Invent. math. 126, 505–563 (1996)

Inventiones mathematicae

New concentration inequalities in product spaces

Michel Talagrand

Equipe d Analyse-Tour 56, E.R.A. au C.N.R.S. no. 754, Université Paris VI, 4 Pl Jussieu, F-75230 Paris Cedex 05, France
Department of Mathematics, The Ohio State University, 231 W. 18th Ave., Columbus, OH 43210-1174, USA

Oblatum 29-IX-1995 & 15-IV-1996

Abstract. We introduce three new ways to measure the "distance" from a point to a subset of a product space and we prove corresponding concentration inequalities. Each of them allows to control the fluctuation of a new class of random variables.

1. Introduction

The idea of concentration of measure in product spaces has undergone striking recent development, that has shed a new light on the notion of independence itself. While exponential inequalities for sums of independent random variables are at the core of classical probabilities, the new abstract inequalities are far reaching extensions that apply to considerably more general functions.

In the case of product spaces (that is the only one we will consider) concentration of measure has been very thoroughly studied in [T3], where a number of directions are investigated. The present work is devoted to new directions that have emerged since then. It is largely independent of [T3]. In fact, only the material of Sects. 3.1 and 4.1 of [T3] will be relevant to us, and the content of these few pages will be explained again in the present introduction (inequalities (1.2) and (1.13) below). Nice introductions to the topic and connected areas can be found in [L1], [L2].

The idea of concentration of measure has an already long history. A detailed account (from this author's perspective) is given in [T4], and there is no point to repeat it here. Rather, we will concentrate upon explaining the material of the present paper and its relevance.

Work partially supported by an NSF grant.

Consider a probability space (Ω, μ) and its product Ω^n, $P = \mu^{\otimes n}$. The basic idea is that if a subset A of Ω^n is not too small (say $P(A) \geqq 1/2$), a "generic" point x of Ω^n is "close" to A. The meaning of "generic" is clear: an inequality will quantify the size of the exceptional set. More subtle is the notion of "closeness". In fact the introduction of appropriate notions has been the main ingredient of past progress. Basically we can say that x is close to A if A contains points y such that x and y differ in "few" coordinates (although simply counting these coordinates is not effective). Throughout the paper, for two points x, y of Ω^n, we will set

$$h_i(x, y) = 1 \quad \text{if } x_i \neq y_i$$

$$h_i(x, y) = 0 \quad \text{if } x_i = y_i \,.$$

Incidentally we could mention that the study of other functions and "penalties" for $x_i \neq y_i$ was one of the main themes of [T3], a theme that turns out to be of considerable subtlety. It will not be considered here, and our results are valid without extra structure on (Ω, μ).

The most apparent way to measure the "distance" of x and y is by their Hamming distance

$$d(x, y) = \sum_{i \leqq n} h_i(x, y)$$

and then to set $d(x,A) = \inf\{d(x, y) : y \in A\}$ as measure of the distance of x to A. This is how early results were formulated [M-S]. There is however a considerably more efficient procedure, described in [T3], Sect. 4.1. There, it is explained as a convexification procedure. We will here explain it with a different perspective that is more in the spirit of the present paper. Consider a probability measure v on Ω^n. We set

$$f(v,x) = \sum_{i \leqq n} \left(\int h_i(x, y)\, dv(y)\right)^2 .$$

For a subset A of Ω^n, we set

$$f(A,x) = \inf\{ f(v,x);\ v(A) = 1, v \text{ atomic}\} . \tag{1.1}$$

It is easy to see that the number $f(A,x)$ of (1.1) coincides with the quantity denoted by $f_c^2(A,x)$ in [T3]. Thus, by [T3] Theorem 4.1.1, we have

$$\int \exp \frac{1}{4} f(A,x)\, dP(x) \leqq \frac{1}{P(A)} \tag{1.2}$$

(here, as in [T3], we do not worry about measurability problems, that are irrelevant in the proof of inequalities). In particular, (1.2) implies that

$$P(f(A, \cdot) \geqq u) \leqq \frac{1}{P(A)} \exp\left(-\frac{u}{4}\right) \tag{1.3}$$

which quantifies the idea that $f(A,x)$ is of order 1 for a "generic point" x. We now try to understand the nature of the information brought by $f(A,x)$. Consider numbers $(\alpha_i)_{i \leqq n}, \alpha_i \geqq 0$.

We observe that

$$\int \sum_{i \leqq n} \alpha_i h_i(x,y)\,d\nu(y) = \sum_{i \leqq n} \alpha_i \int h_i(x,y)\,d\nu(y)$$
$$\leqq \left(\sum_{i \leqq n} \alpha_i^2\right)^{1/2} f^{1/2}(\nu,x) \tag{1.4}$$

using Cauchy–Schwarz.

When $\nu(A) = 1$, it follows that

$$\inf\left\{\sum_{i \leqq n} \alpha_i h_i(x,y) : y \in A\right\} \leqq \left(\sum_{i \leqq n} \alpha_i^2\right)^{1/2} f^{1/2}(\nu,x)\,.$$

Thus

$$\inf\left\{\sum_{i \leqq n} \alpha_i h_i(x,y) : y \in A\right\} \leqq \left(\sum_{i \leqq n} \alpha_i^2\right)^{1/2} f^{1/2}(A,x)\,. \tag{1.5}$$

The fact that (1.5) holds for all choices of numbers (α_i), (rather than only for $\alpha_i = 1$) turns out to be of importance for applications as is demonstrated at length in [T3].

The point y of (1.5) has to depend upon the coefficients $(\alpha_i)_{i \leqq n}$. One way to reformulate (1.5) is by saying that a generic point x can be approximated by many points of A. As will be explained at the beginning of Sect. 2 through analysis of a basic example, (1.5) is certainly not the best that can be done. Much of the present paper is motivated by the desire to understand how far one can go in the same direction. The motivation lie not only in the success of (1.5) with applications, but also in the fact that this is one way to study a question that appears to me of central importance, and about which, quite amazingly, little is known; what does a set of probability 1/2 really look like? (The reader is referred to [T5] for some basic open questions in that general direction, questions that are somewhat related to the present work). One essential difficulty there lies into finding proper questions. Such questions arise immediately from the point of view previously introduced. For example, (1.4) means that

$$\left\|\sum_{i \leqq n} \alpha_i h_i(x,\,\cdot)\right\|_{L_1(\nu)} \leqq \left(\sum_{i \leqq n} \alpha_i^2\right)^{1/2} f^{1/2}(\nu,x) \tag{1.6}$$

and it is natural to try to replace the L_1-norm by a stronger norm. In another direction, it would be natural not only to control $\sum_{i \leqq n}(\int h_i(x,y)\,d\nu(y))^2$ but also, say

$$\sum_{i,j \leqq n} \left(\int h_i(x,y)h_j(x,y)\,d\nu(y)\right)^2.$$

In order to formulate the new principle that will accomplish these goals (and more) we observe that

$$f(\nu,x) = \int \sum_{i \leqq n} h_i(x,y_1)h_i(x,y_2)\,d\nu(y_1)\,d\nu(y_2)$$

so that

$$\exp \frac{1}{4} f(v,x) = \exp\left(\frac{1}{4} \int \sum_{i \leqq n} h_i(x, y_1) h_i(x, y_2)\, dv(y_1)\, dv(y_2)\right).$$

It is hence natural to consider what happens when the exponential is put inside the integral rather than outside. We consider the quantity

$$e(v,x) = \int \left(\frac{5}{4}\right)^{\sum_{i \leqq n} h_i(x,y_1) h_i(x,y_2)} dv(y_1)\, dv(y_2) \tag{1.7}$$

and we define $e(A,x)$ as the infimum of $e(v,x)$ when $v(A) = 1$.

Theorem 1.1. *For all sets $A \subset \Omega^n$, we have*

$$\int e(A,x)\, dP(x) \leqq \frac{1}{P(A)}. \tag{1.8}$$

While it is obvious that (modulo the constants) (1.8) improves upon (1.2), it is not apparent at this stage how to use the information contained in the functional $e(A,x)$, or even why this functional is related to the questions addressed above. Doing this requires a non trivial discussion, and is better delayed to Sect. 2, where we hope we will convince the reader that (1.8) represents a very marked improvement upon (1.5) in the understanding, in particular, of some "combinatorial" aspects of sets A with $P(A) \geqq 1/2$. It also turns out that (1.8) has concrete applications, and, as these are easier to understand, we will mention one here. While at first glance, this application could look specialized, as will be explained, the method of proof is extremely general and provides inequalities for random variables under surprisingly weak conditions.

Consider independent Bernoulli random variables ($P(\varepsilon_i = 1) = P(\varepsilon_i = -1) = 1/2$) and vectors $(x_{ij})_{i,j \leqq n}$ is a Banach space W. We assume that $x_{ii} = 0$ for all $i \leqq n$, and that $x_{ij} = x_{ji}$. We denote by W_1^* the unit ball of W, and we consider the numbers

$$U = \sup\left\{ \sum_{i,j \leqq n} \alpha_i \beta_j x^*(x_{ij}); x^* \in W_1^*, \sum_{i \leqq n} \alpha_i^2 \leqq 1, \sum_{j \leqq n} \beta_j^2 \leqq 1 \right\}.$$

Thus U is the supremum of the operator norms of the matrices $(x^*(x_{ij}))_{i,j}$. Consider the number

$$V = E \sup_{x^* \in W_1^*} \left(\sum_{j \leqq n} \left(\sum_{i \leqq n} \varepsilon_i x^*(x_{ij}) \right)^2 \right)^{1/2}.$$

Theorem 1.2. *There exists a number K with the following property. If we consider the random variable*

$$Z = \left\| \sum_{i,j \leqq n} x_{ij} \varepsilon_i \varepsilon_j \right\|$$

and a median M of Z, we have, for each $t > 0$, that

$$P(|Z - M| \geqq t) \leqq 2\exp\left(-\frac{1}{K}\min\left(\frac{t^2}{V^2}, \frac{t}{U}\right)\right). \tag{1.9}$$

Inequality (1.2) should be compared with previous results [L-T], which, based on a decoupling argument, allowed only to control the probabilities $P(Z \geqq 2M + t)$. Inequalities such as (1.2) play an essential role in the theory of stochastic processes. In particular in the case of Gaussian processes, a corresponding inequality is a key to the description of their sample properties [T1]. Thus (1.9) is of potential importance for a better understanding of the processes $\sum_{i,j} \varepsilon_i \varepsilon_j x_{ij}$.

As claimed above, (1.8) can be considered as a significant improvement upon (1.5), in the sense that it expresses in a tighter way that a generic point x can be approximated by many points of A. It does not seem however that (1.8) is the end of the story. In Sect. 3, we attempt to improve upon (1.8). Unfortunately, we could not discover (or even conjecture) what could reasonably be expected (at this stage of our knowledge) to be the "ultimate" result. Our approach is as follows. One of the motivations of replacing the L_1-norm by a stronger norm, say the Orlicz L_{ψ_2}-norm (corresponding to the Orlicz function $e^{x^2} - 1$) in (1.6), i.e. of having an inequality

$$\left\| \sum_{i \leqq n} \alpha_i h_i(x, \cdot) \right\|_{L_{\psi_2}(v)} \leqq \left(\sum_{i \leqq n} \alpha_i^2 \right)^{1/2} U \tag{1.10}$$

for a certain number U is that, given a family $\mathscr{F}$ of coefficients $\alpha = (\alpha_i)_{i \leqq n}$, with $\sum_{i \leqq n} \alpha_i^2 \leqq 1$, we can deduce from (1.10) that, if $v(A) = 1$,

$$\inf_{y \in A} \sup_{\alpha \in \mathscr{F}} \sum_{i \leqq n} \alpha_i h_i(x, y) \leqq KU\sqrt{\log \operatorname{card} \mathscr{F}} \tag{1.11}$$

where K is a universal constant.

In other words, there exists one point y in A such that the sums $\sum_{i \leqq n} \alpha_i h_i(x, y)$ are controlled for all choices of $\alpha \in \mathscr{F}$. In order to improve upon (1.11), one is then led to replace (1.10) by stronger exponential inequalities. It turns out rather unpleasantly that the basic scheme of proof [T3, Sect. 2.1], that has performed so many times (including in the present paper) a flawless job appears here to reach the very limits of its range. Only in the (important) special case $\Omega = \{0, 1\}$, μ uniform, could we obtain a reasonable result, (Theorem 3.1 below) at the expense of significant work. To demonstrate the power of this new principle it is better again to mention a specific consequence. We consider the set T of $n \times n$ matrices with entries zero or one. We provide T with the canonical probability P (that gives mass 2^{-n^2} to each point). We consider on T the distance d induced by the operator norm when a matrix is seen as an operator on ℓ_2^n.

Theorem 1.3. *There exists a number K with the following property. Consider a subset A of T, such that $P(A) \geqq 1/2$. Then*

$$P(\{x; d(x,A) \geqq Kn^{1/4}(\log n)^{5/4}\}) \leqq \frac{1}{n^2}\,. \tag{1.12}$$

The charm of this statement is created by the term $n^{1/4}$, to be compared with the dimension n^2 of T. We will provide an example of set A such that

$$P\left(\left\{x; d(x,A) \geqq \frac{1}{K'}n^{1/4}(\log n)^{1/4}\right\}\right) \geqq \frac{1}{n}$$

where K' is independent of n (see (3.39) below).

While the control of quantities such as $\sum_{i \leqq n} \alpha_i h_i(x,y)$ gives information on how many coordinates y_i are different from x_i, it says nothing of the "location" of y_i when $y_i \neq x_i$. The purpose of Sect. 4 is to provide such information. This is done in Theorem 4.2, which in particular was designed with the following application in mind.

Theorem 1.4. *There exists a number K with the following property. Consider n independent random variables X_i valued in a measurable space Ω. Consider a (countable) class $\mathscr{F}$ of measurable functions on Ω. Consider the random variable $Z = \sup_{f \in \mathscr{F}} \sum_{i \leqq n} f(X_i)$. Consider*

$$U = \sup_{f \in \mathscr{F}} \|f\|_\infty; \qquad V = E \sup_{f \in \mathscr{F}} \sum_{i \leqq n} f(X_i)^2\,.$$

Then for each $t > 0$, we have

$$P(|Z - EZ| \geqq t) \leqq K \exp\left(-\frac{1}{K}\frac{t}{U}\log\left(1 + \frac{tU}{V}\right)\right). \tag{1.13}$$

We will show how to (pessimistically) bound V in function of simpler parameters. To understand (1.13) better, it should be compared with the inequality

$$P(Z \geqq t) \leqq \exp\left(-\frac{1}{K}\frac{t}{U}\log\left(\frac{tU}{V}\right)\right) \tag{1.14}$$

valid for $t \geqq KM$ (M median of Z), a slightly weaker version of which is proved in [T2].

