropy allowed L. Gross to tensorize inequality (1) to deduce with the central limit theorem the basic logarithmic Sobolev inequality for the canonical Gaussian measure γ on \mathbb{R}

$$\int f^2 \log f^2 d\gamma - \int f^2 d\gamma \log \int f^2 d\gamma \leq 2 \int f'^2 d\gamma \quad (3)$$

(for every smooth f on \mathbb{R}). This is however only optimal in the symmetric case and, as soon as $p \neq q$, the central limit theorem on the basis of (2) only yields the Gaussian logarithmic Sobolev inequality with a worse constant going to infinity as $p \rightarrow 0$ or 1. A second limit theorem of interest is of course the Poisson limit. However, after tensorization, the inequality (2) cannot yield a logarithmic Sobolev inequality for Poisson measures. There is of course a good reason at that, namely that Poisson measure do not satisfy logarithmic Sobolev inequalities! This is well-known to a number of people but let us briefly convince ourselves of this claim. Denote thus by π_λ the Poisson measure on \mathbb{N} with parameter $\lambda > 0$ and let us assume that, for some constant $C > 0$, and all f , say bounded, on \mathbb{N} ,

$$\int f^2 \log f^2 d\pi_\lambda - \int f^2 d\pi_\lambda \log \int f^2 d\pi_\lambda \leq C \int |Df|^2 d\pi_\lambda \quad (4)$$

where here $Df(x) = f(x + 1) - f(x)$, $x \in \mathbb{N}$. Apply then (4) to the indicator function of the interval $[k + 1, \infty)$, $k \in \mathbb{N}$. Hence

$$-\pi_\lambda([k + 1, \infty)) \log \pi_\lambda([k + 1, \infty)) \leq C\pi_\lambda(\{k\})$$

which is clearly impossible as k goes to infinity. A further discussion, with various choices of gradients and Dirichlet forms associated to Poisson measure, is presented in the recent work [G-R]. Similarly, (4) still cannot hold with addition of an extra $C' \int f^2 d\pi_\lambda$ on the right-hand side.

One may therefore be led to consider some variation of inequality (2) that could behave better under the preceding limits. To this aim, let us first introduce some notation. If ν is a probability measure on some measurable space (E, \mathcal{E}) , and if f is an integrable function on E with respect to ν , we denote by $E_\nu(f)$ the mean of f with respect to ν . We also write $\text{Var}_\nu(f)$ for its variance. If f is a non-negative function on E , we denote by

$$\text{Ent}_\nu(f) = \int f \log f d\nu - \int f d\nu \log \int f d\nu$$

the entropy, with respect to ν , of f (possibly infinite). We will make repeated use of the product property of entropy (and similarly of variance) in the sense that if ν^n is the product measure of ν on the product space E^n , for all non-negative f on E^n ,

$$\text{Ent}_{\nu^n}(f) \leq \sum_{i=1}^n \int \text{Ent}_\nu(f_i) d\nu^n \tag{5}$$

where, for every $i = 1, \dots, n$, f_i denotes the function on E which is the restriction of f to i th coordinate, the other coordinates being fixed.

An equivalent formulation of the Gaussian inequality (3) is that, for any smooth f on the line with strictly positive values,

$$\text{Ent}_\gamma(f) \leq \frac{1}{2} E_\gamma \left(\frac{1}{f} f'^2 \right). \tag{6}$$

That (6) is equivalent to (3) simply follows from a change of functions together with the chain rule formula for the usual gradient on \mathbb{R} . Of course, such a change may not be performed equivalently on discrete gradients, so that there is some interest to study an inequality such as

$$\text{Ent}_{\mu_p}(f) \leq C E_{\mu_p} \left(\frac{1}{f} |Df|^2 \right)$$

on $\{0, 1\}$ for the Bernoulli measure μ_p and to ask for the best constant C as a function of p . Our first result will be to show that, for any f with strictly positive values on $\{0, 1\}$,

$$\text{Ent}_{\mu_p}(f) \leq pq E_{\mu_p} \left(\frac{1}{f} |Df|^2 \right) \quad (7)$$

and that the constant 1 in front of pq is optimal. Now, the dependence on p in (7) is much better than in (2) and, as a result, this inequality may be tensorized to yield via the Poisson limit theorem

$$\text{Ent}_{\pi_\lambda}(f) \leq \lambda E_{\pi_\lambda} \left(\frac{1}{f} |Df|^2 \right). \quad (8)$$

This inequality may be considered as a possible form of a logarithmic Sobolev inequality for Poisson measure (since (4) does not hold). We will see in the next sections how natural (8) is, and describe further some alternate forms. Inequality (7) also yields the Gaussian logarithmic Sobolev inequality (6) up to a constant 2 via the Gaussian central limit theorem. Actually, we will give, in the process of the proof of (7), some sharper inequalities which imply (6) with its best constant (as well as the optimal Poincaré inequalities for μ_p^n and π_λ). These results are presented in Section 2.

