Concentration and Relative Entropy for Compound Poisson Distributions

M. Madiman
Division of Applied Mathematics
Brown University
Providence, RI 02912, USA
mokshay@dam.brown.edu

I. Kontoyiannis
Division of Applied Mathematics
& Department of Computer Science
Brown University
Providence, RI 02912, USA
yiannis@dam.brown.edu

***Abstract*— Using a simple inequality about the relative entropy, its so-called "tensorization property," we give a simple proof of a functional inequality which is satisfied by any compound Poisson distribution. This functional inequality belongs to the class of modified logarithmic Sobolev inequalities. We use it to obtain measure concentration bounds for compound Poisson distributions under a variety of assumptions on their tail behavior. In particular, we show how the celebrated "Herbst argument" can be modified to yield *sub-exponential* concentration bounds. For example, suppose Z is a compound Poisson random variable with values on the nonnegative integers, and let f be a function such that $|f(k+1)-f(k)| \leq 1$ for all k. Then, if the base distribution of Z does not have a finite moment-generating function but has finite moments up to some order $L > 1$, we show that the probability that $f(Z)$ exceeds its mean by a positive amount t or more decays approximately like (const)$\cdot t^{-L}$, where the constant is explicitly identified. This appears to be one of the very first examples of concentration bounds with power-law decay.**

I. Introduction

Concentration of measure is a well-studied phenomenon, and in the past 30 years or so it has been explored through a wide array of tools and techniques; see, e.g., [20][15][16][18] for broad introductions to the subject. The concentration phenomenon is equally strongly motivated by theoretical questions (in areas such as geometry, functional analysis and probability), as by its numerous applications in different fields including the analysis of algorithms, mathematical physics and empirical processes in statistics.

A typical measure concentration result is an explicit bound for the probability of an event of the form $\{X - m > t\}$, for some random variable X to exceed its mean m by a certain amount t. Two classical, well-known examples are the following:

(i) If $X = \sum_{i=1}^n X_i$ is the sum of n independent random variables X_i with values in an interval $[a,b]$ of length $\Delta = b-a$, then

$$\Pr\big\{X - E(X) > t\big\} \leq e^{-2t^2/n\Delta^2}.$$

This is a special case of Hoeffding's inequality [16], and it gives a simple exponential bound on the probability that the sum X will exceed its mean by an amount t. Note that the bound depends on the distribution of the summands X_i only through the size of their support.

(ii) If $X \sim N(0,1)$ is a standard normal random variable and f is an arbitrary 1-Lipschitz function, i.e., $|f(x)-f(y)| \leq |x-y|$ for all $x, y \in \mathbb{R}$, then [5][1],

$$\Pr\big\{f(X) - E[f(X)] > t\big\} \leq e^{-t^2/2}. \tag{1}$$

Again we see that, although the distribution of the random variable of interest $f(X)$ may be quite complex we have a simple, explicit bound on the probability that it deviates from its mean by an amount t. This is a general theme in many measure concentration results: Under certain conditions, it is possible to derive useful, accurate bounds for the deviations away from the mean for a large class of random variables with complex and possibly not completely known distributions.

Our purpose in this work is to use methods related to information-theoretic ideas in order to prove concentration bounds similar to (1) when X is a *compound Poisson random variable.*

Over the past 10 years, several authors have drawn interesting connections between different aspects of measure concentration and information theory; see, e.g., [17][6][12][7]. In particular, one of the main strategies for establishing measure concentration bounds is the "entropy method" pioneered by Ledoux; cf. [13][14][15]. An important step in applying this method often involves the use of the "tensorization property of the entropy," which, in information-theoretic terms, can be stated as follows.

Let A be a discrete alphabet. Consider the distribution P of an arbitrary n-tuple of random variables $(X_1, X_2, \ldots, X_n)$ with values in A, and similarly let Q be the distribution of an arbitrary collection of *independent* A-valued random variables $(Y_1, Y_2, \ldots, Y_n)$.