Tail inequalities such as (1.14) are useful in probability, since they are at the root e.g. of laws of large numbers. (In fact, the entire work [T3] has its origin in the search for such inequalities). Inequality (1.14) was applied to Statistics by L. Birgé and P. Massart [B-M1], [B-M2]. It turns out however that what is really needed in Statistics is (1.13). (The questions of Birgé and Massart largely motivated the present paper). At a technical level, we should mention that the proof of (1.14) (and indeed of all comparable inequalities) made essential use of a symmetrization argument, which can give information only on $P(Z \geqq 2M + t)$. That we can prove (1.13) means that we have found a rather new approach to a result at the very center of the theory.

The second most important principle of concentration of measure in product space pertains to what we call q-points inequalities ([T3], Sect. 3.1). For $q \geqq 2$, and $x, y_1, \ldots, y_q$ in Ω^n, we set

$$H_q(x, y_1, \ldots, y_q) = \sum_{i \leqq n} \prod_{\ell \leqq q} h_i(x, y_\ell)$$

and for a set A of Ω^n, we write

$$H_q(A, x) = \inf\{H_q(x, y_1, \ldots, y_q); y_1, \ldots, y_q \in A\} . \tag{1.15}$$

It is proved in [T3], Theorem 3.1.1, that

$$\int q^{H_q(A,x)} dP(x) \leqq \frac{1}{P(A)^q} . \tag{1.16}$$

It is interesting to observe that in the case $q = 2$, a very similar result follows from (1.3). Indeed, we have

$$\begin{aligned} &\int H_2(x, y_1, y_2)\, d\nu(y_1)\, d\nu(y_2) \qquad (1.17)\\ &= \sum_{i \leqq n} \int h_i(x, y_1) h_i(x, y_2)\, d\nu(y_1)\, d\nu(y_2) \\ &= \sum_{i \leqq n} \left(\int h_i(x, y)\, d\nu(y)\right)^2 = f(\nu, x) . \end{aligned}$$

It then follows that $H_2(A, x) \leqq f(A, x)$, so that (1.2) implies $\int \exp(\frac{1}{4} H_2(A, x))\, dP(x) \leqq 1/P(A)$. The advantage of (1.17) is that, rather than just knowing that we can find y_1, y_2 in A that make $H_2(x, y_1, y_2)$ small, we know that it suffices to pick the couple (y_1, y_2) at random according to $\nu \otimes \nu$. This in particular provides an immediate way to obtain interpolation results that were quite laboriously obtained in [T3], Sect. 4.5. To see this, consider numbers $\alpha_i, \beta_i \geqq 0$, and set

$$H(y_1, y_2) = \sum_{i \leqq n} (\alpha_i h_i(x, y_1) + \beta_i h_i(x, y_2)) + H_2(x, y_1, y_2) .$$

Thus

$$\begin{aligned} \inf\{H(y_1, y_2); y_1, y_2 \in A\} &\leqq \int H(y_1, y_2)\, d\nu(y_1)\, d\nu(y_2) \\ &\leqq \left(\left(\sum_{i \leqq n} \alpha_i^2\right)^{1/2} + \left(\sum_{i \leqq n} \beta_i^2\right)^{1/2}\right) f^{1/2}(\nu, x) + f(\nu, x) . \end{aligned}$$

In Sect. 5, we will develop this idea in Theorem 5.1, which is a kind of extension of Theorem 1.1 to the case, where, in the definition of $e(\nu, x)$, the terms $h_i(x, y_1) h_i(x, y_2)$ are replaced by terms $\prod_{\ell \leqq q} h_i(x, y_\ell)$.

In Sect. 5, we will also address the question of finding a result that would play the role of (1.16) in the case $1 < q < 2$ rather than $q \in \mathbb{N},\ q \geqq 2$. This result (that again has natural applications) seems, quite interestingly to be of a somewhat new nature.

Finally, in Sect. 6, we show, in a special case, the possibility of improving upon results such as (1.2) in the case where $P(A)$ is close to one, by replacing the term $1/P(A)$ by a smaller term of optimal order $(1-P(A))/\log(1-P(A))$, which opens one more connection between the area of concentration of measure and questions of a true isoperimetric nature.

The present paper was circulated before publication, and partially motivated further work by Ledoux, Dembo, Dembo and Zeitouni. The papers [D], [D-Z] develop Marton's transportation method; the second paper in particular brings to light the fact that this method does not allow to recover all the results obtained by our method. The important paper of Ledoux [L3] introduces a new method based on the theory of Sobolev inequalities, of remarkable power. While this method does not yet allow to prove abstract statements about concentration of measure such as Theorem 1.1, and yields only deviation inequalities, (that is bounds for $P(Z-EZ \geqq t)$ rather than bounds for $P(|Z-EZ| \geqq t)$), it yields these inequalities (in particular in the setting of Theorem 1.4) with a simplicity and an elegance that pleasantly contrast with the (straightforward but) somewhat laborious derivation from the abstract statements needed here.

2. Higher level

In order to better motivate the need for an improvement upon the functional $f(A,x)$, of (1.2), we will (again) look at the "canonical" situation

$$\Omega = \{0,1\}, \qquad \mu(\{1\}) = p, \qquad A = \left\{ (x_i)_{i \leqq n}, \sum_{i \leqq n} x_i \leqq pn \right\}.$$

The reason why this example is of such considerable importance is that the heuristic content of many of the results of [T3] is that among all possible choices of A, the above one is the worst case. It is thereby not surprising that analysis of this example is a never ending source of inspiration.

For simplicity, we assume $np \in \mathbb{N}$, and we consider $x \in \Omega^n$ such that $\sum_{i \leqq n} x_i = m = q + np$. Quite naturally, the points y of A that are the closest of x in any reasonable sense are those for which $\sum_{i \leqq n} y_i = np$, and $y_i \leqq x_i$ for each i. Consider the set $I = \{i \leqq n; x_i = 1\}$. For each subset J of I, of cardinality q, consider the point y_J such that $y_i = 1$ if $i \in I \backslash J$ and $y_i = 0$ otherwise. Consider subsets $J_1, \ldots, J_r$ of I, of cardinal q, that are disjoint, and such that r is as large as possible, $r = [(m/q)]$. Consider the probability ν_0 on A that gives mass $1/r$ to each of the points $y_{J\ell}$, $\ell \leqq r$. Then an easy calculation shows that $f(\nu_0, x) = q/r$, which is almost the optimum $f(\nu, x)$, where ν gives mass $\binom{m}{q}$ to each y_J. Thus, knowing $f(\nu_0, x)$ does not yield more information than the existence of the points $y_{J\ell}$ in A, and thus certainly does not capture well the fact that very many elements of A approximate x (namely all the y_J for $J \subset I$).

While the functional (1.7) is easy to define, we have found that an equivalent, but somewhat different form is easier to use. We will define first this equivalent form, and only later prove its equivalence with (1.7). While, at the formal level, there is some motivation toward (1.7), the equivalent form we will use is motivated by a somewhat different intuition that could also be of interest. The idea in the definition of the functional $f(v,x)$ could be summarized as follows. To each x, y in Ω^n, we associate an object $\theta_{x,y}$ (namely the sequences $(h_i(x,y))_{i \leqq n}$). For a probability measure v on $\mathbb{R}^n$, we define $\theta_{x,v}$ as the integral $\int \theta_{x,y}\, dv(y)$; we then define $f(v,x)$ as a measure of the size of $\theta_{x,v}$ (here the sum of the squares of the coordinates). Then occurs the thought that if we choose for $\theta_{x,y}$ a more complicated object, it will be harder to make the average $\int \theta_{x,y}\, dv(y)$ small. Thereby knowing that this average is small will provide more information.

We consider the space $\{-1,1\}^n$, provided with the uniform measure, that will simply be denoted by $d\varepsilon$. We define the function r_i by $r_i(\varepsilon) = \varepsilon_i$. We consider a parameter $\beta < 1$. For x, y in Ω^n, we consider the function

$$\theta_{x,y} = \prod_{i \leqq n} (1 + \beta h_i(x,y) r_i)$$

on $\{-1,1\}^n$. Thus $\int \theta_{x,y}\, d\varepsilon = 1$. For a probability measure v on Ω^n, we define

$$\theta_{x,v}(\varepsilon) = \int \theta_{x,y}(\varepsilon)\, dv(y)$$

and we then define

$$e(v,x) = \int (\theta_{x,v})^2\, d\varepsilon\ .$$

Thus (by Cauchy–Schwarz) $e(v,x) \geqq 1$. It will later be shown that when $\beta = 1/2$, this coincides with (1.7). We define $e(A,x)$ as the infimum of $e(v,x)$ over all probability measures v with $v(A) = 1$.

Theorem 2.1. *Assume that* $\beta^2 < 1/2$ *and define* α *by*

$$\alpha = \frac{2\beta^2}{1 - 2\beta^2} \tag{2.1}$$

Then for all subsets A *of* Ω^n *we have*

$$\int_{\Omega^n} e(A,x)\, dP(x) \leqq \frac{1}{P(A)^\alpha}\ . \tag{2.2}$$

We will prove Theorem 2.1, and outline an argument showing that the restriction $\beta^2 < 1/2$ is necessary (a quite interesting fact). Most importantly, we will learn how to use the information contained in the quantity $e(A,x)$. We will show how to relate (2.2) and Theorem 1.1. The reader who wonders how Theorem 2.1 was really discovered can look at [T4], Sect. 5, where the genesis of a previous version is explained.

We first turn to the proof of Theorem 2.1. Central to this proof is the following calculus lemma.

Lemma 2.3. *Consider β with $\beta^2 < 1/2$ and α given by* (2.1). *Define*

$$\gamma = 1 + \beta^2 .$$

Then, for all $0 < r \leqq 1$, we have

$$\inf\{(1-\lambda)^2 r^{-\alpha} + 2\lambda(1-\lambda)r^{-\alpha/2} + \gamma\lambda^2 ; 0 \leqq \lambda \leqq 1\} \leqq 1 + \alpha - \alpha r . \tag{2.3}$$

Proof. If $\gamma r^{\alpha/2} \leqq 1$, the minimum is obtained for $\lambda = 1$ so (2.3) will hold in that case as soon as

$$\gamma r^{\alpha/2} \leqq 1 \Rightarrow \gamma \leqq 1 + \alpha - \alpha r .$$

This holds for all r provided this holds when $\gamma r^{\alpha/2} = 1$. Thus, it suffices to prove (2.3) when $\gamma r^{\alpha/2} \geqq 1$. In that case, the infimum is obtained for

$$\lambda = \frac{r^{-\alpha} - r^{-\alpha/2}}{\gamma + r^{-\alpha} - 2r^{-\alpha/2}}$$

(and this number λ satisfies $0 \leqq \lambda \leqq 1$). After substitution and rearrangement, the conclusion is then equivalent to

$$(1 - r^{\alpha}(\alpha + 1 - \alpha r))(\gamma r^{\alpha} - 2r^{\alpha/2} + 1) \leqq r^{\alpha} - 2r^{\alpha/2} + 1 .$$

But, for $\gamma r^{\alpha/2} \geqq 1$, the function $\gamma r^{\alpha} - 2r^{\alpha/2} + 1$ increases so that it suffices to prove that

$$(\gamma - 1)(1 - r^{\alpha}(\alpha + 1 - \alpha r)) \leqq r^{\alpha} - 2r^{\alpha/2} + 1 .$$

There is equality then for $r = 1$, so that, taking derivatives, it suffices to show that

$$(\gamma - 1)(\alpha + 1)(r^{\alpha} - r^{\alpha-1}) \geqq r^{\alpha-1} - r^{\alpha/2-1}$$

or, equivalently,

$$(\gamma - 1)(\alpha + 1)(1 - r) \leqq r^{-\alpha/2} - 1 .$$

This is true, because our choice of α satisfies

$$(\gamma - 1)(\alpha + 1) = \alpha/2$$

and that

$$\frac{\alpha}{2}(1 - r) \leqq r^{-\alpha/2} - 1$$

since the right-hand side is a convex function. □

We now prove Theorem 2.1 by induction over n. When $n = 1$, the reader will check that $e(A,x) = \gamma$ if $x \notin A$ and $e(A,x) = 1$ if $x \in A$. Thus one has to prove, setting $r = P(A)$ that

$$\gamma(1 - r) + r \leqq r^{-\alpha} .$$

Since $1 + \alpha - \alpha r \leqq r^{-\alpha}$ (by convexity of the function $r^{-\alpha}$) it suffices to see that $\gamma - 1 \leqq \alpha$; but in fact, as used in the proof of Lemma 2.3, we have $\gamma - 1 = \alpha/2(1 + \alpha)$.

We now perform the induction step from n to $n + 1$. We denote by P the product measure on Ω^n. We identify Ω^{n+1} with $\Omega^n \times \Omega$, and the product measure P' on Ω^{n+1} to $P \otimes \mu$.

Consider a subset A of Ω^{n+1}. Consider a point ω in Ω, that will remain fixed for a moment. Consider two atomic probability measures v_0, v_1 on A. We assume that v_1 is supported by the set $\{(x, \omega); x \in \Omega^n\}$. Consider a number $0 \leqq \lambda \leqq 1$, and the measure $v = \lambda v_0 + (1 - \lambda)v_1$; thus $v(A) = 1$. Consider $x \in \Omega^n$, and the point $z = (x, \omega)$ of Ω^{n+1}. We try to estimate $e(v, z)$.

We set for simplicity $g_j = \theta_{z, v_j}$, so that

$$\theta_{z,v} = \lambda g_0 + (1 - \lambda) g_1$$

and

$$e(v,z) = \lambda^2 \int g_0^2 \, d\varepsilon + (1 - \lambda)^2 \int g_1^2 \, d\varepsilon + 2\lambda(1 - \lambda) \int g_0 g_1 \, d\varepsilon \tag{2.4}$$

Consider the operator R that, to each function on $\{-1, 1\}^{n+1}$ associates the function on $\{-1, 1\}^n$ obtained by integrating over the last coordinate. If $\bar{v}_j$ denotes the projection of v_j on Ω^n, it should be clear that $Rg_j = \theta_{x, \bar{v}_j}$.

Since, obviously, g_1 does not depend upon the last coordinate, we have $g_1 = Rg_1$, and for each function h on $\{-1, 1\}^{n+1}$ we have

$$\int h g_1 \, d\varepsilon = \int Rh Rg_1 \, d\varepsilon$$

where we again denote by $d\varepsilon$ the uniform measure on $\{-1, 1\}^n$.