These variations on logarithmic Sobolev inequalities are well-adapted to the family of measures we consider. Indeed, it is by now well-known that from the Gaussian logarithmic Sobolev inequality (3) one may deduce in a simple way that the distribution, with respect to γ , of a Lipschitz function g on \mathbb{R} has a Gaussian tail. The argument goes back to I. Herbst and has been revived recently by several authors [D-S, A-M-S, A-S, L1, R] (see [G-R] for the historical developments and references). It consists in applying (3) to $e^{\tau g/2}$ to deduce a differential inequality in $\tau \in \mathbb{R}$ on the Laplace transform of g that implies a Gaussian tail. Variations on this theme are described in the recent paper [G-R] that also deals with discrete gradients. We will first observe here from (8) that the distribution with respect to π_λ of a Lipschitz function f on \mathbb{N} has a Poisson tail. To this goal, we will turn to the subject of “modified” logarithmic Sobolev inequalities in Section 4. Modified logarithmic Sobolev inequalities were introduced in [B-L] as a tool to adapt Herbst’s argument to a variety of distributions. Formally, a modified logarithmic Sobolev inequality is of the type

$$\text{Ent}_\nu(e^g) \leq C(\tau) E_\nu(|Dg|^2 e^g)$$

where $C(\tau)$ depends on the bound τ on the uniform norm of the gradient of g . The behavior of this constant as a function of τ determines the integrability properties of Lipschitz functions. For example, the (equivalent) change of functions $f = e^{g/2}$ in (3) (or $f = e^g$ in (6)) yields

$$\text{Ent}_\gamma(e^g) \leq \frac{1}{2} E_\gamma(g'^2 e^g), \tag{9}$$

so that $C(\tau)$ may be chosen here to be bounded. As a consequence, Lipschitz functions have Gaussian tails with respect to γ . The dependence is of the order $(1 - \tau)^{-1}$, i.e., bounded for the small values of τ , in case of the exponential measure [B–L]. For the Poisson measure, the logarithmic Sobolev inequality (8) shows that $C(\tau)$ is exponential in order to reflect the integrability properties of Lipschitz functions for Poisson measure. In particular, it follows that $E_{\pi_\lambda}(e^{c|g|\log_+|g|}) < \infty$ for all $c > 0$ sufficiently small. Actually, we will prove in Section 3 (independently of (8) but with a similar proof) that, for any g on $\{0, 1\}^n$,

$$\text{Ent}_{\mu_p^n}(e^g) \leq pq E_{\mu_p^n}((|Dg| e^{|Dg|} - e^{|Dg|} + 1) e^g), \tag{10}$$

and in the limit, for any g on \mathbb{N} ,

$$\text{Ent}_{\pi_\lambda}(e^g) \leq \lambda E_{\pi_\lambda}((|Dg| e^{|Dg|} - e^{|Dg|} + 1) e^g). \tag{11}$$

In particular, we get a sharp form of modified logarithmic Sobolev inequality for Poisson measure as

$$\text{Ent}_{\pi_\lambda}(e^g) \leq \lambda \frac{\tau e^\tau - e^\tau + 1}{\tau^2} E_{\pi_\lambda}(|Dg|^2 e^g) \tag{12}$$

provided $\sup_{x \in \mathbb{N}} |Dg(x)| \leq \tau$. These alternate forms of logarithmic Sobolev inequalities for Bernoulli and Poisson measures also imply the Gaussian logarithmic Sobolev inequality (9) with its best constant, as well as the optimal spectral gaps for μ_p^n and π_λ .

One important feature of modified logarithmic Sobolev inequalities is that they may be tensorized to produce some sharp tail estimates of functions with respect to the product measures μ_p^n and π_λ^n in terms of two parameters on the gradients. Namely, if g , on $\{0, 1\}^n$ for example, is such that, for every $x = (x_1, \dots, x_n) \in \{0, 1\}^n$,

$$\sum_{i=1}^n |g(x + e_i) - g(x)|^2 \leq \alpha \quad \text{and} \quad \max_{1 \leq i \leq n} |g(x + e_i) - g(x)| \leq \beta$$

where (e_1, \dots, e_n) is the canonical basis of \mathbb{R}^n , then, for every $t \geq 0$,

$$\begin{aligned} & \mu_p^n(\{g \geq E_{\mu_p^n}(g) + t\}) \\ & \leq \exp\left(-\left(\frac{t}{\beta} + \frac{pq\alpha^2}{\beta^2}\right) \log\left(1 + \frac{\beta t}{pq\alpha^2}\right) + \frac{t}{\beta}\right). \end{aligned} \quad (13)$$

A similar inequality thus holds for π_λ^n changing pq into λ . Such an inequality may be considered as an extension of the classical exponential inequalities for sums of independent random variables with parameters the size and the variance of the variables, and describing a Gaussian tail for the small values of t and a Poisson tail for its large values. It belongs to the family of concentration inequalities for product measures deeply investigated by M. Talagrand [T]. With respect to [T], the study presented here develops some new aspects related to concentration for Bernoulli measures and penalties, [T, Section 2]. The results we obtain in the last section describe concentration inequalities under logarithmic Sobolev inequalities in the line of investigation of the previous works [L2] and [B-L].

2. LOGARITHMIC SOBOLEV INEQUALITIES FOR BERNOULLI AND POISSON MEASURES

We first establish inequality (7) for the Bernoulli measure, and, in the process of the proof, a number of sharper forms of possible independent interest. Recall μ_p is the measure on $\{0, 1\}$ with $\mu_p(\{1\}) = p$ and $\mu_p(\{0\}) = q = 1 - p$ where $p \in [0, 1]$. For any $n \geq 1$, we denote by μ_p^n the product measure of μ_p on $\{0, 1\}^n$. If f is a function on $\{0, 1\}^n$, and $x = (x_1, \dots, x_n) \in \{0, 1\}^n$, set

$$|Df|^2(x) = \sum_{i=1}^n |f(x + e_i) - f(x)|^2$$

where (e_1, \dots, e_n) is the canonical basis of \mathbb{R}^n and the addition is modulo 2. Our main result here is the following theorem. In this statement, p is arbitrary in $[0, 1]$, and $q = 1 - p$.