Tensorization Property. [14, Prop. 4.1][15, Prop. 5.6][16, Sect. 5] With P and Q as above we have,

$$D(P\|Q) \leq \sum_{j=1}^n [D(P\|Q) - D(P^{(j)}\|Q^{(j)})], \tag{2}$$

where $P^{(j)}$ (resp. $Q^{(j)}$) denotes the $(n-1)$-dimensional distribution of $(X_1, X_2, \ldots, X_{j-1}, X_{j+1}, \ldots, X_n)$ (resp. of $(Y_1, Y_2, \ldots, Y_{j-1}, Y_{j+1}, \ldots, Y_n)$).

The proof of (2) is straightforward and only involves an application of Jensen's inequality and simple algebraic manipulations.

Historically, the entropy method was first developed by Herbst (in an unpublished observation) who used a celebrated logarithmic Sobolev inequality due to Gross [8] in order to prove the Gaussian concentration bound in (1). More recently, Bobkov and Ledoux [2] adopted a similar strategy to prove corresponding bounds for the case of the Poisson distribution. They derived a "modified" logarithmic Sobolev inequality, and adapted Herbst's argument to prove concentration bounds analogous to (1) for Poisson-distributed random variables X.

In this work we continue along the same line of investigation and obtain concentration bounds for the much richer family of nonnegative integer-valued *compound Poisson* distributions. Given $\lambda > 0$ and a probability distribution $Q = (Q_j)$ on the positive integers $j \in \mathbb{N} = \{1, 2, 3, \ldots\}$, the random variable Z has a *compound Poisson distribution with parameters* λ *and* Q, denoted CP(λ, Q), if its law can be written as that of the sum of a Poisson(λ) number of independent random variables with distribution Q, i.e.,

$$Z \stackrel{\mathcal{D}}{=} \sum_{i=1}^{W} X_i,$$

where $W \sim$ Poisson(λ) and the X_i are independent and identically distributed according to Q [by convention we take the empty sum from $i = 1$ to $i = 0$ to be equal to zero]. An alternative representation for the law of Z which we will find useful below, is in terms of the series

$$Z \stackrel{\mathcal{D}}{=} \sum_{j=1}^{\infty} j\, W_j, \qquad W_j \sim \text{Poisson}(\lambda Q_j), \tag{3}$$

where the W_j are independent random variables, each W_j having the Poisson(λQ_j) distribution.

The class of compound Poisson distributions is of course much richer that the simple Poisson law, and they are, in a certain sense, dense among all infinitely divisible laws [19]. In particular, the compound Poisson law CP(λ, Q) inherits its tail behavior from that of the so-called "base distribution" Q. For example if Q has a finite moment-generating function, then so does CP(λ, Q) for any λ; similarly, if Q has finite moments up to some order L, then so does CP(λ, Q); and so on. It is in part from this versatility of tail behavior that the CP distribution draws its importance in many applications.

Our main result is a concentration inequality for the CP(λ, Q) measure when Q only has finite moments up to a certain order $L > 1$, Theorem 2 below. This is one of the first results on *sub-exponential* concentration bounds. To our knowledge, the only other works on such bounds are [10], [4] and [3]. In [10], the idea of covariance representations was combined with truncation and explicit computations to prove concentration bounds for the class of stable laws on $\mathbb{R}^d$. The recent preprint [4] extends these results to a large class of functionals on Poisson space, including infinitely divisible random variables satisfying certain conditions. Both the methods of [10][4] as well as the form of the results themselves are very different from those derived here; however, more detailed comparisons are made in [11]. Another recent paper [3] contains moment inequalities for functions of independent random variables, primarily with statistical applications in mind. This relates to our work in that [3] also extends the Herbst argument to certain situations where exponential moments do not exist, but their extension as well as the results they obtain are very different from those presented here.