Consider now any u in $\{0, 1\}^n$, and for j in $\{-1, 1\}$, let $a_j = g_0((u, j))$. The reader will be convinced that $a_0 \geqq a_1$, $a_0 \leqq (1 + \beta)/(1 - \beta) a_1$. Therefore,

$$\frac{1}{2}(a_0^2 + a_1^2) \leqq \gamma((a_0 + a_1)/2)^2$$

and, by integration in u,

$$\int g_0^2 \, d\varepsilon \leqq \gamma \int (Rg_0)^2 \, d\varepsilon \, .$$

Combining with (2.4), we get

$$e(v,z) \leqq \gamma\lambda^2 \int (Rg_0)^2 \, d\varepsilon + (1 - \lambda)^2 \int (Rg_1)^2 \, d\theta + 2\lambda(1 - \lambda) \int Rg_0 Rg_1 \, d\varepsilon \, . \tag{2.5}$$

We now observe that $\bar{v}_0$ is an arbitrary atomic probability measure on

$$B = \{x \in \Omega^n; \exists \omega' \in \Omega, (x, \omega') \in A\}$$

while $\bar{v}_1$ is an arbitrary atomic probability measure on

$$A_\omega = \{x \in \Omega^n; (x, \omega) \in A\} \, .$$

Thus, using Cauchy–Schwarz on the term $\int Rg_1 Rg_0 \, d\varepsilon$ and taking the infimum over $\bar{v}_0, \bar{v}_1$, we see from (2.5) that

$$e(A,z) \leqq \gamma\lambda^2 e(B,x) + (1-\lambda)^2 e(A_\omega,x) + 2\lambda(1-\lambda)(e(B,x)e(A_\omega,x))^{1/2} .$$

We now recall that $z = (x,\omega)$. Integration in x, using Cauchy–Schwarz and the induction hypothesis yield

$$\begin{aligned}\int e(A,(x,\omega))\, dP(x) &\leqq \frac{\gamma\lambda^2}{P(B)^\alpha} + \frac{(1-\lambda)^2}{P(A_\omega)^\alpha} + 2\lambda(2-\lambda)\left(\frac{1}{P(B)^\alpha P(A_\omega)^\alpha}\right)^{1/2} \\ &= \frac{1}{P(B)^\alpha}((1-\lambda)^2 r^{-\alpha} + 2\lambda(1-\lambda) r^{-\alpha/2} + \lambda^2\gamma)\end{aligned}$$

where $r = P(A_\omega)/P(B) \leqq 1$. Since λ is arbitrary, we appeal to Lemma 2.2 to get

$$\int e(A,(x,\omega))\, dP(x) \leqq \frac{1}{P(B)^\alpha}\left(1 + \alpha - \alpha\frac{P(A_\omega)}{P(B)}\right).$$

Integrating over ω and using Fubini theorem gives

$$\int e(A,z)\, dP'(z) \leqq \frac{1}{P(B)^\alpha}\left(1 - \alpha - \alpha\frac{P'(A)}{P(B)}\right) \leqq \frac{1}{P'(A)^\alpha}$$

since $P(B) \geqq P'(A)$ and by the inequality $1 - \alpha - \alpha x \leqq x^{-\alpha}$ for $x > 0$. $\square$

To start our discussion of Theorem 2.1, we will show that the condition $\beta^2 < 1/2$ is natural. This will follow from a computation; we unfortunately have no intuitive explanation to offer. Not surprisingly, we will simply analyze what happens in the canonical situation

$$\Omega = \{0,1\}, \qquad \mu(\{1\}) = p, \qquad A = \left\{(x_i)_{i\leqq n}; \sum_{i\leqq r} x_i \leqq np\right\}.$$

Assume for simplicity that $np \in \mathbb{N}$. Consider $q \geqq 0$, and $x \in \Omega^n$ such that $\sum_{i\leqq n} x_i = q + np$. Obviously $e(A,x)$ depends only upon n, p, q.

Given any ε in $\{-1,1\}^n$, we have

$$\theta_{x,y}(\varepsilon) = (1+\beta)^a(1-\beta)^b$$

where $a = \operatorname{card}\{i \leqq n; \varepsilon_i = 1 = h_i(x,y)\}$, $b = \operatorname{card}\{i \leqq n; h_i(x,y) = 1 = -\varepsilon_i\}$. On the other hand, by averaging over all permutations of the coordinates that leave x invariant, we see that the best measure v on A one can choose to approximate x gives equal weight to each of the points y_J for $J \subset I$, $\operatorname{card} J = q$, where, for a subset J of $I = \{i \leqq n; x_i = 1\}$, we set as before $y_{J,i} = 1$ if $i \in I \backslash J$ and $y_{J,i} = 0$ otherwise. Thus

$$\theta_{x,v}(\varepsilon) = \text{average}\,(1+\beta)^{a(\varepsilon,J)}(1-\beta)^{b(\varepsilon,J)} \tag{2.6}$$

where the average is over all choices of J, and where

$$a(\varepsilon, J) = \operatorname{card}\{i \in J; \varepsilon_i = 1\}; \qquad b(\varepsilon, J) = \operatorname{card}\{i \in J; \varepsilon_i = -1\}.$$

It is hard to compute the average (2.6). So, rather, let us consider independent random variables $(\xi_i)_{i\in I}$, $\xi_i \in \{0,1\}$, $P(\xi_i = 1) = q/m$, where $m = np + q = \operatorname{card} I$ and consider the expression

$$h(z) = E\left((1+\beta)^{\sum_{i\in I}\xi_i(1+\varepsilon_i)/2}(1-\beta)^{\sum_{i\in I}\xi_i(1-\varepsilon_i)/2}\right). \tag{2.7}$$

That is, we pretend that $J = \{i \in I; \xi_i = 1\}$. By independence

$$h(z) = (E(1+\beta)^{\xi})^r (E(1-\beta)^{\xi})^{m-r}$$

where $r = \operatorname{card}\{i \in I, \varepsilon_i = 1\}$ and $P(\xi = 1) = q/m$, $P(\xi = 0) = 1 - q/m$.

Now, using algebra

$$H(x) = \int h(\varepsilon)^2\, d\varepsilon = 2^{-m}\sum_r \binom{m}{r}(E(1+\beta)^{\xi})^{2r}(E(1-\beta)^{\xi})^{2m-2r}$$

$$= \left(\frac{1}{2}(E(1+\beta)^{\xi})^2 + \frac{1}{2}(E(1-\beta)^{\xi})^2\right)^m$$

$$= \left(1 + \frac{q^2}{m^2}\beta^2\right)^m .$$

We then let $n \to \infty$, $q \to \infty$, so that if $tq/\sqrt{p(1-p)n}$ has a limit t, we have

$$\lim H(x) = \exp(1-p)\beta^2 t .$$

Thus, by the central limit theorem $\int H(x)\, dP(x)$ can stay bounded independently of n only if $(1-p)\beta^2 < 1/2$. Requiring this for all p means $\beta^2 \leqq 1/2$. The only problem with the previous argument is that it is not quite true that $H(x) = e(A,x)$. It is however possible to make this argument completely rigorous, by showing e.g. that $H(x)$ dominates $e(A',x)$, where $A' = \{(x_i); \sum_{i\leqq n} x_i \leqq np - \sqrt{np}\}$. We leave this to the interested reader.

At a given value of β, Theorem 2.1 quantifies the size of the set where $e(A,x)$ is large, i.e.

$$P(e(A,\cdot) \geqq u) \leqq \frac{1}{uP(A)^{\alpha}} . \tag{2.8}$$

It remains to learn how to use the information contained in $e(A,x)$. We first observe that

$$\int r_i\theta_{x,y}\, d\varepsilon = \beta h_i(x,y) . \tag{2.9}$$

For $k \geqq 1$, let us denote by S_k the set of subsets I of $\{1,\ldots,n\}$ of cardinality k. For $I \in S_k$, we denote by $|I|$ its cardinal. We set

$$r_I = \prod_{i\in I} r_i; h_I(x,y) = \prod_{i\in I} h_i(x,y) .$$

Using (2.9) and independence we get that

$$\int r_I \theta_{x,y}\, d\varepsilon = h_I(x,y)\beta^{|I|} . \tag{2.10}$$

Consider now $S = \bigcup_{k \leqq n} S_k$, and numbers α_I, $I \in S$. Consider the function h on Ω^n given by

$$h(y) = \sum_{I \in S} \alpha_I h_I(x,y)$$

so that, by (2.10),

$$h(y) = \int T(h)\theta_{x,y}\, d\varepsilon \tag{2.11}$$

where

$$T(h) = \sum_{I \in S} \alpha_I \beta^{-|I|} r_I . \tag{2.12}$$

Consider $q \geqq 1$. Using (2.11) and Hölder's inequality we get

$$|h(y)|^q \leqq \int |T(h)|^q \theta_{x,y}\, d\varepsilon .$$

Given a set $B \subset \{-1,1\}^n$ we denote by $|B|$ its measure. Given any $t \geqq 0$, we have, setting $B = \{|T(h)| \geqq t\}$,

$$\begin{aligned} \int |h(y)|^q d\nu(y) &\leqq \int |T(h)|^q \theta_{x,\nu}\, d\varepsilon \\ &\leqq t^q + \int_B |T(h)|^q \theta_{x,\nu}\, d\varepsilon \\ &\leqq t^q + \left(\int_B |T(h)|^{2q}\, d\varepsilon \right)^{1/2} e(\nu,x)^{1/2} \\ &\leqq t^q + |B|^{1/4} \left(\int |T(h)|^{4q}\, d\varepsilon \right)^{1/4} e(\nu,x)^{1/2} \end{aligned} \tag{2.13}$$

where we have used twice Cauchy's inequality. We then get, taking q-th roots, that

$$\|h\|_q \leqq t + \|T(h)\|_{4q} e(\nu,x)^{1/2q} |\{|T(h)| \geqq t\}|^{1/4q} . \tag{2.14}$$

This relates the tail properties of h with those of $T(h)$.

In order to use this inequality, we must estimate $|B|$, that is the tails of $T(h)$. We will first use hypercontractivity to get estimates that are very useful although not always of the optimal order. The following result is due to A. Bonami [B].

Lemma 2.4. *Consider numbers α_I, $I \in S$. Then for $p \geqq 2$ we have*

$$\left\| \sum_{I \in S} \alpha_I r_I \right\|_p \leqq \left(\sum_{I \in S} \alpha_I^2 (p-1)^{|I|} \right)^{1/2} . \tag{2.15}$$

We now specialize (2.14) to the case where $h(y) = \sum_{I \in S_k} \alpha_I h_I(x, y)$, $\sum \alpha_I^2 \leqq 1$, so that, by (2.15)

$$\|T(h)\|_p \leqq \beta^{-k}(p-1)^{k/2} .$$

We have, for any $p \geqq 1$

$$|B| = |\{|T(h)| \geqq t\}| \leqq \frac{\|T(h)\|^p}{t^p} \leqq \left(\frac{\beta^{-k}(p-1)^{k/2}}{t}\right)^p$$

and taking t such that $p = \beta^2 t^{2/k}/e$, we see that if $p \geqq 2$ we have

$$|B| \leqq \exp\left(-\frac{kp}{2}\right)$$

so that (2.14) gives

$$\|h\|_q \leqq \beta^{-k}\left((ep)^{k/2} + (4q)^{k/2} e(v,x)^{1/2q} \exp\left(-\frac{kp}{8q}\right)\right). \tag{2.16}$$

Proposition 2.5. *For each $k \geqq 1$, each $q \geqq 1$, each numbers $(\alpha_I)_{I \in S_k}$ with $\sum_{I \in S_k} \alpha_I^2 \leqq 1$, we have*

$$\left(\int\left(\sum_{I \in S_k} \alpha_I h_I(x, y)\right)^q dv(y)\right)^{1/q}$$
$$\leqq \beta^{-k}((4q)^{k/2} + \left(2e \max\left(1, \frac{2}{k}\log(e(v,x))\right)^{k/2}\right) . \tag{2.17}$$

Proof. In (2.16), we choose $p = \max(2, \frac{4}{k}\log(e(v,x)))$. □

The case $q = k = 1$ of this statement is very comparable to (1.6). If we fix k, take $q = 1$, optimization over α_I show that we control

$$\left(\sum_{I \in S_k}\left(\int h_I(x, y) dv(y)\right)^2\right)^{1/2}$$

by the right-hand side of (2.17). If we fix $k = 1$, (2.17) indicates a rate of growth of the moments of order q of the function $\sum_{i \leqq n} \alpha_i h_i(x, y)$ in $q^{1/2}$; this indicates a control in the L_{ψ_2} norm. More generally, using Chebyshev inequality, (2.17) implies that for $\sum_{I \in S_k} \alpha_I^2 \leqq 1$, we have

$$v\left(\left\{y; \sum_{I \in S_k} \alpha_I h_I(x, y) \geqq t\right\}\right) \leqq \exp\left(-\frac{k\beta^2}{2e}\left(\frac{t}{2}\right)^{2/k}\right) \tag{2.18}$$

whenever

$$t \geqq (\sqrt{e}/\beta)^k 2e \max\left(1, \frac{2}{k}\log(e(v,x))\right) . \tag{2.19}$$

We now turn to the reformulation of $e(\nu,x)$. First we observe that

$$e(\nu,x) = \int \left(\int \theta_{x,y}(\varepsilon)d\nu(y)\right)^2 d\varepsilon$$
$$= \int \theta_{x,y_1}(\varepsilon)\theta_{x,y_2}(\varepsilon)d\nu(y_1)d\nu(y_2)d\varepsilon . \tag{2.20}$$

Now, by expending the product,

$$\theta_{x,y}(\varepsilon) = \sum_{I\in S} \beta^{|I|} r_I(\varepsilon)h_I(x,y)$$

so that integrating in (2.20) in ε gives, (since $\int r_I r_J d\varepsilon = 1$ if $I = J$ and zero otherwise)

$$e(\nu,x) = \sum_{I\in S} \int \beta^{2|I|} h_I(x,y_1)h_I(x,y_2)d\nu(y_1)d\nu(y_2) . \tag{2.21}$$

so that

$$e(\nu,x) = \sum_{I\in S} \beta^{2|I|} \left(\int h_I(x,y_1)d\nu(y_1)\right)^2 . \tag{2.22}$$

Another expression of interest following from (2.21) is

$$e(\nu,x) = \int \prod_i (1+\beta^2 h_i(x,y_1)h_i(x,y_2))d\nu(y_1)d\nu(y_2) .$$

We now observe that $h_i(x,y_1)h_i(x,y_2)$ takes only values zero or one. For $u \in \{0,1\}$, we have $1+\beta^2 u = (1+\beta^2)^u$ so that

$$e(\nu,x) = \int (1+\beta^2)^{\sum_{i\leqq n} h_i(x,y_1)h_i(x,y_2)} d\nu(y_1)d\nu(y_2) \tag{2.23}$$

and, as promised, Theorem 1.1 corresponds to the case $\beta = 1/2$ of Theorem 2.1.

We now turn to the proof of Theorem 1.2, of which we keep the notation. We set $\Omega = \{-1,1\}$, with the uniform measure.

For $\varepsilon \in \Omega^n$, let us consider

$$Y(\varepsilon) = \sup_{x^*\in W_1^*} \left(\sum_{i\leqq n}\left(\sum_{j\leqq n} \varepsilon_j x^*(x_{ij})\right)^2\right)^{1/2} .$$

This auxiliary random variable will play an important part, and we need a deviation inequality for it. This inequality could be deduced from the results of [T3], but it is more instructive to proceed directly as a preparation for the main argument.