THEOREM 1. *For any positive function f on $\{0, 1\}^n$,*

$$\text{Ent}_{\mu_p^n}(f) \leq pq E_{\mu_p^n} \left(\frac{1}{f} |Df|^2 \right).$$

Proof. We first establish a general calculus lemma.

LEMMA 2. Consider a function

$$U(p) = \text{Ent}_{\mu_p}(f) - pq E_{\mu_p}(g), \quad 0 \leq p \leq 1,$$

where f and g are arbitrary non-negative functions on $\{0, 1\}$. Then $U(p) \leq 0$ for every p if and only if

$$U'(0) \leq 0 \leq U'(1). \tag{14}$$

If, additionally, $f(0) \geq f(1)$ and $g(0) \geq g(1)$ (respectively, $f(0) \leq f(1)$ and $g(0) \leq g(1)$), then the condition (14) may be weakened into $U'(0) \leq 0$ (respectively, $U'(1) \geq 0$).

Proof. Set $a = f(1)$, $b = f(0)$, $\alpha = g(1)$, $\beta = g(0)$, so that

$$U(p) = (pa \log a + qb \log b) - (pa + qb) \log(pa + qb) - pq(p\alpha + q\beta).$$

Since $U(0) = U(1) = 0$, the condition (14) is necessary for U to be non-positive. Now, assume (14) is fulfilled. Differentiating in p , we have

$$U'(p) = (a \log a - b \log b) - (a - b)(\log(pa + qb) + 1)$$

$$+ (p - q)(p\alpha + q\beta) - pq(\alpha - \beta),$$

$$U''(p) = -(a - b)^2 (pa + qb)^{-1} + 2(p\alpha + q\beta) + 2(p - q)(\alpha - \beta),$$

$$U'''(p) = (a - b)^3 (pa + qb)^{-2} + 6(\alpha - \beta),$$

$$U''''(p) = -2(a - b)^4 (pa + qb)^{-3}.$$

Since $U'''' \leq 0$, U'' is concave. Hence, formally three situations are possible.

(1) $U'' \geq 0$ on $[0, 1]$. In this case, U is convex, and thus $U \leq 0$ on $[0, 1]$ in view of $U(0) = U(1) = 0$.

(2) $U'' \leq 0$ on $[0, 1]$. By (14), this case is only possible if U is identically 0.

(3) For some $0 \leq p_0 < p_1 \leq 1$, $U'' \leq 0$ on $[0, p_0]$, $U'' \geq 0$ on $[p_0, p_1]$, and $U'' \leq 0$ on $[p_1, 1]$. In this case, U is concave on $[0, p_0]$, and, due to the assumption $U'(0) \leq 0$, one may conclude that U is non-increasing on $[0, p_0]$. In particular, $U \leq 0$ on $[0, p_0]$. It is then necessary that $U(p_1) \leq 0$. Indeed, U is concave on $[p_1, 1]$, hence the assumption $U(p_1) > 0$ together with $U(1) = 0$ would imply $U'(1) < 0$ which contradicts (14). As a result, by convexity of U on $[p_0, p_1]$, we get $U \leq 0$ on $[p_0, p_1]$. At last, $U \leq 0$ on $[p_1, 1]$, since U is concave on $[p_1, 1]$, $U(p_1) \leq 0$ and $U'(1) \geq 0$ (in particular, U is non-decreasing on this interval). The first part of Lemma 2 is thus proved.

We turn to the second part. Again, since $U(0) = U(1) = 0$, any of the conditions $U'(0) \leq 0$ or $U'(1) \geq 0$ is necessary for U to be non-positive on $[0, 1]$. Now, assume that $a \geq b$, $\alpha \geq \beta$, and $U'(0) \leq 0$ (the other case is similar). Then $U''' \geq 0$, and hence U'' is non-decreasing on $[0, 1]$. Again three cases are formally possible.

(1) $U'' \geq 0$ on $[0, 1]$. In this case, U is convex, and thus $U \leq 0$ on $[0, 1]$ in view of $U(0) = U(1) = 0$.

(2) $U'' \leq 0$ on $[0, 1]$. This case is not possible unless $U \equiv 0$.

(3) For some $0 \leq p_0 \leq 1$, $U'' \leq 0$ on $[0, p_0]$ and $U'' \geq 0$ on $[p_0, 1]$. In this case, U is concave on $[0, p_0]$, and, due to the fact that $U'(0) \leq 0$, one may conclude that U is non-increasing on $[0, p_0]$. In particular $U \leq 0$ on $[0, p_0]$. At last, $U \leq 0$ on $[p_0, 1]$ since U is convex on this interval and $U(p_0) \leq 0$ and $U(1) = 0$. Lemma 2 is established. ■

Before turning to the applications of this lemma, let us note the following. In the notation of the proof of Lemma 2, set

$$R(a, b) = a \log \frac{a}{b} - (a - b).$$

Clearly, $R(a, b) \geq 0$ for all $a, b \geq 0$. Then,

$$U'(0) \leq 0 \quad \text{if and only if} \quad \beta \geq R(a, b) \quad (15)$$

while

$$U'(1) \geq 0 \quad \text{if and only if} \quad \alpha \geq R(b, a). \quad (16)$$

The following consequence of Lemma 2 implies Theorem 1 for $n = 1$. The full statement of Theorem 1 then follows from the product property (5) of entropy.