Our methods are general but also elementary. Following along the same path as Bobkov and Ledoux, in Section II we first establish a modified logarithmic Sobolev inequality for an arbitrary CP law. Although this can be immediately deduced from the recent (and more general) results of [21], we give a much simpler proof using the tensorization property of the entropy, thus avoiding Wu's sophisticated methods involving stochastic calculus.

Section III contains our main contribution, which is a new modification of the Herbst argument to obtain concentration inequalities with power-law decay. When the base distribution Q of a random variable Z with CP(λ, Q) distribution only has finite moments up to order $L > 1$, we give explicit bounds for the probability of event $\{f(Z) - E[f(Z)] > t\}$ that are approximately of order t^{-L}; f here is an arbitrary Lipschitz function on the nonnegative integers. We also show how our methods can be readily adapted to recover some recent results on the concentration of CP(λ, Q) laws when Q has finite support or finite exponential moments.

II. A Modified Logarithmic Sobolev Inequality

To state the main result of this section we find it convenient to introduce the following minor generalization of the relative entropy. Given a probability distribution $P = (P_i)$ on the nonnegative integers $\mathbb{Z}_+ = \{0, 1, 2, \ldots\}$ and a nonnegative function $f : \mathbb{Z}_+ \to \mathbb{R}$, define the *entropy functional* $\text{Ent}_P(f)$ as

$$\text{Ent}_P(f) = \sum_i P_i f_i \log f_i \; - \; [\sum_i P_i f_i] \log[\sum_i P_i f_i].$$

Note that if f is the discrete density between some probability distribution Q and P, i.e., $f_i = Q_i/P_i$, then $\text{Ent}_P(f)$ reduces to the familiar relative entropy $D(Q\|P)$.

The following result is an immediate consequence of [21, Cor. 4.2]. As discussed above, we provide an alternative proof.

Theorem 1. (Modified Logarithmic Sobolev Ineq.) If Z has distribution CP(λ, Q), then for any bounded function $f : \mathbb{Z}_+ \to \mathbb{R}$ we have,

$$\begin{aligned}&\text{Ent}_{\text{CP}(\lambda,Q)}(e^f)\\ &\quad \le \lambda \sum_{j\ge 1} Q_j E\Big[e^{f(Z)}\Big\{e^{|D^j f(Z)|}(|D^j f(Z)| - 1) + 1\Big\}\Big],\end{aligned}$$

where $D^j f(x) = f(x + j) - f(x)$.

PROOF OUTLINE. Our starting point is the following modified logarithmic Sobolev inequality due to Bobkov and Ledoux [2]: If W has Poisson(λ) distribution, denoted Po(λ), then for any bounded function $f : \mathbb{Z}_+ \to \mathbb{R}$ we have,

$$\begin{aligned} &\text{Ent}_{\text{Po}(\lambda)}(e^f) &(4)\\ &\quad \le \lambda E\Big[e^{f(W)}\Big\{e^{|Df(W)|}(|Df(W)|-1)+1\Big\}\Big], &(5)\end{aligned}$$

where D is the difference operator, $Df(x) = f(x+1) - f(x)$.

In order to prove the corresponding statement for the CP(λ, Q) distribution we first rewrite the tensorization property in terms of the entropy functional. Let $(Y_1, Y_2, \ldots, Y_n)$ be independent random variables with each $Y_j \sim$ Poisson(λQ_j), and let $\mathbb{Q}_n$ denote their joint distribution. For an arbitrary function $G : \mathbb{Z}_+^n \to \mathbb{R}$, the tensorization property (2) can be expressed as

$$\text{Ent}_{\mathbb{Q}_n}(G) \le \sum_{j=1}^n E\Big\{\text{Ent}_{\text{Po}(\lambda Q_j)}\big(G_j(Y_1^{j-1}, \cdot, Y_{j+1}^n)\big)\Big\}, \quad (6)$$

where the entropy functional on the right hand side is applied to the restriction G_j of G to its jth co-ordinate.