We have defined $V = EY$, so $P(Y \leqq 2V) \geqq 1/2$, and the set $B = \{\varepsilon; Y(\varepsilon) \leqq 2V\}$ has probability at least 1/2. We will prove that

$$Y(\varepsilon) \leqq 2V + 2U\sqrt{f(B,\varepsilon)} \tag{2.24}$$

where the functional f is defined in (1.1). First, we observe that

$$
\begin{aligned}
Y(\varepsilon) &= \sum_{x^* \in W_1^*} \sup_{\Sigma \beta_j^2 \leqq 1} \sum_{i,j \leqq n} \beta_j x^*(x_{ij}) \varepsilon_i \\
&= \sup_{\tau \in \mathscr{F}} \sum_{i \leqq n} \tau_i \varepsilon_i
\end{aligned}
$$

where

$$
\mathscr{F} = \left\{ \left(\sum_{j \leqq n} \beta_j x^*(x_{ij}) \right)_{i \leqq n} ; \sum_{j \leqq n} \beta_j^2 \leqq 1, x^* \in W_1^* \right\} .
$$

We observe that

$$
\tau \in \mathscr{F} \Rightarrow \sum_{i \leqq n} \tau_i^2 \leqq U^2 . \tag{2.25}
$$

Let us now fix ε, and $\tau \in \mathscr{F}$ such that $Y(\varepsilon) = \sum_{i \leqq n} \tau_i \varepsilon_i$.

Consider a probability measure ν on B. Then

$$
\begin{aligned}
Y(\varepsilon) - 2V &\leqq \int \left(\sum_{i \leqq n} \tau_i \varepsilon_i - \sum_{i \leqq n} \tau_i \eta_i \right) d\nu(\eta) \\
&\leqq 2 \sum_{i \leqq n} \tau_i \int h_i(\varepsilon, \eta) d\nu(\eta) \\
&\leqq 2U \sqrt{f(\nu, \varepsilon)}
\end{aligned}
$$

where we have used that $|\varepsilon_i - \eta_i| = 2h_i(\varepsilon, \eta)$, (2.25) and Cauchy–Schwarz. Thus (2.24) is proved.

We now start the main argument.

For $\eta \in \Omega^n$, we set $Z(\eta) = \|\sum_{i,j \leqq n} x_{ij} \eta_i \eta_j\|$. We consider a number a, and we set $A = \{\eta \in \Omega^n ; Z(\eta) \leqq a\}$. Let us now fix $\varepsilon \in \Omega^n$, and fix y^* in W_1^* such that

$$
Z(\varepsilon) = \sum_{i,j \leqq n} y^*(x_{ij}) \varepsilon_i \varepsilon_j .
$$

Consider a probability measure ν on A. Then

$$
Z(\varepsilon) - a \leqq \int \sum_{i,j \leqq n} y^*(x_{ij})(\varepsilon_i \varepsilon_j - \eta_i \eta_j) d\nu(\eta) \tag{2.26}
$$

We write, since $\varepsilon_i - \eta_i = 2\varepsilon_i h_i(\varepsilon, \eta)$,

$$
\begin{aligned}
\varepsilon_i \varepsilon_j - \eta_i \eta_j &= (\varepsilon_i - \eta_i)\varepsilon_j + (\varepsilon_j - \eta_j)\varepsilon_i - (\varepsilon_i - \eta_i)(\varepsilon_j - \eta_j) \\
&= \varepsilon_i \varepsilon_j [2h_i(\varepsilon, \eta) + 2h_j(\varepsilon, \eta) - 4h_i(\varepsilon, \eta) h_j(\varepsilon, \eta)] .
\end{aligned}
$$

Let us set

$$
\begin{aligned}
\alpha_i &= \int h_i(\varepsilon, \eta) d\nu(\eta) \\
\alpha_{ij} &= \int h_i(\varepsilon, \eta) h_j(\varepsilon, \eta) d\nu(\eta) .
\end{aligned}
$$

Thus

$$
\int (\varepsilon_i \varepsilon_j - \eta_i \eta_j) d\nu(\eta) = 2\varepsilon_i \varepsilon_j (\alpha_i + \alpha_j - 2\alpha_{ij}) .
$$

Combining with (2.26) and the fact that $x_{ij} = x_{ji}$ we get

$$Z(\varepsilon) - a \leqq 2 \sum_{i,j \leqq n} (\varepsilon_i \varepsilon_j \alpha_i y^*(x_{ij}) + \varepsilon_i \varepsilon_j \alpha_j y^*(x_{ij}) - 4\varepsilon_i \varepsilon_j \alpha_{ij} y^*(x_{ij}))$$

and thus

$$Z(\varepsilon) - a \leqq 2Z_1(\varepsilon) + Z_2(\varepsilon)$$

where

$$Z_1(\varepsilon) = 2 \left(\sum_{i \leqq n} \alpha_i^2 \right)^{1/2} Y(\varepsilon)$$

$$Z_2(\varepsilon) = 4 \left| \sum_{i,j \leqq n} y^*(x_{ij}) \varepsilon_i \varepsilon_j \alpha_{ij} \right| .$$

We will now use the functional of Theorem 2.1. We choose $\beta = 1/2$, so that the right-hand side of (2.2) is at most $P(A)^{-1}$. We will denote by K a universal constant, not necessarily the same at each occurrence.

As pointed out after Proposition 2.5, the case $q = k = 1$ of this proposition implies

$$\left(\sum_{i \leqq n} \alpha_i^2 \right)^{1/2} \leqq K(1 + \sqrt{\log e(v, \varepsilon)})$$

and combining with (2.24) we get

$$Z_1(\varepsilon) \leqq K(1 + \sqrt{\log e(A, \varepsilon)})(V + U\sqrt{f(B, \varepsilon)}) . \tag{2.27}$$

Thus, we turn to the study of $Z_2(\varepsilon)$. We have $Z_2(\varepsilon) \leqq 4\|h\|_1$, where

$$h(\eta) = \sum_{i,j \leqq n} y^*(x_{ij}) \varepsilon_i \varepsilon_j h_i(\varepsilon, \eta) h_j(\varepsilon, \eta)$$

and where the norm is computed in $L_1(v)$. It then follows from (2.14) that for any $t > 0$, we have

$$\|h\|_1 \leqq t + \|T(h)\|_4 e(v, x)^{1/2} \{|T(h)| \geqq t\}^{1/4} \tag{2.28}$$

where

$$T(h) = \sum_{i,j \leqq n} y^*(x_{ij}) \varepsilon_i \varepsilon_j r_i r_j \beta^{-2}$$

To estimate the last term of (2.28), we need the following.

Lemma 2.6. [L-T, 3.7] *Let us set*

$$U' = \sup \left\{ \sum_{i,j \leqq n} \alpha_i y^*(x_{ij}) \beta_j ; \sum_{i \leqq n} \alpha_i^2 \leqq 1 \sum_{j \leqq n} \beta_j^2 \leqq 1 \right\}$$

$$V' = \left(\sum_{i,j \leqq n} y^*(x_{ij})^2 \right)^{1/2} .$$

Then

$$P\{|T(h)| \geqq t\} \leqq 2 \exp \left(-\frac{1}{K} \min \left(\frac{t^2}{V'^2}, \frac{t}{U'} \right) \right) . \tag{2.29}$$

It must be pointed out that the estimate given by Lemma 2.4 yields only a bound $2\exp(-t/KV')$, which is not sufficient.

To continue the argument, we observe that $U' \leqq U$; also $V' \leqq KV$. The second inequality is less obvious, but follows from the fact that if

$$H = \left(\sum_{j \leqq n} \left(\sum_{i \leqq n} \varepsilon_i y^*(x_{ij}) \right)^2 \right)^{1/2}, \quad \text{then } EH^2 \leqq K(EH)^2 \leqq KV^2$$

(See [L-T, 3.2]).

We now combine (2.28), (2.29), with

$$t = K \max(U \log e(v,x), V(\log e(v,x))^{1/2})$$

to get (using (2.15) for $p = 4$)

$$Z_2(\varepsilon) \leqq K \max(UX^2, VX)$$

where $X = X(\varepsilon) = (1 + \log e(A,\varepsilon))^{1/2}$. Thus, by (2.21),

$$\int e^{X^2} dP \leqq \frac{e}{P(A)} . \tag{2.30}$$

We now set $F = F(\varepsilon) = \sqrt{f(B,\varepsilon)}$, so that, by (2.2),

$$\int e^{F^2/4} dP \leqq \frac{1}{P(B)} \leqq 2 \leqq \frac{e}{P(A)} \tag{2.31}$$

and we rewrite (2.27) as

$$Z_1(\varepsilon) \leqq KX(V + UF) .$$

Thus we have shown that

$$\begin{aligned} Z(\varepsilon) - a &\leqq KX(V + UF) + K\max(UX^2, VX) \\ &\leqq K \max(U(X^2 + F^2), VX) . \end{aligned}$$

We now use (2.30), (2.31) (and let ε vary) to get

$$P(Z - a \geqq t) \leqq \frac{K}{P(A)} \exp\left(-\frac{1}{K} \min\left(\frac{t}{U}, \frac{t^2}{V} \right) \right) .$$

If we take $a = M$, we have $P(A) \geqq \frac{1}{2}$, and thus

$$P(Z \geqq M + t) \leqq K \exp\left(-\frac{1}{K} \min\left(\frac{t}{U}, \frac{t^2}{V} \right) \right) .$$

If we now take $a = M - t$, then $A = \{Z \leqq M - t\}$ and $P(Z - a \geqq t) \geqq 1/2$, so that

$$P(Z \leqq M - t) \leqq K \exp\left(-\frac{1}{K} \min\left(\frac{t}{U}, \frac{t^2}{V} \right) \right) . \qquad \square$$

We now sketch a general method to provide bounds for $|Z-M|$, where Z is a random variable on a product space Ω^n, of median M. Let us say that a function

$$h_x(y)=\sum_{I\in S}\alpha_I h_I(x,y)$$

on Ω^n is a *subchaos* for Z at x if it satisfies the inequality

$$\forall y\in\Omega^n,\qquad Z(y)\geqq Z(x)-h_x(y)\,.$$

Then, if $A=\{y\in\Omega^n;Z(y)\leqq a\}$, and if v is a probability measure on A, we have

$$\begin{aligned}Z(x)&\leqq a+\int h_x(y)dv(y)\\&\leqq a+\|h_x\|_{L_1(v)}\,.\end{aligned}$$

Now, the norm $\|h_x\|_{L_1(v)}$ can be controlled in function of $e(A,x)$ and of the function $T_x=\sum_{I\in S}\alpha_I\beta^{-|I|}r_I$ (that depends upon x through the coefficients α_I) on $\{0,1\}^n$, as is done e.g. in (2.14). The simpler the subchaos can be found, the better the resulting bound. The simplest possible case is when at each point x a subchaos for Z of the type

$$h_x(y)=\sum_{i\leqq n}\alpha_i h_i(x,y)$$

can be found, where $\sum_{i\leqq n}\alpha_i^2\leqq\sigma^2$ for all x in Ω^n. This is in particular the case when $\Omega\subset[-1,1]$ and

$$Z(x)=\sup_{\tau\in\mathscr{F}}\sum_{i\leqq n}\tau_i x_i$$

where $\sum_{i\leqq n}\tau_i^2\leqq\sigma^2/4$ for each $\tau=(\tau_i)_{i\leqq n}$ in $\mathscr{F}$. This is the situation of (2.24) (although we did use the approach through the functional f, approach that is simpler in this case). The case of Theorem 1.2 is somehow the second simplest possible case.

3. Varying patterns

Yet another way to formulate concentration of measure in product spaces is as follows. The sets $\Delta(x,y)=\{i;h_i(x,y)=1\}$ vary widely as y varies in A, producing many different patterns. For example, a consequence of (2.18) is that, for any set I,

$$v(\{y;\operatorname{card}(I\cap\Delta(x,y))\geqq t\})\leqq\exp\left(-\frac{\beta^2t^2}{8es}\right)\tag{3.1}$$

when $t/\sqrt{s}$ satisfies (2.19), and where $s=\operatorname{card}I$. This means that the possible patterns $\Delta(x,y)$ do not concentrate too much in any given set I.

In this section we show how to improve upon (3.1). Why this is a natural question should be obvious; its importance will be demonstrated through the proof of Theorem 1.3.

Somewhat unexpectedly, it turns out that the basic scheme of proof that works flawlessly in [T3] and through the rest of the present paper breaks down. At the expense of considerable refinements we will be able to prove a near optimal result in the most important case, $\Omega = \{0,1\}$, $\mu(\{1\}) = \mu(\{0\}) = \frac{1}{2}$. The proof however does not extend to the case $\mu(\{0\}) = p$, p small, with the correct dependence in p.

Theorem 3.1. *There exists a number L with the following property. For each set $A \subset \{0,1\}^n$, and each point x of $\{0,1\}^n$, we can find a probability measure $\nu_{A,x}$ on A, and a number $p(A,x)$ such that for all numbers $(\alpha_i)_{i \leqq n}, \alpha_i \leqq 1$, we have*

$$\int \exp \frac{1}{L} \sum_{i \leqq n} \alpha_i h_i(x,y) d\nu_{A,x}(y) \leqq \exp \sqrt{\sum_{i \leqq n} \alpha_i^2}\,(p(A,x) + \sqrt{\log en}) \tag{3.2}$$

and moreover

$$\int \exp p(A,x)^2 dP(x) \leqq \frac{1}{P(A)} \,. \tag{3.3}$$

We first comment upon this result. One of the purposes of (3.2) is to obtain for any set I

$$\nu_{A,x}(\{y; \operatorname{card}(I \cap \Delta(x,y)) \geqq t\}) \leqq \exp\left(-\frac{t}{2L}\right)$$

when

$$t \geqq 2L\sqrt{\operatorname{card} I}\,(p(A,x) + \sqrt{\log en}) \,. \tag{3.4}$$

Under (3.4), this is a marked improvement upon (3.1).

It is very likely that the term $\sqrt{\log en}$ is not needed in (3.2). We will explain in the course of the proof while this term apparently cannot be removed with the current approach. On the other hand it can be argued that, for applications, this term is not disastrous. In such applications, we often make n vary, and we ask that the measures of the exceptional sets be summable; it is precisely the values of $p(A,x)$ of order $\sqrt{\log n}$ that achieve this result.

Before we start the main chain of arguments, we make a simple observation. We have, for any probability ν on A that

$$\int \exp \frac{1}{L} \sum_{i \leqq n} \alpha_i h_i(x,y) d\nu(y) \leqq \exp \frac{\sum_{i \leqq n} \alpha_i}{L} \leqq \exp \sqrt{\sum_{i \leqq n} \alpha_i^2}\left(\frac{\sqrt{n}}{L}\right) .$$

Thus, we can assume that

$$x \notin A \Rightarrow p(A,x) \leqq \frac{\sqrt{n}}{L} \,. \tag{3.5}$$

Obviously, we can take $p(A,x) = 0$ if $x \in A$. Thus

$$\int \exp p(A,x)^2 dP(x) \leqq (1 - P(A)) \exp \frac{n}{L^2} + P(A) \,.$$

We note that

$$x \leqq 1 \Rightarrow 2(1-x) + x = 2 - x \leqq \frac{1}{x} \,.$$

We have thus proved

Lemma 3.2. *If $n \leqq L^2 \log 2$, conditions* (3.2), (3.3) *hold.*

The proof of Theorem 3.1 will be by induction upon n; in the induction step, it will be essential to use a clever choice of the "new" coordinate. This coordinate will be chosen through the next Lemma, that is part of the folklore. We fix $n \geqq 1$; we denote by P' the product measure on $\{0,1\}^{n+1}$, and by P the measure on $\{0,1\}^n$.