COROLLARY 3. *For any positive function f on $\{0, 1\}$,*

$$\text{Ent}_{\mu_p}(f) \leq pq \left(1 - \frac{1}{2M(f)} \right) E_{\mu_p} \left(\frac{1}{f} |Df|^2 \right)$$

where

$$M(f) = \max \left\{ \frac{f(1)}{f(0)}, \frac{f(0)}{f(1)} \right\}.$$

Proof. By Lemma 2 with $g = C/f$, and according to (15) and (16), the optimal value of $C > 0$ in the inequality

$$\text{Ent}_{\mu_p}(f) \leq Cpq E_{\mu_p} \left(\frac{1}{f} \right)$$

provided $p \in [0, 1]$ is arbitrary is given by

$$C = \max\{bR(a, b), aR(b, a)\},$$

where $a = f(1)$, $b = f(0)$. By symmetry, one may assume that $a > b > 0$. Then, $bR(a, b) \leq aR(b, a)$. Indeed, for fixed $b > 0$, the function $\rho(a) = aR(b, a) - bR(a, b)$ has derivative $\rho'(a) = 2R(b, a) \geq 0$. Hence $\rho(a) \geq \rho(b) = 0$. Thus, $C = aR(b, a)$, $a > b > 0$. Now, fixing $b > 0$, consider

$$u(a) = aR(b, a) = a \left(b \log \frac{b}{a} - (b - a) \right), \quad a > b.$$

We have $u'(a) = b \log(b/a) - 2(b - a)$, thus $u(b) = u'(b) = 0$ and, for every $a > 0$,

$$u''(a) = 2 - \frac{b}{a} \leq 2 - \frac{1}{M(f)}.$$

Hence, by a Taylor expansion, denoting by a_0 some middle point between a and b , we get

$$\begin{aligned} C = u(a) &= u(b) + u'(b)(a - b) + \frac{1}{2} u''(a_0)(a - b)^2 \\ &\leq \left(1 - \frac{1}{2M(f)} \right) (a - b)^2. \end{aligned}$$

Since $(a - b)^2 = |f(1) - f(0)|^2$, Corollary 3, and thus Theorem 1, are established. ■

In the same way, one can show on the basis of Lemma 2 that for any f on $\{0, 1\}$ with positive values,

$$\text{Ent}_{\mu_p}(f^2) \leq 2pq(1 + \log M(f)) E_{\mu_p}(f^2). \tag{17}$$

This inequality reflects in another way the constant in (2).

As announced, the logarithmic Sobolev inequality of Theorem 1 may be used in the limit to yield a logarithmic Sobolev inequality for Poisson

measure. Take namely φ on \mathbb{N} such that $0 < c \leq \varphi \leq C < \infty$ and apply Theorem 1 to

$$f(x) = f(x_1, \dots, x_n) = \varphi(x_1 + \dots + x_n),$$

$$x = (x_1, \dots, x_n) \in \{0, 1\}^n,$$

with this time $p = \lambda/n$, $\lambda > 0$ (for every n large enough). Then, setting $S_n = x_1 + \dots + x_n$,

$$|Df|^2(x) = (n - S_n)[\varphi(S_n + 1) - \varphi(S_n)]^2 + S_n[\varphi(S_n) - \varphi(S_n - 1)]^2.$$

Therefore,

$$\text{Ent}_{\mu_p^n}(\varphi(S_n)) \leq \frac{\lambda}{n} \left(1 - \frac{\lambda}{n}\right) \text{E}_{\mu_p^n} \left(\frac{1}{\varphi(S_n)} \left((n - S_n)[\varphi(S_n + 1) - \varphi(S_n)]^2 + S_n[\varphi(S_n) - \varphi(S_n - 1)]^2 \right) \right).$$

The distribution of S_n under $\mu_{\lambda/n}^n$ converges to π_λ . Using that $0 < c \leq \varphi \leq C < \infty$ and that $(1/n) \text{E}_{\mu_p^n}(S_n) \rightarrow 0$, we immediately obtain the following corollary.

COROLLARY 4. *For any f on \mathbb{N} with strictly positive values,*

$$\text{Ent}_{\pi_\lambda}(f) \leq \lambda \text{E}_{\pi_\lambda} \left(\frac{1}{f} |Df|^2 \right)$$

where we recall that here $Df(x) = f(x + 1) - f(x)$, $x \in \mathbb{N}$.

Again, no better proportion of λ can hold in this inequality as can be checked with the functions $f(x) = e^{cx}$, $x \in \mathbb{N}$, as $c \rightarrow \infty$.

Theorem 1 may also be used to imply, as in [G], the Gaussian logarithmic Sobolev inequality (6) up to a constant 2. Actually, using the refined forms of Corollary 3 or of (17), one can recover (6) with its optimal value. Indeed, tensorizing Corollary 3 first shows that for any f with strictly positive values on $\{0, 1\}^n$,

$$\text{Ent}_{\mu_p^n}(f) \leq pq \left(1 - \frac{1}{2M(f)}\right) \text{E}_{\mu_p^n} \left(\frac{1}{f} |Df|^2 \right) \quad (18)$$

where

$$M(f) = \max_{x \in \{0, 1\}^n} \max_{1 \leq i \leq n} \frac{f(x + e_i)}{f(x)}.$$