Now given an f as in the statement of the theorem, define the functions $G : \mathbb{Z}_+^n \to \mathbb{R}$ and $H : \mathbb{Z}_+^n \to \mathbb{R}$ by

$$H(y_1^n) = f\Big(\sum_{k=1}^n k y_k\Big), \quad y_1^n \in \mathbb{Z}_+^n,$$

and $G = e^H$. Let $\overline{Q}_n$ denote the distribution of $\sum_{k=1}^n kY_k$ and write $H_j : \mathbb{Z}_+ \to \mathbb{R}$ for the restriction of H to the variable y_j with the remaining y_i's held fixed. Applying (6) to this G we obtain

$$\begin{aligned}\text{Ent}_{\overline{Q}_n}(e^f) &= \text{Ent}_{\mathbb{Q}_n}(e^H)\\ &= \text{Ent}_{\mathbb{Q}_n}(G)\\ &\le \sum_{j=1}^n E\Big\{\text{Ent}_{\text{Po}(\lambda Q_j)}\Big(G_j(Y_1^{j-1}, \cdot, Y_{j+1}^n)\Big)\Big\}\\ &= \sum_{j=1}^n E\Big\{\text{Ent}_{\text{Po}(\lambda Q_j)}\big(e^{H_j(Y_1^{j-1}, \cdot, Y_{j+1}^n)}\big)\Big\},\end{aligned}$$

and using the Bobkov-Ledoux logarithmic Sobolev inequality (5) to bound each term in the above right hand side,

$$\text{Ent}_{\overline{Q}_n}(e^f) \le \lambda \sum_{j=1}^n Q_j E\Big\{e^{H(Y_1^n)}\eta\big(DH_j(Y_1^n)\big)\Big\},$$

where

$$\eta(x) = |x|e^{|x|} - e^{|x|} + 1, \quad x \in \mathbb{R}. \quad (7)$$

Observing that, trivially,

$$DH_j(y_1^n) = D^j f\Big(\sum_{k=1}^n k y_k\Big),$$

and writing S_n for the sum $S_n = \sum_{k=1}^n kY_k$, we obtain that

$$\text{Ent}_{\overline{Q}_N}(e^f) \le \lambda \sum_{j=1}^n Q_j E\Big\{e^{f(S_n)}\eta\big(D^j f(S_n)\big)\Big\}. \quad (8)$$

Finally, we recall from the Introduction that the compound Poisson law CP(λ, Q) of the random variable Z can be expressed as the limit (in the sense of convergence of distribution) as $n \to \infty$ of the laws of the sums S_n; equivalently, $\overline{Q}_n \to$ CP(λ, Q) in distribution, as $n \to \infty$. Therefore, taking the limit in (8) yields,

$$\text{Ent}_{\text{CP}(\lambda,Q)}(e^f) \le \lambda \sum_{j=1}^\infty Q_j E\Big\{e^{f(Z)}\eta\big(D^j f(Z)\big)\Big\}.$$

as claimed. □

III. MAIN RESULTS: CONCENTRATION FOR COMPOUND POISSON DISTRIBUTIONS

Here we state our main results, namely three concentration of measure inequalities for Lipschitz functions of CP random variables. Their complete proofs will be given in an extended version of this paper [11].

Our main result gives a sharp concentration bound for the tails of a CP(λ, Q) random variable, when Q only has finite moments up to some order L.

Theorem 2. (POWER-LAW CONCENTRATION) Suppose Z has distribution CP(λ, Q), where the base distribution has finite moments up to order L,

$$L = \sup\{t \ge 1 : \sum_{j \ge 1} Q_j j^t < \infty\} \in (1, \infty).$$

Let $f : \mathbb{Z}_+ \to \mathbb{R}$ be K-Lipschitz,

$$|Df(x)| = |f(x+1) - f(x)| \le K \quad \text{for all } x.$$

Then, for any $t > 0$ and any $\epsilon > 0$, the probability $\Pr\{|f(Z) - E[f(Z)]| > t\}$ is bounded above by

$$\exp\Big\{\min_{0<\alpha<L}\Big[I_\epsilon(\alpha) - \alpha\log\Big(\frac{t}{2|f(0)| + 2\lambda K q_1 + \epsilon}\Big)\Big]\Big\}$$

where

$$\begin{aligned} q_1 &= \sum_{j=1}^\infty jQ_j\\ I_\epsilon(\alpha) &= \lambda\sum_{j=1}^\infty Q_j\big[C_{j,\epsilon}^\alpha - 1 - \alpha\log C_{j,\epsilon}\big]\\ \text{and} \quad C_{j,\epsilon} &= 1 + \frac{jK}{\epsilon}.\end{aligned}$$