Lemma 3.3. *Consider a set* $A \subset \{0,1\}^{n+1}$, *and* $a = P'(A)$. *For* $j \in \{0,1\}$, *set*

$$A_j = \{x \in \{0,1\}^n; (x,j) \in A\}$$

and $a_j = P(A_j)$. *After permutation of the coordinates, we can assume that*

$$|a_0 - a_1| \leqq \frac{ea}{\sqrt{n}} \sqrt{\log \frac{e}{a}} \ . \tag{3.6}$$

Proof. Consider the functions $r_i = 2x_i - 1$ on Ω^{n+1}, and set $\alpha_i = \int_A r_i dP'$. Then, for $p \geqq 1$

$$\sum_{i \leqq n+1} \alpha_i^2 = \int_A \sum_{i \leqq n+1} \alpha_i r_i dP' \leqq \left\| \sum_{i \leqq n+1} \alpha_i r_i \right\|_p \|1_A\|_q \ .$$

Using (2.15) we have, if $p \geqq 2$,

$$\left\| \sum_{i \leqq n+1} \alpha_i r_i \right\|_p \leqq \sqrt{p-1} \left(\sum_{i \leqq n+1} \alpha_i^2 \right)^{1/2} .$$

Thus

$$\sqrt{n} \min_{i \leqq n+1} |\alpha_i| \leqq \left(\sum_{i \leqq n+1} \alpha_i^2 \right)^{1/2} \leqq \sqrt{p-1} a^{1/q} = a\sqrt{p-1} \exp \left(\frac{\log 1/a}{p} \right) .$$

Taking $p = 1 + \log e/a \geqq 2$ concludes the proof. □

Lemma 3.4. *To prove Theorem* 3.1, *one can assume*

$$|a_0 - a_1| \leqq \frac{a}{10} \ .$$

Proof. Otherwise, by (3.6) we have $e/a \geqq \exp \frac{n}{100e^2}$.

Now, by (3.5) we have

$$\int \exp p^2(A,x) dP(x) \leqq \exp \left(\frac{n}{L} \right) \leqq \frac{1}{a}$$

as soon as $L \geqq 200e^2$, $n \geqq 200e^2$. But by Lemma 3.2 there is nothing to prove if $n \leqq 200e^2$. □

We now start the induction step from n to $n+1$, with the notation of Lemma 3.3. Thus, for $j \in \{0,1\}$, and $x \in \{0,1\}^n$, we are given a number $p_j(x)$ and a probability measure $\nu_{j,x}$ on A_j such that for each numbers $(\alpha_i)_{i \leqq n}$,

$$\int \exp \frac{1}{L} \sum_{i \leqq n} \alpha_i h_i(x,y) d\nu_{j,x}(y) \leqq \exp \sqrt{\sum_{i \leqq n} \alpha_i^2} (p_j(x) + \sqrt{\log en}) \tag{3.7}$$

and moreover

$$\int \exp p_j^2(x)\,dP(x) \leqq \frac{1}{a_j}\,. \tag{3.8}$$

We think to L as a parameter, that will be determined in the course of the proof. We assume without loss of generality that $a_0 \leqq a_1$.

For simplicity, we will identify A_j with $A_j \times \{j\}$, to view it as a subset of A.

For each $x \in \Omega^n$, we will define

$$v_{(x,0)} = v_{0,x}$$

and

$$v_{(x,1)} = \mu v_{0,x} + (1-\mu)v_{1,x}$$

where $\mu = \mu(x)$ will be chosen appropriately.

First, we note that for any $x \in \{0,1\}^n$, and since $h_{n+1}((x,0),z) = 0$ for z in A_0, we have

$$\int \exp \frac{1}{L}\sum_{i\leqq n+1} \alpha_i h_i((x,0),z)\,dv_{(x,0)}(z) \leqq \exp\sqrt{\sum_{i\leqq n+1}\alpha_i^2}\left(\sqrt{\log en} + p_0(x)\right)\,. \tag{3.10}$$

We will denote by $p(x)$ the smallest number such that, for a suitable choice of μ, for all $(\alpha_i)_{i\leqq n+1}$, $\alpha_i \leqq 1$, we have

$$\int \exp \frac{1}{L}\sum_{i\leqq n+1} \alpha_i h_i((x,1),z)\,dv_{(x,1)}(z) \leqq \exp\sqrt{\sum_{i\leqq n+1}\alpha_i^2}\left(\sqrt{\log en} + p(x)\right). \tag{3.11}$$

We do not take advantage of the fact that it would suffice to have a term $\sqrt{\log e(n+1)}$ rather that $\sqrt{\log en}$ in (3.11). We will find upper bounds for $p(x)$, and then we will show that

$$\int_{\Omega^n} \exp p_0^2(x)\,dP(x) + \int_{\Omega^n} \exp p^2(x)\,dP(x) \leqq \frac{2}{a}\,. \tag{3.12}$$

This will conclude the proof.

Let us denote by $R((\alpha_i)_{i\leqq n+1})$ the left-hand side of (3.11). Then, by definition of $v_{(x,1)}$

$$\begin{aligned} R((\alpha_i)_{i\leqq n+1}) &= \mu \exp\frac{\alpha_{n+1}}{L}\int \exp\left(\frac{1}{L}\sum_{i\leqq n}\alpha_i h_i(x,y)\right) dv_{0,x}(y) \\ &\quad + (1-\mu)\int \exp\left(\frac{1}{L}\sum_{i\leqq n}\alpha_i h_i(x,y)\right) dv_{1,x}(y)\,. \end{aligned}$$

We now use the induction hypothesis. For simplicity, we write p_j rather than $p_j(x)$ when no ambiguity arises. We have

$$\begin{aligned} R((\alpha_i)_{i \leqq n+1}) \ &\leqq \mu \exp\left(\frac{\alpha_{n+1}}{L} + \alpha(\sqrt{\log en} + p_0)\right) \\ &\quad + (1-\mu)\exp(\alpha(\sqrt{\log en} + p_1)) \end{aligned} \tag{3.13}$$

where we have set $\alpha^2 = \sum_{i \leqq n} \alpha_i^2$.

Lemma 3.5. $p(x) \leqq p_1(x)$

Proof. Take $\mu = 0$. □

Lemma 3.6. *We have*

$$p(x) \leqq p_0(x) + \frac{1}{2L^2(p_0(x) + \sqrt{\log en})} . \tag{3.14}$$

Proof. To prove this, we take $\mu = 1$ in (3.13). We set

$$\beta = \sqrt{\sum_{i \leqq n+1} \alpha_i^2} \tag{3.15}$$

so that

$$\alpha = \sqrt{\beta^2 - \alpha_{n+1}^2} = \beta\sqrt{1 - \frac{\alpha_{n+1}^2}{\beta^2}} \leqq \beta\left(1 - \frac{\alpha_{n+1}^2}{2\beta^2}\right) \tag{3.16}$$

an observation that will be used again. Thus (3.13) yields

$$R((\alpha_i)_{i \leqq n+1}) \leqq \exp(\beta(\sqrt{\log en} + p_0 + C))$$

where

$$C = -\frac{\alpha_{n+1}^2}{2\beta^2}(p_0 + \sqrt{\log en}) + \frac{1}{L}\frac{\alpha_{n+1}}{\beta}$$

The result follows by optimization over α_{n+1}/β. □

We now turn to the investigation of what can be done with a choice of $0 < \mu < 1$. It would be nice at this point if the exponential function were concave. But the inequality

$$\mu \exp u + (1-\mu)\exp v \leqq \exp(\mu u + (1-\mu)v)$$

is not quite true, and our first task is to look for substitutes.

Lemma 3.7. *If* $0 \leqq \mu \leqq 1$ *and* $u < v$, *we have*

$$\mu \exp u + (1-\mu)\exp v \leqq \exp\left[v - \frac{\mu}{e}\min(1, v-u)\right] . \tag{3.17}$$

Proof. We set $t = v - u$. Assume first that $v \leqq u + 1$. Then (3.17) reduces to

$$\mu + (1-\mu)\exp t \leqq \exp\left(t\left(1 - \frac{\mu}{e}\right)\right) . \tag{3.18}$$

Now, for $t' \leqq t$, we have

$$\exp(t - t') \geqq \exp t - t'e$$

and thus

$$\exp t\left(1 - \frac{\mu}{e}\right) \geqq (\exp t) - t\mu \geqq \exp t - \mu(\exp t - 1) ,$$

so that (3.18) is true.

Assume now that $v \geqq u + 1$. Then (3.17) means that

$$\mu + (1-\mu)\exp t \leqq \exp\left(t - \frac{\mu}{e}\right) .$$

But, since $1 \leqq \frac{1}{e}\exp t$ we have

$$\begin{aligned}(1-\mu)\exp t + \mu &\leqq (1-\mu)\exp t + \frac{\mu}{e}\exp t \\ &= \left(1 - \mu\left(1 - \frac{1}{e}\right)\right)\exp t \\ &\leqq \exp\left(t - \mu\left(1 - \frac{1}{e}\right)\right)\end{aligned}$$

□

Lemma 3.8. *If* $0 \leqq \mu \leqq 1,\ u < v,\ w \leqq 1,$ *we have*

$$\mu\exp(u+w) + (1-\mu)\exp v \leqq \exp\left(v - \frac{\mu}{3}\min(1, v-u) + 2\mu w\right) . \tag{3.19}$$

Proof. If $u + w < v$, we use Lemma 3.7 and we observe that

$$\min(1, v - (u+w)) \geqq \min(1, v-u) - w .$$

If $u + w > v$, the left-hand side is

$$\begin{aligned}e^v(1 + \mu(e^{u+w-v} - 1)) &\leqq e^v(1 + 2\mu(u + w - v)) \\ &\leqq \exp(v + 2\mu(u + w - v)) \\ &\leqq \exp(v - 2\mu(v-u) + 2\mu w) ,\end{aligned}$$

where we have used that $e^x \leqq 1 + 2x$ for $x \leqq 1$. □

We now pursue the study of (3.13). We use (3.19) to get

$$\begin{aligned}R((\alpha_i)_{i \leqq n+1}) &\leqq \exp\left[\alpha(\sqrt{\log en} + p_1) + \frac{2\mu\alpha_{n+1}}{L} - \frac{\mu}{e}\min(1, \alpha(p_1 - p_0))\right] \\ &\leqq \exp\beta(\sqrt{\log en} + p_1 - H((\alpha_i)_{i \leqq n+1}))\end{aligned}$$

where

$$H((\alpha_i)_{i\leqq n+1}) = \left(1-\frac{\alpha}{\beta}\right)(\sqrt{\log en}+p_1) - \frac{2\alpha_{n+1}\mu}{L\beta} \tag{3.20}$$
$$+\frac{\mu}{e\beta}\min(1,\alpha(p_1-p_0))\,.$$

Thus we have

$$p \leqq p_1 - \inf H((\alpha_i)_{i\leqq n+1})\,. \tag{3.21}$$

Lemma 3.9. *If*

$$L(p_1(x)-p_0(x))(p_1(x)+\sqrt{\log en}) \leqq 1 \tag{3.22}$$

we have

$$p(x) \leqq p_1(x) - \frac{L}{24}(p_1(x)+\sqrt{\log en})(p_1(x)-p_0(x))$$
$$\times\min\left(\frac{1}{\sqrt{n}}, p_1(x)-p_0(x)\right)\,.$$

Proof. Under (3.22),

$$\mu = \frac{L}{4}(p_1-p_0)(p_1+\sqrt{\log en}) \leqq 1\,. \tag{3.23}$$

We note also that (3.22) implies $(p_1-p_0)L \leqq 1$.

Case 1. $\alpha(p_1-p_0) \geqq 1$. Assuming $L \geqq 2e$, since $\alpha_{n+1} \leqq 1$, we see from (3.20) that

$$H((\alpha_i)_{i\leqq n+1}) \geqq \frac{\mu}{2e\beta} \geqq \frac{\mu}{2e\sqrt{n+1}}$$
$$\geqq \frac{L}{24\sqrt{n}}(p_1-p_0)(p_1+\sqrt{\log n})\,.$$

Case 2. $\alpha(p_1-p_0) \leqq 1$, $\alpha \geqq \beta/2$.

Then using (3.16) we have

$$H((\alpha_i)_{i\leqq n+1}) \geqq \frac{\mu}{e\beta}\alpha(p_1-p_0) + \inf_t\left\{\frac{t^2}{2}(\sqrt{\log en}+p_1) - \frac{2t\mu}{L}\right\}$$
$$\geqq \frac{\mu(p_1-p_0)}{2e} - \frac{\mu^2}{L^2}\frac{1}{(p_1+\sqrt{\log en})}$$
$$\geqq \frac{\mu(p_1-p_0)}{4e}$$
$$\geqq \frac{L}{16e}(p_1-p_0)^2(p_1+\sqrt{\log en})\,.$$

Case 3. $\alpha(p_1 - p_0) \leqq 1$, $\alpha \leqq \beta/2$. Then

$$\begin{aligned} H((\alpha_i)_{i\leqq n+1}) &\geqq \frac{1}{2}(p_1 + \sqrt{\log en}) - \frac{2\mu}{L} \\ &\geqq \frac{1}{4}(p_1 + \sqrt{\log en}) \\ &\geqq \frac{L^2}{4}(p_1 - p_0)^2(p_1 + \sqrt{\log en}) . \end{aligned} \tag{3.24}$$

$\square$

It remains to prove (3.12). It is convenient to proceed by contradiction. So we assume that (3.12) fails. Using (3.8), we then get

$$\int (\exp p^2 - \exp p_0^2)\, dP > 2\left(\frac{1}{a} - \frac{1}{a_0}\right) \tag{3.25}$$

$$\int (\exp p_1^2 - \exp p^2)\, dP \leqq \frac{1}{a_0} + \frac{1}{a_1} - \frac{2}{a} = \frac{4(a_1 - a_0)^2}{aa_0a_1} . \tag{3.26}$$

We will show that these two relations are not compatible. Certainly it seems natural to find a lower bound for $\exp p_1^2 - \exp p^2$; this will be done using Lemmas 3.6 and 3.9.

We consider a parameter M that will be adjusted later.