Let then $\varphi > 0$ be smooth enough on \mathbb{R} , for example C^2 with bounded derivatives, and apply (18) to

$$f(x_1, \dots, x_n) = \varphi \left(\frac{x_1 + \dots + x_n - np}{\sqrt{npq}} \right)$$

for fixed $p, 0 < p < 1$. Under the smoothness properties on φ , it is easily seen that $M(f) \rightarrow 1$ as $n \rightarrow \infty$. Therefore, by the Gaussian central limit theorem, we deduce in the classical way inequality (6) for φ . Changing φ into φ^2 , and using a standard approximation procedure, we get Gross's logarithmic Sobolev inequality (3) with its best constant. The same reasoning can be performed on (17). Other consequences of these sharp forms are the spectral gap inequalities for μ_p^n and π_λ . Applying (18) (or the product form of (17)) to $1 + \varepsilon f$ and letting ε go to 0, we get, since $M(1 + \varepsilon f) \rightarrow 1$,

$$\text{Var}_{\mu_p^n}(f) \leq pq E_{\mu_p^n}(|Df|^2) \tag{19}$$

and

$$\text{Var}_{\pi_\lambda}(f) \leq \lambda E_{\pi_\lambda}(|Df|^2). \tag{20}$$

3. MODIFIED LOGARITHMIC SOBOLEV INEQUALITIES FOR BERNOULLI AND POISSON MEASURES

In this section, we investigate some related logarithmic Sobolev inequalities for Bernoulli and Poisson measures that will lead to some sharp form of modified logarithmic Sobolev inequalities. As in the previous section, we start with the Bernoulli measure. The following statement will be our basic result.

THEOREM 5. *For any function g on $\{0, 1\}^n$,*

$$\text{Ent}_{\mu_p^n}(e^g) \leq pq E_{\mu_p^n}((|Dg| e^{|Dg|} - e^{|Dg|} + 1) e^g).$$

Proof. It is similar to the proof of Theorem 1 and relies on the next lemma.

LEMMA 6. *The optimal constant C in the inequality*

$$\text{Ent}_{\mu_p}(e^g) \leq Cpq E_{\mu_p}(e^g)$$

provided p is arbitrary in $[0, 1]$ and $g: \{0, 1\} \rightarrow \mathbb{R}$ is fixed is given by

$$C = h e^h - e^h + 1$$

where $h = |g(1) - g(0)|$.

Proof. One may assume that $g(1) = h \geq 0 = g(0)$. By Lemma 2 (second part), the optimal constant $C \geq 0$ in the inequality

$$U(p) = \text{Ent}_{\mu_p}(e^g) - Cpq E_{\mu_p}(e^g) \leq 0$$

can be found from $U'(0) \leq 0$. According to (15) with $a = e^h$, $b = 1$, $\alpha = Ce^h$, $\beta = C$, this condition is just

$$C \geq R(a, b) = he^h - e^h + 1$$

which is the result. Lemma 6 is proved. ■

According to Lemma 6, the theorem is proved in dimension one. We now simply observe that the inequality may be tensorized. By the product property of entropy (5), we get namely, for every g on $\{0, 1\}^n$,

$$\text{Ent}_{\mu_p^n}(e^g) \leq pq \int \sum_{i=1}^n (|g(x + e_i) - g(x)| e^{|g(x + e_i) - g(x)|} - e^{|g(x + e_i) - g(x)|} + 1) e^{g(x)} d\mu_p^n(x) \quad (21)$$

where we recall that (e_1, \dots, e_n) is the canonical basis of \mathbb{R}^n and that $x + e_i$ is understood here modulo 2. The function $Q(v) = \sqrt{v} e^{\sqrt{v}} - e^{\sqrt{v}} + 1$, $v \geq 0$, is increasing and convex on $[0, \infty)$ with $Q(0) = 0$. Hence, setting $a_i = |g(x + e_i) - g(x)|$, $i = 1, \dots, n$,

$$\sum_{i=1}^n Q(a_i^2) \leq Q\left(\sum_{i=1}^n a_i^2\right) = Q(|Dg(x)|^2) = |Dg(x)| e^{|Dg(x)|} - e^{|Dg(x)|} + 1.$$

Theorem 5 is therefore established. ■

As for Corollary 4, the Poisson limit theorem on (21) yields the following consequence for π_λ .

COROLLARY 7. For any function g on \mathbb{N} ,

$$\text{Ent}_{\pi_\lambda}(e^g) \leq \lambda E_{\pi_\lambda}((|Dg| e^{|Dg|} - e^{|Dg|} + 1) e^g).$$

Corollary 7 is sharp in many respect. It becomes an equality for linear functions of the type $g(x) = ax + b$, $a \geq 0$. Furthermore, applying Theorem 5 and Corollary 7 to εg with $\varepsilon \rightarrow 0$ yields the Poincaré inequalities

(19) and (20) for μ_p^n and π_λ respectively since $he^h - e^h + 1 \sim \frac{1}{2}h^2$ as $h \rightarrow 0$. Similarly, (21) contains the Gaussian logarithmic Sobolev inequality in the form of inequality (9) by applying (21), in a classical way, to

$$g(x_1, \dots, x_n) = \varphi \left(\frac{x_1 + \dots + x_n - np}{\sqrt{npq}} \right)$$

for φ smooth enough on \mathbb{R} . The same argument may be developed on the product form of Corollary 7 together with the central limit theorem for sums of independent Poisson random variables. Theorem 5 and Corollary 7 admit furthermore a number of variations, with similar proofs. One of them is an interpolation inequality between variance and entropy. Let us only briefly present it on the discrete cube. Namely, when $1 \leq s \leq 2$, for any g on $\{0, 1\}^n$,

$$E_{\mu_p^n}(e^{sg}) - (E_{\mu_p^n}(e^g))^s \leq pq E_{\mu_p^n}((e^s |Dg| - s e^{|Dg|} + s - 1) e^{sg}). \tag{22}$$

Letting $s \rightarrow 1$, we recover Theorem 5. By the central limit theorem, for every g smooth enough on \mathbb{R} ,

$$E_\gamma(e^{sg}) - (E_\gamma(e^g))^s \leq \frac{1}{2}s(s-1) E_\gamma(g'^2 e^{sg}),$$

that is, with the change of function $f = e^g$, the family of interpolation inequalities for Gaussian measure between variance ($s=2$) and entropy ($s=1$) put forward in [Be]. By the Poisson limit theorem, there is an analogous inequality for π_λ .