Remarks.

1) This upper bound is meaningful (less than 1) when $t > 2|f(0)| + 2\lambda K q_1 + \epsilon$, and in this case, one has the alternate representation

$$\exp\Big\{-\int_0^a i_\epsilon^{-1}(s)ds\Big\}$$

where $i_\epsilon(\alpha) = I'_\epsilon(\alpha)$ and $a = \log(t/[2|f(0)| + 2\lambda K q_1 + \epsilon])$.

2) Although the constants in the bound given in the theorem are expressed in an implicit form, it is straightforward to get exact bounds by taking specific values for the parameters α and ϵ. In particular, it is easy to see that this bound decays approximately like t^{-L} for t large. A simpler form of the bound in Theorem 2 indicating this is given in the corollary below. Moreover, a different, sometimes better result than the corollary can be obtained in the case when $L > 2$.
3) From the theorem we immediately obtain useful integrability properties of Lipschitz functions of a CP random variable; e.g., for any K-Lipschitz function f, $E[|f(Z)|^\tau] < \infty$ for all $\tau < L$.
4) In particular, using $f(x) = x$ we can recover the following well-known result: The tail of CP(λ, Q) decays like the tail of Q, whenever the tail of Q has a power-law decay for some power $L > 1$; cf. [19].
5) Since in the proof we do not use explicitly the fact that Z has a CP distribution, except to say that we can apply the logarithmic Sobolev inequality of Theorem 1, it follows that analogous concentration bounds hold for *any* random variable Z with values in $\mathbb{Z}_+$ whose law satisfies the same logarithmic Sobolev inequality.

PROOF OUTLINE. The main idea is to extend the Herbst argument from its usual (exponential) form to the present setting of power-law tails. Therefore, unlike the typical case where the goal is to first obtain bounds on the moment generating function of $f(Z)$, we instead obtain bounds for the function

$$G(\tau) = E[|f(Z)|^\tau], \quad \tau \in (1, L).$$

Roughly speaking, this is achieved by substituting the function

$$f_\tau(x) = \tau \log |f(x)|$$

in the logarithmic Sobolev inequality of Theorem 1, to get a bound of the form

$$\mathrm{Ent}_{\mathrm{CP}(\lambda,Q)}\big(|f|^\tau\big) \le \lambda G(\tau) \sum_{j=1}^{\infty} \eta(\tau \log C_{j,\epsilon}),$$

where the function $\eta(x)$ is defined as in (7). We also note that

$$\mathrm{Ent}_{\mathrm{CP}(\lambda,Q)}\big(|f|^\tau\big) = \tau G'(\tau) - G(\tau) \log G(\tau),$$

so that expanding the last two expressions and rearranging terms appropriately yields a differential inequality for $G(\tau)$. Solving this inequality gives us an explicit upper bound for $G(\tau)$ which, in conjunction with Markov's inequality, proves the result. □

Corollary. Under the assumptions of Theorem 2, suppose $n \ge 1$ is integer with $n < L$. Then, for any $t > 0$,

$$\Pr\{|f(Z) - E[f(Z)]| > t\} \le A\Big(\frac{B}{t}\Big)^n$$

where

$$A = \exp\Big\{\lambda \sum_{r=1}^{n} \binom{n}{r} K^r q_r - \lambda n \log K\Big\}$$
$$B = 1 + 2|f(0)| + 2\lambda K q_1,$$

and q_k denotes the kth moment of the base distribution Q,

$$q_k = \sum_{j\ge1} j^k Q_j.$$

Next we show how elementary techniques – the standard Herbst argument in conjunction with the logarithmic Sobolev inequality of Theorem 1 – can be used to give a simple proof for a recent result of [9], without requiring the more sophisticated covariance representations used there. Theorem 3 gives concentration bounds for Lipschitz functions of CP random variables under the assumption of exponential moments.