We now define a partition $(U_\ell)_{\ell \leqq 5}$ of Ω^n. We set

$$\begin{aligned} U_1 &= \{x;\, p_1(x) \leqq p_0(x)\} \\ U_2 &= \{x \notin U_1;\, L(p_1(x) - p_0(x))(p_1(x) + \sqrt{\log en}) > 1\} \\ U_3 &= \left\{x \notin U_1 \cup U_2;\, p_1(x) - p_0(x) \leqq \frac{1}{\sqrt{n}}\right\} \\ U_4 &= \{x \notin U_1 \cup U_2 \cup U_3;\, p_1(x) + \sqrt{\log en} \geqq M\} \\ U_5 &= \{0,1\}^n \backslash (U_1 \cup U_2 \cup U_3 \cup U_4) . \end{aligned}$$

For $1 \leqq \ell \leqq 5$, we set

$$\begin{aligned} J_\ell &= \int_{U_\ell} \exp p^2(x)\, dP(x) \\ S_{\ell,j} &= \int_{U_\ell} \exp p_j^2(x)\, dP(x) . \end{aligned}$$

Thus we can rewrite (3.25) and (3.26) respectively as

$$\sum_{1 \leqq \ell \leqq 5} J_\ell - S_{\ell,0} > 2\left(\frac{1}{a} - \frac{1}{a_0}\right) \tag{3.27}$$

$$\sum_{1 \leqq \ell \leqq 5} S_{\ell,1} - J_\ell \leqq \frac{4(a_1 - a_0)^2}{aa_1a_0} . \tag{3.28}$$

By Lemma 3.5, we have $J_1 - S_{1,0} \leqq 0$. Thus there exists $2 \leqq \ell \leqq 5$ such that $J_\ell - S_{\ell,0} \geqq (a_0 - a)/2aa_0$. We will show that the corresponding term

$S_{\ell_1,1} - J_\ell$ is larger than the right-hand side of (3.28). Since, by Lemma 3.5, each of the terms $S_{\ell,1} - J_\ell$ is positive, this finishes the proof. Thereby the task is for $2 \leqq \ell \leqq 5$ to find a lower bound of $S_{\ell,1} - J_\ell$ related to $J_\ell - S_{\ell,0}$.

Lemma 3.10. *We have* $J_2 - S_{2,0} \leqq S_{2,1} - J_2$.

Proof. This means that $J_2 \leqq (S_{2,1} + S_{2,0})/2$. By convexity of the exponential, it suffices to prove that $p^2(x) \leqq (p_0^2(x) + p_1^2(x))/2$ for x in U_2, i.e.

$$p^2(x) - p_0^2(x) \leqq \frac{1}{2}(p_1^2(x) - p_0^2(x)) . \tag{3.29}$$

First, by (3.14), we have

$$p^2 - p_0^2 = (p + p_0)(p - p_0) \leqq 2p(p - p_0) \leqq \frac{p}{L^2(p_0 + \sqrt{\log en})} .$$

Next, since $x \in U_2$, we have

$$L(p_1 - p_0)(p_1 + \sqrt{\log en}) \geqq 1$$

so that

$$p_1^2 - p_0^2 \geqq p_1(p_1 - p_0) \geqq \frac{p_1}{L(p_1 + \sqrt{\log en})} .$$

Therefore to prove (3.29) it suffices to show that

$$\frac{1}{2}\left(\frac{p_1}{p_1 + \sqrt{\log en}}\right) \geqq \frac{p}{L(p_0 + \sqrt{\log en})} .$$

If $p_1 \leqq \sqrt{\log en}$, we have

$$\frac{1}{2}\frac{p_1}{p_1 + \sqrt{\log en}} \geqq \frac{1}{4}\frac{p_1}{\sqrt{\log en}} \geqq \frac{p_1}{L(p_0 + \sqrt{\log en})}$$

provided $L \geqq 4$. If $p_1 \geqq \sqrt{\log en}$, we have

$$\frac{1}{2}\frac{p_1}{p_1 + \sqrt{\log en}} \geqq \frac{1}{4} \geqq \frac{1}{L}\frac{p}{(p_0 + \sqrt{\log en})}$$

since $p \leqq p_0 + 1$ by (3.14). □

Lemma 3.11. *We have*

$$S_{3,1} - J_3 \geqq \frac{L}{K}\frac{(S_{3,1} - S_{3,0})^2}{S_{3,1}} .$$

Proof. For x in U_3, by Lemma 3.9 we have

$$p \leqq p_1 - \frac{L}{24}(p_1 - p_0)^2(p_1 + \sqrt{\log en}) .$$

We will however only use the weaker information

$$p \leqq p_1 - \frac{L}{24} p_1(p_1 - p_0)^2 . \tag{3.30}$$

Since $x \notin U_2$, we have

$$L(p_1 - p_0)(1 + p_1) \leqq 1 ,$$

so in particular $L(p_1 - p_0)^2 \leqq 1$. For $u \leqq v$, we have $(v-u)^2 \leqq v - uv$. Thus, from (3.30) we have

$$p^2 \leqq p_1^2 - \frac{L}{24} p_1^2(p_1 - p_0)^2 .$$

We use that $e^{-x} \leqq 1 - x/2$ for $x \leqq 1$. Thus,

$$J_3 \leqq S_{3,1} - \frac{L}{48} \int_{U_3} p_1^2(p_1 - p_0)^2 \exp p_1^2 dP. \tag{3.31}$$

Using Cauchy–Schwarz, we have

$$\left(\int_{U_3} p_1(p_1 - p_0) \exp p_1^2 dP \right)^2 \leqq S_{3,1} \int_{U_3} p_1^2(p_1 - p_0)^2 \exp p_1^2 dP . \tag{3.32}$$

Now using the fact that $e^x - 1 \leqq 2x$ for $x \leqq 1$, we have

$$\begin{aligned} p_1(p_1 - p_0) \geqq \frac{1}{2}(p_1 + p_0)(p_1 - p_0) &\geqq \frac{1}{2}(p_1^2 - p_0^2) \\ &\geqq \frac{1}{4}(\exp(p_1^2 - p_0^2) - 1) \end{aligned}$$

so that

$$\begin{aligned} \int_{U_3} p_0(p_1 - p_0) \exp p_1^2 dP &\geqq \int_{U_3} p_0(p_1 - p_0) \exp p_0^2 dP \\ &\geqq \frac{1}{4}(S_{3,1} - S_{3,0}) . \end{aligned} \tag{3.33}$$

Combining with (3.31) and (3.32) yield the result. □

Lemma 3.12. *We have*

$$S_{4,1} - J_4 \geqq \frac{ML}{8\sqrt{n}}(S_{4,1} - S_{4,0})$$

Proof. Since $x \in U_4$, by Lemma 3.9, we have

$$p \leqq p_1 - \frac{LM}{\sqrt{n}}(p_1 - p_0) . \tag{3.34}$$

Now, $L(p_1 - p_0)(p_1 + \sqrt{\log en}) \leqq 1$ since $x \notin U_1$, so that $L(p_1 - p_0)M \leqq 1$. Since $p_1 \geqq p_0 + 1/\sqrt{n}$, the right-hand side of (3.34) is nonnegative, so that

$$p^2 \leqq p_1^2 - \frac{LM}{\sqrt{n}} p_1(p_1 - p_0) .$$

Since $p_1 \leqq \sqrt{n}$, we have $LMp_1(p_1 - p_0)/\sqrt{n} \leqq 1$, and using that $e^{-x} \geqq 1 - x/2$ for $x \leqq 1$ we get

$$J_4 \leqq S_{4,1} - \frac{ML}{2\sqrt{n}} \int_{U_4} p_1(p_1 - p_0) \exp p_1^2 dP$$

and we conclude the proof using an inequality similar to (3.33). □

It is in this lemma that lies the weakness of the argument. Analysis of examples show that enough of the contribution to $\int(\exp p_1^2 - \exp p_0^2)dP$ occurs where the values of p_1 of order $\sqrt{\log e/a}$. If we could show that this is always the case, we could remove the parasitic $\sqrt{\log n}$ factor. Unfortunately we do not see how to carry this information in the induction hypothesis.

Lemma 3.13. *We have*

$$J_5 - S_5 \leqq \exp M^2 .$$

Proof. $J_5 \leqq \int_{U_5} \exp p_1^2 \leqq \exp M^2$. □

We are now ready to finish the proof of Theorem 3.1. Let us set $\rho = (a_0 - a_1)/a$, and recall that by (3.5)

$$\rho \leqq \frac{e}{\sqrt{n}} \sqrt{\log \frac{e}{a}} . \tag{3.35}$$

We now choose M as follows. If

$$\rho \leqq \frac{e}{\sqrt{n}} \sqrt{\log en^2} , \tag{3.36}$$

we choose $M = \frac{1}{2}\sqrt{\log en}$. If $\rho > \frac{2e}{\sqrt{n}}\sqrt{\log en^2}$, by (3.35), we have $a < n^{-2}$. We then choose $M = \frac{1}{2}\sqrt{\log 1/a}$. We observe that (at least for $n > 1$) we have

$$\rho \leqq \frac{KM}{\sqrt{n}} \tag{3.37}$$

As we have already observed, there exists $2 \leqq \ell \leqq 5$ such that $J_\ell - S_{\ell,0} \geqq (a_0 - a)/2aa_0 = \rho/4a_0$. We observe that $(a_0 - a_1)^2/aa_1a_0 = \rho^2 a/a_0a_1 \leqq \rho^2/a_1$.

Case 1. $\ell = 2$. By Lemma 3.10, we have $S_{2,1} - J_2 \geqq \rho/4a_0$, so it is enough to check that

$$\frac{\rho}{4a_0} \geqq \frac{\rho^2}{a_1} .$$

But, by (3.6) and Lemma 3.4, we can assume, $\rho \leqq \frac{1}{10}$, and $a_0 \leqq a_1 + \frac{1}{20}(a_0 + a_1)$, so

$$a_0 \leqq \frac{11}{9} a_1 , \tag{3.38}$$

and the conclusion follows.

Case 2. $\ell = 3$. By Lemma 3.11, and since $S_{3,1} \leqq 1/a_1$,

$$S_{3,1} - J_3 \geqq \frac{L}{32} \left(\frac{\rho e}{4a_0} \right)^2 a_1 \geqq \frac{\rho^2}{a_1}$$

using (3.38), and for L large enough.

Case 3. $\ell = 4$. Then by Lemma 3.12 we have

$$S_{4,1} - J_4 \geqq \frac{ML}{\sqrt{n}} \frac{\rho}{4a_0} .$$

Using (3.37), (3.38), this is bigger than ρ^2/a_1 if L is large enough.

Case 4. $\ell = 5$. Then we must have $U_5 \neq \emptyset$, and this occurs only in the case where (3.36) fails, and where $M = \frac{1}{2}\sqrt{\log 1/a}$. Then $a < n^{-2}$.

On the other hand, using Lemma 3.13, we have

$$\frac{\rho}{4a_0} \leqq J_5 - S_5 \leqq \exp M^2 = \frac{1}{a^{1/4}}$$

so that, since (3.36) fails, and using (3.38)

$$\frac{2e}{\sqrt{n}} \sqrt{\log en^2} \leqq \rho \leqq \frac{4a_0}{a^{1/4}} \leqq 5a^{3/4} \leqq 5n^{-3/2}$$

and thus $2ne\sqrt{\log en^2} \leqq 5$, which is a contradiction. □

We now turn to the discussion of Theorem 1.3. One essential feature of this theorem is that the matrices have entries zero or one. Suppose, for example, that we would replace $\{0, 1\}$ by $[0, 1]$ with Lebesgue measure. Then given a subset A of T, with $P(A) \geqq 1/2$, we have

$$P(\{x; d(x, A) \geqq t\}) \leqq \exp -\frac{t^2}{K} .$$

This follows from example from [P] after one dominates the operator norm by the Hilbert–Schmit norm. Going back to our situation, consider the set

$$A = \left\{ (y_{ij})_{i,j \leqq n}; \sum_{j \leqq n} y_{1j} \leqq \frac{n}{2} \right\} .$$

Consider x with $\sum_{j \leqq n} x_{1j} = \frac{n}{2} + q$. Then for any $y \in A$, the matrix $x - y$ has at least q ones on the first row, so that its norm is at least $\sqrt{q}$. On the other hand, for $0 \leqq u \leqq \sqrt{n}/2$, we have

$$P\left(\left\{x; \sum_{j \leqq n} x_{1j} \geqq \frac{n}{2} + u\sqrt{n}\right\}\right) \geqq \exp -\frac{u^2}{K}$$

so that

$$P(\{x; d(x,A) \geqq u\}) \geqq \exp\left(-\frac{u^4}{nK}\right) \tag{3.39}$$

for $u \leqq \sqrt{n}/K$.

The natural set one would like to try is

$$B = \left\{(y_{ij})_{i,j \leqq n}; \sum_{i,j \leqq n} y_{ij} \leqq \frac{n^2}{2}\right\}.$$

Consider x_{ij} with $\sum_{i,j \leqq n} x_{ij} = \frac{n^2}{2} + q$. Consider $I = \{(i,j);\ x_{ij} = 1\}$. For any susbet J of I of cardinal q, the matrix z_J that has ones in J and zeroes elsewhere is of the type $x - y$ for y in A. But we can locate J such that $\|z_J\|$ is of order $\frac{q}{n}$ and then we get

$$P(\{x; d(x,A) \geqq u\}) \leqq \exp\left(-\frac{u^2}{K}\right).$$

The basic fact there is that we had a greater freedom on the location of the points of J. This is very much what (3.2) expresses.

The basic ingredient in the proof of Theorem 1.3 is as follows, when h_{ij} has the obvious meaning.

Lemma 3.14. *Consider a set $A \subset T$, a number $1 \leqq M \leqq n^{1/4}/(\log n)^2$, a point x of T and a probability measure ν on A such that for each set $I \subset \{1, \ldots, n\}^2$ we have*

$$\int \exp\left(\frac{1}{L} \sum_{(i,j) \in I} h_{ij}(x,y)\right) d\nu(y) \leqq \exp \sqrt{s}M \tag{3.40}$$

where $s = \operatorname{card} I$. Then we have

$$d(x,A) \leqq KL\sqrt{M}n^{1/4} \log n\,. \tag{3.41}$$

Proof. We define $D(y) = d(x,y)$, and we will show that

$$\int D(y) d\nu(y) \leqq KLn^{1/4} \log n \sqrt{M}\,. \tag{3.42}$$

We consider the smallest integer m such that $2^{2m} \geqq n$.

We observe that

$$\begin{aligned} D(y) &= \sup\left\{\sum_{i,j\leqq n} h_{ij}(x,y)\alpha_i\beta_j;\ \sum_{i\leqq n}\alpha_i^2 \leqq 1,\ \sum_{j\leqq n}\beta_j^2 \leqq 1\right\} \\ &\leqq \sup\left\{\sum_{k,\ell\leqq m} 2^{-k}2^{-\ell}\sum_{i\in A_k, j\in B_\ell} h_{ij}(x,y)\right\} \end{aligned}$$

where the supremum is over all families $(A_k)_{k\leqq m}, (B_\ell)_{\ell\leqq m}$ of sets for which $\operatorname{card} A_k \leqq 2^{2k}$, $\operatorname{card} B_\ell \leqq 2^{2\ell}$.