As announced, the preceding statements actually describe sharp forms of modified logarithmic Sobolev inequalities in this context. As a consequence of Theorem 5 and Corollary 7, we namely get

COROLLARY 8. *For any function g on $\{0, 1\}^n$ with $\max_{1 \leq i \leq n} |g(x + e_i) - g(x)| \leq \tau$ for every x in $\{0, 1\}^n$,*

$$\text{Ent}_{\mu_p^n}(e^g) \leq pq \frac{\tau e^\tau - e^\tau + 1}{\tau^2} E_{\mu_p^n}(|Dg|^2 e^g).$$

The case $n=1$ is just Lemma 6 together with the fact that $\tau^{-2}[\tau e^\tau - e^\tau + 1]$ is non-decreasing in $\tau \geq 0$. The corollary follows by tensorization. Similarly,

COROLLARY 9. *For any function g on \mathbb{N} with $\sup_{x \in \mathbb{N}} |Dg(x)| \leq \tau$,*

$$\text{Ent}_{\pi_\lambda}(e^g) \leq \lambda \frac{\tau e^\tau - e^\tau + 1}{\tau^2} E_{\pi_\lambda}(|Dg|^2 e^g).$$

4. MODIFIED LOGARITHMIC SOBOLEV INEQUALITIES AND POISSON TAILS

Logarithmic Sobolev inequalities of the type of those described in Sections 2 and 3 entail some information on the Poisson behavior of Lipschitz functions. The following proposition describes a first result in this direction. For simplicity, we only deal with the case of measures on \mathbb{N} . It applies in particular to π_λ .

Let ν be a probability measure on \mathbb{N} such that, for some constant $C > 0$,

$$\text{Ent}_\nu(f) \leq C E_\nu \left(\frac{1}{f} |Df|^2 \right) \quad (23)$$

for all functions f on \mathbb{N} with positive values, where $Df(x) = f(x+1) - f(x)$, $x \in \mathbb{N}$. Let now g on \mathbb{N} such that $\sup_{x \in \mathbb{N}} |Dg(x)| \leq \tau$, and apply (23) to $f = e^g$. Since

$$|Df(x)| = |e^{g(x+1)} - e^{g(x)}| = |Dg(x)| e^\theta$$

for some $\theta \in]g(x), g(x+1)[$ or $]g(x+1), g(x)[$, and since $|Dg(x)| = |g(x+1) - g(x)| \leq \tau$, we get that

$$\text{Ent}_\nu(e^g) \leq C e^{2\tau} E_\nu(|Dg|^2 e^g). \quad (24)$$

In particular,

$$\text{Ent}_\nu(e^g) \leq C \tau^2 e^{2\tau} E_\nu(e^g). \quad (25)$$

On the basis of (25), we obtain a first result on Poisson tails of Lipschitz functions.

PROPOSITION 10. *Let ν be a probability measure on \mathbb{N} such that, for some constant $C > 0$,*

$$\text{Ent}_\nu(f) \leq C E_\nu \left(\frac{1}{f} |Df|^2 \right)$$

for all functions f on \mathbb{N} with positive values, where $Df(x) = f(x+1) - f(x)$, $x \in \mathbb{N}$. Then, for any g such that $\sup_{x \in \mathbb{N}} |Dg(x)| \leq 1$, we have $E_\nu(|g|) < \infty$ and, for all $t \geq 0$,

$$\nu(\{g \geq E_\nu(g) + t\}) \leq \exp \left(-\frac{t}{4} \log \left(1 + \frac{t}{2C} \right) \right).$$

In particular, $E_\nu(e^{c|g| \log_+ |g|}) < \infty$ for sufficiently small $c > 0$.

The inequality of Proposition 10 describes the classical Gaussian tail behavior for the small values of t and the Poisson behavior for the large values of t (with respect to C). The constants have no reason to be sharp.

Proof. Assume first that g is bounded. We may assume that $E_\nu(g) = 0$. Apply (25) to τg to get that, for every $\tau \geq 0$,

$$\text{Ent}_\nu(e^{\tau g}) \leq C\tau^2 e^{2\tau} E_\nu(e^{\tau g}). \tag{26}$$

Provided with this inequality, the conclusion is easy. We need simply follow the corresponding argument for Gaussian logarithmic Sobolev inequalities (cf. e.g. [L1]). Let $G(\tau) = E_\nu(e^{\tau g})$, $\tau \geq 0$, be the Laplace transform of g . Then, by the definition of entropy, (26) reads as

$$\tau G'(\tau) - G(\tau) \log G(\tau) \leq C\tau^2 e^{2\tau} G(\tau), \quad \tau \geq 0.$$