Theorem 3. (EXPONENTIAL CONCENTRATION) Suppose Z has distribution CP(λ, Q), where the base distribution has finite exponential moments up to order M,

$$M = \sup\{t \ge 0 : \sum_{j\ge1} Q_j e^{tj} < \infty\} \in (0, \infty).$$

Let $f : \mathbb{Z}_+ \to \mathbb{R}$ be K-Lipschitz,

$$|Df(x)| = |f(x+1) - f(x)| \le K \quad \text{for all } x.$$

Then, for any $t > 0$,

$$\begin{aligned}\Pr\{f(Z) - E[f(Z)] \ge t\} &\le \exp\Big(\min_{0<\alpha<M}\big[H(\alpha) - \alpha t\big]\Big)\\ &= \exp\Big(-\int_0^t h^{-1}(s)ds\Big)\end{aligned}$$

where

$$H(\alpha) = \lambda \sum_{j=1}^{\infty} Q_j\big[e^{\alpha K j} - 1 - \alpha K j\big],$$

and h^{-1} is the inverse of the function

$$h(\alpha) = H'(\alpha) = \lambda K \sum_{j=1}^{\infty} Q_j\big[j\big(e^{\alpha K j} - 1\big)\big].$$

Remarks.

1) Simple calculations show that for t small this bound behaves like the Gaussian tail. However, for t large, we get exponential decay of order approximately $e^{-M/K}$.
2) As for Theorem 2, integrability properties of Lipschitz functions are a simple consequence of the proof: For any K-Lipschitz function f,

$$E[e^{\tau f(Z)}] < \infty \quad \text{for all} \quad \tau < \frac{M}{K}.$$

Finally we show, extending a different result from [9], that better results can be obtained under the assumption that the base distribution Q corresponding to some CP(λ, Q) law has bounded support. In that case, Lipschitz functions have Poisson-like tails of the form $e^{-ct\log t}$.

Theorem 4. (POISSON CONCENTRATION) Suppose Z has distribution CP(λ, Q), where the base distribution has finite support,

$$m = \max\{j \in \mathbb{Z}_+ : Q_j > 0\} < \infty.$$

Let $f : \mathbb{Z}_+ \to \mathbb{R}$ be K-Lipschitz,

$$|Df(x)| = |f(x+1) - f(x)| \leq K \quad \text{for all } x.$$

Then, for any $t > 0$,

$$\Pr\{f(Z) - E[f(Z)] \geq t\} \leq \exp\{\rho(t)\},$$

where $\rho(t)$ is defined as the minimum between the two expressions

$$-\frac{t}{2m}\log\left(1 + \frac{t}{2m\lambda}\right)$$
$$\text{and} \qquad \frac{t}{m} - \left(\frac{t}{m} + \frac{q_2^2\lambda}{m^2}\right)\log\left(1 + \frac{mt}{q_2^2\lambda}\right),$$

with q_2 denoting the second moment of the base distribution Q as before.

Furthermore, for $t \in [0, \frac{3\lambda m^2}{2}]$,

$$\Pr\{f(Z) - E[f(Z)] \geq t\}) \leq \exp\left\{-\frac{t^2}{3\lambda m^3}\right\}.$$

Remark. In fact, given *any* finite-range interval (that is, restricting t to $[0, T]$ for some $T < \infty$), the method of proof can be easily modified to yield a Gaussian bound for deviations in that range. As we would expect from our result that the "far" tails of Lipschitz functions decay in a Poisson-like manner and not in a Gaussian fashion, the variance of these bounds increases without bound as the size of the finite range increases.