It then follows that

$$\int D(y)d\nu(y) \leqq 2\sum_{1\leqq\ell\leqq k\leqq m} 2^{-k-\ell}E_{k,\ell} \tag{3.43}$$

where

$$E_{k,\ell} = \int \sup\left\{\sum_{i\in A, j\in B} h_{ij}(x,y)\right\} d\nu(y)$$

and where the supremum is taken over all sets A, B with $\operatorname{card} A = 2^{2k}$, $\operatorname{card} B = 2^{2\ell}$.

We now come to the main point.

Lemma 3.15. *Consider a family $\mathscr{T}$ of subsets of $\{1,\ldots,n\}^2$ such that* $\operatorname{card} I \leqq s$ *for $I \in \mathscr{I}$. Then, under* (3.40)

$$\int \sup_{I\in\mathscr{T}}\left\{\sum_{(i,j)\in I} h_{ij}(x,y)\right\} d\nu(y) \leqq KL(\sqrt{s}M + \log\operatorname{card}\mathscr{T}) \tag{3.44}$$

Proof. By (3.40)

$$\nu\left(\sup_{I\in\mathscr{T}}\left\{\sum_{(i,j)\in I} h_{ij}(x,y)\right\} \geqq t\right) \leqq \operatorname{card}\mathscr{T}\exp\left(-\frac{t}{L} + \sqrt{s}M\right)$$

from which (3.44) follows by a standard computation. □

To bound $E_{k,\ell}$ we use (3.44) with $\mathscr{T} = \{A\times B\}$, so that

$$\operatorname{card}\mathscr{T} = \binom{n}{2^{2k}}\binom{n}{2^{2\ell}} \leqq n^{2^{2k}+2^{2\ell}} \leqq n^{2^{2k+1}}$$

and thus

$$E_{k,\ell} \leqq KL(2^{k+\ell}M + 2^{2k}\log n)\,. \tag{3.45}$$

Another bound can be obtained by observing that

$$\sum_{i\in A, j\in B} h_{ij}(x,y) \leqq \sup\left(\sum_{i\leqq n, j\in B} h_{ij}(x,y)\right)$$

and using (3.44) with $\mathscr{T} = \{\{1,\ldots,n\} \times B\}$ where B is any set of cardinal $2^{2\ell}$. Thus $s = n2^{2\ell}$, $\operatorname{card}\mathscr{T} \leqq n^{2^{2\ell}}$, and (3.44) gives

$$E_{k,\ell} \leqq KL(2^{\ell}\sqrt{n}M + 2^{2\ell}\log n)\,. \tag{3.46}$$

Consider now an integer k_0 to be determined later. Using (3.45) we get

$$\begin{aligned}\sum_{\ell \leqq k \leqq k_0} 2^{-k-\ell}E_{k,\ell} &\leqq KL\sum_{k\leqq k_0}\left(\sum_{\ell\leqq k} M + 2^{k-\ell}\log n\right)\\ &\leqq KL\sum_{k\leqq k_0}(k_0M + 2^k\log n)\\ &\leqq KL(k_0^2M + 2^{k_0}\log n)\,.\end{aligned}$$

Using (3.46) we get

$$\begin{aligned}\sum_{k_0\leqq k\leqq m}\sum_{\ell\leqq k} 2^{-k-\ell}E_{k,\ell} &\leqq KL\sum_{k_0\leqq k\leqq m}\left(\sum_{\ell\leqq k}(2^{-k}\sqrt{n}M + 2^{\ell-k}\log n)\right)\\ &\leqq KL\sum_{k_0\leqq k\leqq m}(k2^{-k}\sqrt{n}M + \log n)\\ &\leqq KL(m2^{-k_0}\sqrt{n}M + m\log n)\,.\end{aligned}$$

Thus, going back to (3.43) we have

$$\int D(y)d\nu(y) \leqq KL(2^{k_0}\log n + (\log n)2^{-k_0}\sqrt{n}M + (\log n)^2M)\,.$$

The choice of 2^{k_0} of order $n^{1/4}\sqrt{M}$ gives a bound

$$KL(n^{1/4}\log n\sqrt{M} + (\log n)^2M)$$

and this implies the result.

If we combine Lemma 3.14 with Theorem 3.1, we obtain the following, that contains Theorem 1.3.

Proposition 3.16. *Consider a subset A of $n\times n$ matrices with $0,1$ entries such that $P(A) \geqq 1/2$. Then for $Kn^{1/4}(\log n)^{5/4} \leqq u \leqq n^{1/2}/(\log n)^2$ we have*

$$P(x; d(x,A) \geqq u) \leqq \exp\left(-\frac{u^4}{Kn(\log n)^4}\right).$$

To conclude this section, we should point out a possibly interesting direction of research. Consider the space $T = \{0,1\}^n$, provided with the uniform measure. Consider vectors x_i in a Banach space and the distance d defined on T by $d(\varepsilon,\eta) = \|\sum_{i\leqq n}(\varepsilon_i - \eta_i)x_i\|$. What are the properties of (T,d) with respect to concentration of measure?

4. Controlling missed points

Consider a probability measure ν on Ω^n and x in Ω^n. We want to understand the "location" of y_i when y is distributed like ν. Quite naturally, if $C_i = \{y \in \Omega^n; y_i \neq x_i\}$, we can consider the image of the restriction of ν to C_i by the map $y \to y_i$ and its Radon–Nikodym derivative d_i with respect to μ. It can happen that d_i does not exist. In order to avoid this problem, as well as any other measurability problem, let us assume that Ω is finite, and that each point is measurable of positive measure. Thus, for each function g, we have

$$\int_{C_i} g(y_i)d\nu(y) = \int g d_i d\mu \tag{4.1}$$

or, equivalently

$$\int g(y_i)h_i(x,y)d\nu(y) = \int g d_i d\mu\ . \tag{4.2}$$

In order to measure the size of d_i, it is good to consider the function ψ on $\mathbb{R}^+$ given by

$$\begin{aligned} \psi(x) &= \tau x^2 \quad \text{if } x \leqq 2 \\ \psi(x) &= x\log x \quad \text{if } x \geqq 2 \end{aligned} \tag{4.3}$$

where $\tau = (\log 2)/2$. Thus ψ is convex.

We set

$$m(\nu,x) = \sum_{i \leqq n} \int \psi(d_i)d\mu\ .$$

To get some feelings for this definition, we compare $m(\nu,x)$ with $f(\nu,x)$.

Lemma 4.1. $f(\nu,x) \leqq Km(\nu,x)$

Proof. We first observe that

$$f(\nu,x) = \sum_{i \leqq n} \left(\int d_i d\mu\right)^2 .$$

Consider then $d_i' = d_i 1_{\{d_i \leqq 2\}}$ and $d_i'' = d_i - d_i'$. Thus

$$\begin{aligned} \left(\int d_i d\mu\right)^2 &\leqq 2\left(\int d_i' d\mu\right)^2 + 2\left(\int d_i'' d\mu\right)^2 \\ &\leqq 2\int d_i'^2 d\mu + 2\int d_i'' d\mu \\ &\leqq \left(\frac{2}{\tau} + \frac{2}{\log 2}\right)\int \psi(d_i)d\mu \end{aligned}$$

where we have used Cauchy–Schwarz and the fact that $\int d_i d\mu \leqq 1$. □

We define $m(A,x)$ as the infimum of $m(\nu,x)$ when $\nu(A) = 1$.

The following is a kind of improvement upon (1.2).

Theorem 4.2. *There exists a number L with the following property. For any n and any $A \subset \Omega^n$, we have*

$$\int \exp \frac{1}{L} m(A,x) dP(x) \leqq \frac{1}{P(A)} . \tag{4.4}$$

We first prove the case $n = 1$, that will help to understand the reason for the function ψ. If $x \in A$, it is clear that $m(A,x) = 0$. If $x \notin A$, consider the measure ν on A given by $\nu(B) = \mu(A \cap B)/\mu(A)$. Its Radon Nikodym derivative is $d = 1_A \mu(A)^{-1}$ and

$$\int \psi(d) d\mu = \int_A \psi(\mu(A)^{-1}) .$$

If $\mu(A) \geqq \frac{1}{2}$, this is $\tau\mu(A)^{-1}$; if $\mu(A) \leqq 1/2$, this is $\log(1/\mu(A))$. Thus, if $L \geqq 1$, (4.4) holds when $\mu(A) \leqq 1/2$. When $\mu(A) \geqq 1/2$, we need to have, setting $x = \mu(A)$,

$$x + (1-x)\left(\exp \frac{\tau}{Lx}\right) \leqq \frac{1}{x}$$

and it suffices that

$$\exp \frac{\tau}{Lx} \leqq \frac{1}{x}$$

which holds if $x \geqq 1/2$ for $L \geqq 2\tau/\log 2 = 4$.

The center of the proof of Theorem 4.2 is the following result, that asserts the existence of a certain kernel.

Proposition 4.2. *There exists a number L with the following property. Given a function g from Ω to $[0,1]$, we can find a function k from $\Omega \times \Omega$ to $\mathbb{R}^+$ with the following properties*

$$\forall \omega \in \Omega, \quad \int_\Omega k(\omega', \omega) d\mu(\omega') \leqq 1 \tag{4.5}$$

$$\int_\Omega \exp I(\omega) d\mu(\omega) \leqq \frac{1}{\int_\Omega g d\mu} \tag{4.6}$$

where

$$I(\omega) = \log \frac{1}{g(\omega)} \left(1 - \int k(\omega', \omega) d\mu(\omega')\right) - \int_\Omega \log \frac{1}{g(\omega')} k(\omega', \omega) d\mu(\omega')$$
$$+ \frac{1}{L} \int_\Omega \psi(k(\omega', \omega)) d\omega' .$$

It is part of the statement that $I(\omega)$ is well defined, i.e. $\int_\Omega k(\omega', \omega) = 1$ if $g(\omega) = 0$, and $k(\omega', \omega) = 0$ if $g(\omega') = 0$.

It should be obvious that to prove Proposition 4.2, it suffices to prove the following.

Proposition 4.3. *There exists a number L with the following property. Consider a nonincreasing function g from $[0,1]$ to $[0,1]$. Then we can find a function k from $[0,1]^2$ to $\mathbb{R}^+$ with the following properties*

$$\forall t \in [0,1], \quad \int_0^t k(s,t)ds \leqq 1 \tag{4.7}$$

$$\int_0^1 \exp I(t) d\mu \leqq \frac{1}{\int_0^1 g(t)dt} \tag{4.8}$$

where

$$I(t) = \log \frac{1}{g(t)} \left(1 - \int_0^t k(s,t)ds\right) + \int_0^t \log \frac{1}{g(s)} k(s,t)ds + \frac{1}{L} \int_0^t \psi(k(s,t))ds .$$

The most interesting part of the proof is contained in the following lemma.

Lemma 4.4. *If L is large enough, the following occurs. Consider a number b; $0 < b < 1/2$, and a nonincreasing function g from $[0,1]$ to $[b/2, 2b]$. Then*

$$\int_0^1 \frac{1}{g(t)} \exp\left(-\frac{L}{8}\int_0^t \left(\log \frac{1}{g(s)} - \log \frac{1}{g(t)}\right)^2 ds\right) dt \leqq \frac{1}{\int_0^1 g(t)dt} . \tag{4.9}$$

Proof. We set $a = \int_0^1 g(t)dt$. We consider

$$t_0 = \sup \left\{ t : \frac{L}{8} \int_0^t \left(\log \frac{g(t)}{g(s)}\right)^2 ds \leqq 2 \right\} .$$

For $t > t_0$, the integrand in (4.9) is at most

$$\frac{2}{b} e^{-2} \leqq \frac{1}{2b} \leqq \frac{1}{a} .$$

Consider

$$c = \frac{1}{t_0} \int_0^{t_0} g(t)dt \geqq a .$$

It suffices to show that

$$\int_0^{t_0} \frac{1}{g(t)} \exp\left(-\frac{L}{8}\int_0^t \left(\log \frac{g(t)}{g(s)}\right)^2 ds\right) dt \leqq \frac{t_0}{c} .$$

For $x \leqq 2$, we have $\exp -x \leqq 1 - x/4$. Since $g(t)^{-1} \geqq 1/2b \geqq 1/4c$, it suffices to show that

$$\int_0^{t_0} \left(\frac{1}{g(t)} - \frac{L}{128c} \int_0^t \log \left(\frac{g(t)}{g(s)}\right)^2 ds \right) dt \leqq \frac{t_0}{c} . \tag{4.10}$$

We write $g(t) = c(1 + u(t))$, so that $\int_0^{t_0} u(t)dt = 0$, and thus (4.10) reduces to

$$\int_0^{t_0} \left(\frac{1}{1+u(t)} - \frac{L}{128} \int_0^t \left(\log \frac{1+u(t)}{1+u(s)} \right)^2 ds \right) dt \leqq t_0 . \tag{4.11}$$

Now since the derivative of $\log x$ is at least $1/2$ when $x \in [1/2, 2]$, we have

$$\left(\log \frac{1+u(t)}{1+u(s)} \right)^2 \geqq \frac{1}{4}(u(s) - u(t))^2 .$$

Moreover,

$$\frac{1}{1+u} = 1 - u + \frac{u^2}{1+u} \leqq 1 - u + u^2$$

so that it suffices to prove that

$$\int_0^{t_0} u^2(s)ds - \frac{L}{512} \int_0^{t_0} \int_0^t (u(s) - u(t))^2 ds dt \leqq 0 .$$

But, since $\int_0^{t_0} u(s)ds = 0$, we have

$$\begin{aligned} \int_0^{t_0} \int_0^t (u(s) - u(t))^2 ds dt &= \frac{1}{2} \int \int_{0 \leqq s,t \leqq t_0} (u(s) - u(t))^2 ds dt \\ &\geqq \frac{1}{2} \int_0^{t_0} u(t)^2 dt . \end{aligned}$$

□

The connection with (4.8) is as follows. Writing

$$\begin{aligned} I(t) &= \log \frac{1}{g(t)} - \int_0^t \left(\log \frac{1}{g(t)} - \log \frac{1}{g(s)} \right) k(s,t) ds \\ &\quad + \frac{1}{L} \int_0^t \psi(k(s,t)) ds \end{aligned}$$

and recalling that $\psi(u) = \tau u^2$ for $u \leqq 2$, we see that a good choice for k is $k(s,t) = \frac{L}{2r} \log \frac{g(s)}{g(t)}$; But this is appropriate only when g varies in a narrow range as in Lemma 4.4. The other cases require different arguments.

We start the proof of Proposition 4.3. We set $a = \int_0^1 g(t)dt$. We set

$$t_1 = \sup\{t; g(t) \geqq 2a\}$$

(and $t_1 = 0$ if the set above is empty). We set

$$a_1 = \int_0^{t_1} g(s)ds .$$

Lemma 4.5. *Proposition* 4.3 *holds when* $L \geqq 2$, $a_1 \geqq a 2^{-L/2}$.