Setting $H(\tau) = (1/\tau) \log G(\tau)$, $H(0) = E_\nu(g)$, we see that, for every $\tau \geq 0$, $H'(\tau) \leq C e^{2\tau}$. Hence, $H(\tau) \leq H(0) + (C/2)(e^{2\tau} - 1)$, that is

$$E_\nu(e^{\tau g}) \leq e^{(C\tau/2)(e^{2\tau} - 1)}, \quad \tau \geq 0.$$

By Chebyshev's inequality, for every $t \geq 0$ and $\tau \geq 0$,

$$\nu(\{g \geq E_\nu(g) + t\}) \leq e^{-\tau t + (C\tau/2)(e^{2\tau} - 1)}.$$

When $t \leq 2C$ (for example), choose $\tau = t/4C$ so that

$$e^{-\tau t + (C\tau/2)(e^{2\tau} - 1)} \leq e^{-\tau t + 2C\tau^2} = e^{-t^2/8C},$$

while, when $t \geq 2C$, choose $\tau = \frac{1}{2} \log(t/C)$ for which

$$e^{-\tau t + (C\tau/2)(e^{2\tau} - 1)} \leq e^{-(t/4) \log(t/C)}.$$

These two cases together yield the tail estimate of the proposition, at least when g is bounded. Let now g be arbitrary such that $\sup_{x \in \mathbb{N}} |Dg(x)| \leq 1$. Let $g^N = \max(-N, \min(g, N))$, $N \geq 0$. Then g^N is bounded for every N and $\sup_{x \in \mathbb{N}} |Dg^N(x)| \leq 1$, and thus satisfies the conclusion of the proposition. The same is true for $|g^N|$ and $-|g^N|$. Applying first this result to $-|g^N|$, for every $t \geq 0$,

$$\nu(\{|g^N| \leq E_\nu(|g^N|) - t\}) \leq \exp\left(-\frac{t}{4} \log\left(1 + \frac{t}{2C}\right)\right).$$

Choose t_0 large enough independent of N such that the right-hand-side of this inequality is less than $\frac{1}{2}$. Choose furthermore m so that $\nu(\{|g| \geq m\}) < \frac{1}{2}$. Since $|g^N| \leq |g|$, we get, by comparing the probabilities, that $E_\nu(|g^N|) \leq m + t_0$ independently of N . Moreover, for every $t \geq 0$,

$$\nu(\{|g^N| \geq m + t_0 + t\}) \leq \exp\left(-\frac{t}{4} \log\left(1 + \frac{t}{2C}\right)\right)$$

from which it is easily seen that $\sup_N E_\nu(|g^N|^2) < \infty$. We then conclude by uniform integrability that g is integrable and satisfies the inequality of the statement. The proof of Proposition 10 is complete. ■

The inequality (24) that precedes (25) (on which Proposition 10 is based) is part of the family of so-called modified logarithmic Sobolev inequalities described in the introduction. The advantage of (24) over (25) is that it may be tensorized to yield concentration properties for the product measures ν^n in terms of two parameters, one on the “Euclidean” norm of the gradient, and one on the sup-norm. These concentration properties are part of the general study of [T]. The results we obtain describe in particular some further aspects of concentration for Bernoulli measures. The following proposition is the general result on tensorization of modified logarithmic Sobolev inequalities and application to concentration. We establish next the sharp form of (13) for Bernoulli and Poisson measures.

PROPOSITION 11. *Let ν be some measure on \mathbb{N} . Assume that for every g on \mathbb{N} with $\sup_{x \in \mathbb{N}} |Dg(x)| \leq \tau$,*

$$\text{Ent}_\nu(e^g) \leq C(\tau) E_\nu(|Dg|^2 e^g) \tag{27}$$

where, as function of $\tau \geq 0$,

$$C(\tau) \leq c_1 e^{c_2 \tau}$$

for some $c_1, c_2 > 0$. Denote by ν^n the product measure on \mathbb{N}^n . Let g be a function on \mathbb{N}^n such that, for every $x \in \mathbb{N}^n$,

$$\sum_{i=1}^n |g(x + e_i) - g(x)|^2 \leq \alpha^2 \quad \text{and} \quad \max_{1 \leq i \leq n} |g(x + e_i) - g(x)| \leq \beta.$$

Then $E_{\nu^n}(|g|) < \infty$ and, for every $t \geq 0$,

$$\nu^n(\{g \geq E_{\nu^n}(g) + t\}) \leq \exp\left(-\frac{t}{2c_2\beta} \log\left(1 + \frac{\beta c_2 t}{4c_1 \alpha^2}\right)\right).$$

Proof. We may first tensorize (27) to get that for every g on \mathbb{N}^n such that $\max_{1 \leq i \leq n} |g(x + e_i) - g(x)| \leq \tau$ for every $x \in \mathbb{N}^n$,

$$\text{Ent}_{\nu^n}(e^g) \leq C(\tau) E_{\nu^n} \left(\sum_{i=1}^n |D_i g|^2 e^g \right) \tag{28}$$

where $D_i g(x) = g(x + e_i) - g(x)$, $i = 1, \dots, n$. Fix then g on \mathbb{N}^n satisfying the hypotheses of the statement. We may assume, by homogeneity, that $\beta = 1$. From now on, the proof is similar to the proof of Proposition 10. In particular, we may assume throughout the argument that g is bounded. Apply (28) to τg for every $\tau \in \mathbb{R}$. Setting $G(\tau) = E_{\nu^n}(e^{\tau g})$, we get

$$\tau G'(\tau) - G(\tau) \log G(\tau) \leq \alpha^2 \tau^2 C(\tau) G(\tau). \tag{29}$$