ACKNOWLEDGMENTS

We wish to thank Christian Houdré for sending us the preprint [4] and for numerous interesting comments on these and related results. We also wish to thank Andrew Barron and an anonymous referee for bringing [3] to our attention. I.K. was supported in part by a Sloan Foundation Research Fellowship and by NSF grant #0073378-CCR.

REFERENCES

[1] S. Aida, T. Masuda, and I. Shigekawa. Logarithmic Sobolev inequalities and exponential integrability. *J. Funct. Anal.*, 126(1):83–101, 1994.
[2] S.G. Bobkov and M. Ledoux. On modified logarithmic Sobolev inequalities for Bernoulli and Poisson measures. *J. Funct. Anal.*, 156(2):347–365, 1998.
[3] S. Boucheron, O. Bousquet, G. Lugosi, and P. Massart. Moment inequalities for functions of independent random variables. *Annals of Prob.*, 33(2):514–560, 2005.
[4] J-C. Breton, C. Houdré and N. Privault. Dimension free and infinite variance tail estimates on Poisson space. *Preprint*, 2004.
[5] E.B. Davies and B. Simon. Ultracontractivity and the heat kernel for Schrödinger operators and Dirichlet Laplacians. *J. Funct. Anal.*, 59(2):335–395, 1984.
[6] A. Dembo. Information inequalities and concentration of measure. *Ann. Probab.*, 25(2):927–939, 1997.
[7] E. Figueroa-Lopez and C. Houdré. On the asymptotic redundancy of lossless block coding with two codeword lengths. *Preprint*, 2004.
[8] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4):1061–1083, 1975.
[9] C. Houdré. Remarks on deviation inequalities for functions of infinitely divisible random vectors. *Ann. Probab.*, 30(3):1223–1237, 2002.
[10] C. Houdré and P. Marchal. On the concentration of measure phenomenon for stable and related random vectors. *Ann. Probab.*, 32(2):1496–1508, 2004.
[11] I. Kontoyiannis and M. Madiman. Measure concentration for Compound Poisson distributions. *Preprint*, 2005.
[12] I. Kontoyiannis and A.D. Sezer. A remark on unified error exponents: Hypothesis testing, data compression and measure concentration. In *Progress in Probability*, volume 56, pages 23–32. Birkhauser, Basel, Switzerland, 2003.
[13] M. Ledoux. Isoperimetry and Gaussian analysis. In *Lectures on probability theory and statistics (Saint-Flour, 1994)*, volume 1648 of *Lecture Notes in Math.*, pages 165–294. Springer, Berlin, 1996.
[14] M. Ledoux. On Talagrand's deviation inequalities for product measures. *ESAIM Probab. Statist.*, 1:63–87 (electronic), 1997.
[15] M. Ledoux. *The concentration of measure phenomenon*. American Mathematical Society, Providence, RI, 2001.
[16] G. Lugosi. Concentration-of-measure inequalities. *Presented at the Winter School on Probabilistic Methods in High Dimension Phenomena*, Toulouse, January 2005. Notes available at www.econ.upf.es/~lugosi.
[17] K. Marton. Bounding $\bar{d}$-distance by informational divergence: a method to prove measure concentration. *Ann. Probab.*, 24(2):857–866, 1996.
[18] C. McDiarmid. Concentration. In *Probabilistic methods for algorithmic discrete mathematics*, pages 195–248. Algorithms Combin., 16, Springer, Berlin, 1998.
[19] K-I. Sato. *Lévy processes and infinitely divisible distributions*, volume 68. Cambridge University Press, Cambridge, 1999. Translated from the 1990 Japanese original, Revised by the author.
[20] M. Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. *Inst. Hautes Études Sci. Publ. Math.*, No. 81:73–205, 1995.
[21] L. Wu. A new modified logarithmic Sobolev inequality for Poisson point processes and several applications . *Probab. Theory Related Fields*, 118:427–438, 2000.