Therefore, if $H(\tau) = (1/\tau) \log G(\tau)$, $H(0) = E_{\nu^n}(g)$,

$$H'(\tau) \leq \alpha^2 C(\tau) \leq \alpha^2 c_1 e^{c_2 \tau}.$$

It follows that, for every $\tau \geq 0$,

$$H(\tau) \leq H(0) + \alpha^2 \frac{c_1}{c_2} (e^{c_2 \tau} - 1).$$

In other words,

$$E_{\nu^n}(e^{\tau g}) \leq e^{\tau E_{\nu^n}(g) + c_1 \alpha^2 \tau (e^{c_2 \tau} - 1)/c_2} \tag{30}$$

which holds for every $\tau \in \mathbb{R}$ (changing g into $-g$). We conclude with Chebyshev's exponential inequality. For every τ ,

$$\nu^n(\{g \geq E_{\nu^n}(g) + t\}) \leq e^{-\tau t + c_1 \alpha^2 \tau (e^{c_2 \tau} - 1)/c_2}.$$

If $c_2 t \leq 4c_1 \alpha^2$ (for example), choose $\tau = t/4c_1 \alpha^2$ whereas when $c_2 t \geq 4c_1 \alpha^2$, take

$$\tau = \frac{1}{c_2} \log \left(\frac{c_2 t}{2c_1 \alpha^2} \right).$$

The proof is easily completed. ■

Corollary 9 describes the sharp modified logarithmic Sobolev inequality for Poisson measure. Due to the sharp behavior of $C(\tau)$, the tail estimates of Proposition 10 and 11 may be improved accordingly. If we start for example in Proposition 11 with μ_p^n or π_λ , and with Corollaries 8 or 9 instead of (27), one may improve as announced the bound (30) on the

Laplace transform and the corresponding tail estimate. We namely get instead of (29),

$$\tau G'(\tau) - G(\tau) \log G(\tau) \leq \lambda \alpha^2 (\tau e^\tau - e^\tau + 1) G(\tau).$$

Therefore, for every τ ,

$$H'(\tau) \leq \lambda \alpha^2 \frac{\tau e^\tau - e^\tau + 1}{\tau^2}.$$

Since

$$\int_0^\tau \frac{u e^u - e^u + 1}{u^2} du = \frac{e^\tau - 1 - \tau}{\tau},$$

it follows that, for every $\tau \in \mathbb{R}$,

$$E_{\mu_p^n}(e^{\tau g}) \leq e^{\tau E_{\mu_p^n}(g) + \lambda \alpha^2 (e^\tau - 1 - \tau)}.$$

The same holds for π_λ^n and this bound is sharp since, when $n = 1$ for example, it becomes an equality for $g(x) = x$, $x \in \mathbb{N}$. Together with Chebyshev's inequality and a straightforward minimization procedure, it implies the tail estimate (13) for either μ_p^n or π_λ^n (with pq or λ).

ACKNOWLEDGMENT

We are most grateful to I. Pinelis for pointing out an error in a first version of this work.

REFERENCES

- [A-M-S] S. Aida, T. Masuda, and I. Shigekawa, Logarithmic Sobolev inequalities and exponential integrability, *J. Funct. Anal.* **126** (1994), 83–101.
- [A-S] S. Aida and D. Stroock, Moment estimates derived from Poincaré and logarithmic Sobolev inequalities, *Math. Res. Lett.* **1** (1994), 75–86.
- [Be] W. Beckner, A generalized Poincaré inequality for Gaussian measures, *Proc. Amer. Math. Soc.* **105** (1989), 397–400.
- [B-L] S. Bobkov and M. Ledoux, Poincaré's inequalities and Talagrand's concentration phenomenon for the exponential measure, *Probab. Theory related Fields* **107** (1997), 383–400.
- [Bo] A. Bonami, Étude des coefficients de Fourier des fonctions de $L^p(G)$, *Ann. Inst. Fourier* **20** (1970), 335–402.
- [D-S] E. B. Davies and B. Simon, Ultracontractivity and the heat kernel for Schrödinger operators and Dirichlet Laplacians, *J. Funct. Anal.* **59** (1984), 335–395.
- [D-SC] P. Diaconis and L. Saloff-Coste, Logarithmic Sobolev inequalities for finite Markov chains, *Ann. Appl. Prob.* **6** (1996), 695–750.

- [G] L. Gross, Logarithmic Sobolev inequalities, *Amer. J. Math.* **97** (1975), 1060–1083.
- [G–R] L. Gross and O. Rothaus, Herbst inequalities for supercontractive semigroups, preprint, 1997.
- [H–Y] Y. Higuchi and N. Yoshida, Analytic conditions and phase transition for Ising models, *Lecture Notes in Japanese* (1995).
- [L1] M. Ledoux, Remarks on logarithmic Sobolev constants, exponential integrability and bounds on the diameter, *J. Math. Kyoto Univ.* **35** (1995), 211–220.
- [L2] M. Ledoux, On Talagrand’s deviation inequalities for product measures, *ESAIM Prob. & Stat.* **1** (1996), 63–87.
- [R] O. Rothaus, Logarithmic Sobolev inequalities and the growth of L^p norms (1996), *Proc. Amer. Math. Soc.*, to appear.
- [SC] L. Saloff-Coste, “Lectures on Finite Markov Chains, École d’Été de Probabilités de St-Flour 1996,” *Lecture Notes in Math.*, Vol. 1685, pp. 301–413, Springer-Verlag, Berlin/New York, 1997.
- [T] M. Talagrand, Concentration of measure and isoperimetric inequalities in product spaces, *Publ. Math. I.H.E.S.* **81** (1995), 73–205.