Proof. We set $k(s,t) = 0$ whenever $t \leqq t_1$ or $s \geqq t_1$. If $s < t_1 < t$, we set $k(s,t) = g(s)/a_1$. We observe that $k(s,t) \geqq 2a/a_1 \geqq 2$. Thus, for $t \leqq t_1$, we have

$$I(t) = \log \frac{1}{g(t)} \leqq \log \frac{1}{a} \,. \tag{4.12}$$

For $t > t_1$, since $t_1 \leqq 1/2$ and since $\psi(x) = x \log x$ for $x \geqq 2$, we have

$$\begin{aligned} I(t) &= \frac{1}{a_1} \int_0^{t_1} g(s) \log \frac{1}{g(s)} ds + \frac{1}{L} \int_0^{t_1} \psi\left(\frac{g(s)}{a_1}\right) ds \\ &= \frac{1}{a_1} \int_0^{t_1} g(s) \log \left(\frac{1}{g(s)} \left(\frac{g(s)}{a_1}\right)^{1/L} \right) ds \,. \end{aligned}$$

Now, since $g(s) \geqq 2a$ for $s \leqq t_1$, since $L \geqq 2$, and since we assume $a/a_1 \leqq 2^{L/2}$, we have

$$\begin{aligned} \frac{1}{g(s)} \left(\frac{g(s)}{a_1}\right)^{1/L} &= \frac{1}{g(s)^{1-1/L}} \frac{1}{a_1^{1/L}} \leqq \frac{1}{(2a)^{1-1/L}} \frac{1}{a_1^{1/L}} \\ &= \frac{1}{a} \left(\frac{a}{a_1}\right)^{1/L} \frac{1}{2^{1-1/L}} \leqq \frac{1}{a} \frac{\sqrt{2}}{2^{1-1/L}} \leqq \frac{1}{a} \,. \end{aligned}$$

Thereby $I(t) \leqq \log 1/a$ for each value of t. The result follows. □

In words, Lemma 4.5 takes care of the case where g has a lot of weight on the set $\{g \geqq 2a\}$; On the other hand Lemma 4.4 assumes there is no such weight. Of course we will need to interpolate between both arguments; but we also need to handle the values of t where $g(t) \leqq a/2$. We now assume

$$a_1 \leqq a 2^{-L/2} \tag{4.13}$$

since otherwise the proof is finished. We define t_2 by

$$t_2 = \sup \left\{ t \leqq 1 ; \frac{L}{2} \int_{t_1}^{t} \log \frac{g(s)}{g(t)} ds \leqq 1 \right\} . \tag{4.14}$$

We define d by

$$\frac{L}{2} \int_{t_1}^{t_2} \log \frac{g(s)}{d} ds = 1 \,. \tag{4.15}$$

We first settle a technical point.

Lemma 4.6. *If $L \geqq 16$, we have $d \geqq a/2$.*

Proof. First, we observe that $g(t) \leqq d$ if $t > t_2$; this is rather obvious comparing (4.14) and (4.15), and since g is non increasing. Thus

$$\int_{t_1}^{t_2} g(t) dt = a - \int_0^{t_1} g(t) dt - \int_{t_2}^{1} g(t) dt \geqq a - a_1 - (1 - t_2) d \,. \tag{4.16}$$

Next, since $d < g(t) < 2a$ for $t_1 < t < t_2$, we can find $u, v > 0$ with $u + v = 1$ and

$$2ua + vd = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} g(t)dt \ .$$

The convexity of the function $x \to \log(1/x)$ shows that

$$\frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \log \frac{1}{g(t)} dt \leqq u \log \frac{1}{2a} + v \log \frac{1}{d} \ . \tag{4.17}$$

Now, since $v = 1 - u$, and using (4.17)

$$\begin{aligned} \frac{2}{(t_2 - t_1)L} &= \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \log \frac{g(t)}{d} dt \\ &= \log \frac{1}{d} - \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \log \frac{1}{g(t)} dt \\ &\geqq u \left(\log \frac{1}{d} - \log \frac{1}{2a} \right) = u \log \frac{2a}{d} \ . \end{aligned} \tag{4.18}$$

Also, using (4.16)

$$\begin{aligned} u(2a - d) &= 2ua + vd - d \\ &= -d + \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} g(t)dt \geqq \frac{1}{t_2 - t_1} (a - a_1 - (1 - t_1)d) \ . \end{aligned}$$

Let us assume for contradiction that $d \leqq a/2$. Then, since we assume $a_1 \leqq a/4$, we have

$$u \geqq \frac{1}{t_2 - t_1} \frac{1}{\frac{3a}{2}} \cdot \frac{a}{4} = \frac{1}{6(t_2 - t_1)} \ .$$

On the other hand, (4.18) implies

$$u \log 4 \leqq \frac{2}{L(t_2 - t_1)} \ .$$

This is impossible if L is large enough. □

We now define $\bar{g}(t) = g(t)$ if $t \leqq t_2$ and $\bar{g}(t) = d$ if $t \geqq t_2$. Thus $\bar{g} \geqq g$, and $\bar{g} \geqq a/2$. We will show that in fact k can be constructed so that $\int_0^1 \exp I(t)dt \leqq 1/\bar{a}$, where $\bar{a} = \int_0^1 \bar{g}(t)dt$.

We consider now a parameter $\xi, 0 < \xi < 1$, which role is to decide the "weight" of the argument of Lemma 4.5 in the final agrument. We define

$$k(s,t) = \frac{\xi g(s)}{a_1} \quad \text{if } s \leqq t_1 \leqq t$$

$$k(s,t) = \frac{L(1 - \xi)}{2} \log \frac{\bar{g}(s)}{\bar{g}(t)} \quad \text{if } t_1 < s \leqq t$$

and $k(s,t) = 0$ in the other cases.

We observe that $\int_{t_1}^{t} k(s,t)ds$ is an increasing function of t, and that the definition of t_2, of d, and of $\bar{g}$ ensures that (4.7) holds.

When $t \leqq t_1$, we have

$$I(t) = \log \frac{1}{g(t)} \ . \tag{4.19}$$

When $t \geqq t_1$, we have $I(t) = \xi I_1 + (1-\xi) I_2(t)$, where

$$\begin{aligned} I_1 &= \int_0^{t_1} \left(\frac{g(s)}{a_1} \log \frac{1}{g(s)} + \frac{1}{L\xi} \psi\left(\frac{\xi g(s)}{a_1}\right)\right) ds \\ I_2(t) &= \log \frac{1}{g(t)} - \frac{L}{2} \int_{t_1}^{t} \log \frac{g(s)}{g(t)} \log \frac{\bar{g}(s)}{\bar{g}(t)} dt \\ &\quad + \frac{1}{L(1-\xi)} \int_{t_1}^{t} \psi\left(\frac{L}{2}(1-\xi) \log \frac{\bar{g}(s)}{\bar{g}(t)}\right) dt \ . \end{aligned}$$

Thus,

$$\begin{aligned} \int_0^1 \exp I(t) dt &\leqq \int_0^{t_1} \frac{1}{g(s)} ds + \int_{t_1}^{1} \exp\left(\xi I_1(t) + (1-\xi) I_2(t)\right) dt \\ &\leqq \frac{t_1}{2a} + \left(\int_{t_1}^{1} \exp I_1(t) dt\right)^{\xi} \left(\int_{t_1}^{1} \exp I_2(t) dt\right)^{1-\xi} \ . \end{aligned} \tag{4.20}$$

Lemma 4.7. *We have*

$$I_1 \leqq J + \log \frac{1}{a} \quad \textit{where } J = \log \left(\frac{1}{2}\right)^{1-1/L} \left(\frac{\xi a}{2a_1}\right)^{1/L} \ . \tag{4.21}$$

When $t \geqq t_1$ we have

$$I_2(t) \leqq \log \frac{1}{\bar{g}(t)} - \frac{L}{4} \int_{t_1}^{t} \left(\log \frac{\bar{g}(t)}{\bar{g}(s)}\right)^2 ds \ . \tag{4.22}$$

Proof. To prove (4.21), we note that if we assume $\xi a \geqq a_1$, we have

$$\xi^{-1} \psi\left(\frac{\xi g(s)}{a_1}\right) \leqq \frac{g(s)}{a_1} \log \frac{\xi g(s)}{a_1}$$

since $g(s) \geqq 2a_1$ for $s \leqq t_1$. Thus

$$I_1(t) \leqq \int_0^{t_1} \frac{g(s)}{a_1} \log \left(\frac{1}{g(s)}\right)^{1-1/L} \left(\frac{\xi}{a_1}\right)^{1/L} ds \ .$$

Now

$$\log \left(\frac{1}{g(s)}\right)^{1-1/L} \left(\frac{\xi}{a_1}\right)^{1/L} \leqq \log \frac{1}{a} + \log \left(\frac{1}{2}\right)^{1-1/L} \left(\frac{\xi a}{2a_1}\right)^{1/L} \ .$$

To prove (4.22), we observe that

$$\log \frac{g(s)}{g(t)} \log \frac{\bar{g}(s)}{\bar{g}(t)} \geqq \left(\log \frac{\bar{g}(s)}{\bar{g}(t)}\right)^2 \ .$$

Indeed, if $s > t_2$, both sides are zero, while, if $s < t_2$, $g(s) = \bar{g}(s)$ and $g(t) \leqq \bar{g}(t)$. To conclude, we then use that $\psi(x) \leqq x^2$. □

We now observe that

$$\int_{t_1}^{1} \exp\left(\log\frac{1}{\bar{g}(t)} - \frac{L}{4}\int_{t_1}^{t}\left(\log\frac{\bar{g}(t)}{\bar{g}(s)}\right)^2 ds\right) dt \leqq \frac{(1-t_1)^2}{\int_{t_1}^{1}\bar{g}(t)dt} \ . \tag{4.23}$$

To see this, we first observe that $a/2 \leqq \bar{g}(t) \leqq 2a$ on $[t_1, 1]$. We then use the transformation $t \to (1-t_1)^{-1}(t - t_1)$ from $[t_1, 1]$ to $[0, 1]$ to reduce to Lemma 4.4, since $t_1 \leqq 1/2$.

We note that

$$\int_{t_1}^{1}\bar{g}(t)dt = \int_{0}^{1}\bar{g}(t)dt - a_1 = \bar{a} - a_1 = \bar{a}\left(1 - \frac{a_1}{\bar{a}}\right) \ .$$

Since $a_1 \leqq a/4 \leqq \bar{a}/4$, we have

$$\left(\int_{t_1}^{1}\bar{g}(t)dt\right)^{-1} \leqq \frac{1}{\bar{a}}\left(1 + \frac{2a_1}{\bar{a}}\right) \ . \tag{4.24}$$

Combining (4.20) to (4.24), we have shown that

$$\int_0^1 \exp I(t)dt \leqq \frac{t_1}{2a} + (1-t_1)\left(\frac{1}{a}\right)^{\xi}\exp\xi J\left(\frac{1}{\bar{a}}\left(1+\frac{2a_1}{\bar{a}}\right)\right)^{1-\xi}$$

as soon as $\xi \leqq 1/2$. Thus it suffices that

$$(\exp\xi J)\left(\frac{1}{\bar{a}}\left(1+\frac{2a_1}{\bar{a}}\right)\right)^{1-\xi} \leqq \left(\frac{1}{a}\right)^{1-\xi}$$

or even

$$\left(1+\frac{2a_1}{\bar{a}}\right)^{1-\xi}\exp\xi J \leqq 1 \ .$$

Since

$$\left(1+\frac{2a_1}{\bar{a}}\right)^{1-\xi} \leqq \exp\frac{2a_1}{\bar{a}}$$

it suffices that

$$\frac{2a_1}{\bar{a}} + \xi J \leqq 0 \ .$$

If we look for ξ of the type $\eta a_1/a$, it suffices that

$$2 + \eta\log\left(\left(\frac{1}{2}\right)\eta^{1/L}\right) < 0$$

This holds if $\eta = 4$, and L is large enough (and $\xi \leqq 1/2$ if $L \geqq 32$). □

We now turn to the proof of Theorem 1.4, by induction upon n. Assuming that the result holds for n, we will prove it for $n+1$. We denote by P the product measure on Ω^n, by P' the product measure on Ω^{n+1}. We identify Ω^{n+1} with $\Omega^n \times \Omega$ and P' with $P' \otimes \mu$. For ω in Ω we set

$$A_\omega = \{x \in \Omega^n;\ (x, \omega) \in A\}$$

and $g(\omega) = P(A_\omega)$. We appeal to Proposition 4.2. The number L there will be the number L of Theorem 4.1. We consider a kernel $k(\omega', \omega)$ that satisfies (4.5) and (4.6).

Consider for each $\omega' \in \Omega$ and each $x \in \Omega^n$ a measure $\nu_{x,\omega'}$ on Ω^n such that $\nu_{x,\omega'}(A_{\omega'}) = 1$.

Consider the measure $\eta_{x,\omega}$ given by

$$\eta_{x,\omega} = \left(1 - \int_\Omega k(\omega', \omega)\, d\mu(\omega')\right)\nu_{x,\omega} \otimes \delta_\omega + \int k(\omega', \omega)\nu_{x,\omega'} \otimes \delta_{\omega'}\, d\mu(\omega')\ .$$

Thus $\eta_{x,\omega}$ is a probability measure supported by A. The convexity of ψ should make it obvious that

$$\begin{aligned} m(\eta_{x,\omega}, (x, \omega)) \leqq &\int \psi(k(\omega', \omega))d\mu(\omega') + (1 - \int k(\omega', \omega)d\mu(\omega'))m(\nu_{x,\omega}, x) \\ &+ \int k(\omega', \omega)m(\nu_{x,\omega'}, x)d\mu(\omega')\ . \end{aligned}$$

Thus, taking infimum over all possible choices of $\nu_{x,\omega'}$, we get

$$\begin{aligned} m(A, (x, \omega)) \leqq &\int \psi(k(\omega', \omega))d\mu(\omega') + \left(1 - \int k(\omega', \omega)d\mu(\omega')\right) m(A_{\omega'}, x) \\ &+ \int k(\omega', \omega)m(A_{\omega'}, x)d\mu(\omega')\ . \end{aligned}$$

Now, using Hölder's inequality, we get

$$\int \exp \frac{1}{L} m(A, (x, \omega))dP(x) \leqq U \exp \frac{1}{L} \int \psi(k(\omega', \omega))d\mu(\omega')$$

where

$$\begin{aligned} U = &\left(\int \exp \frac{1}{L} m(A_\omega, x)dP(x)\right)^{1 - \int k(\omega', \omega)d\mu(\omega')} \\ &\times \prod_{\omega'} \left(\int \exp \frac{1}{L} m(A_{\omega'}, x)dP(x)\right)^{\mu(\{\omega'\})k(\omega', \omega)}\ . \end{aligned}$$

Using the induction hypothesis,

$$U \leqq \left(\frac{1}{g(\omega)}\right)^{1 - \int k(\omega', \omega)d\mu(\omega')} \prod_{\omega'} \left(\frac{1}{g(\omega')}\right)^{\mu(\{\omega'\})h(\omega', \omega)}$$

so that

$$\int \exp \frac{1}{L} m(A, (x, \omega))dP(x) \leqq \exp I(\omega)\ .$$

Integrating in ω, using Fubini theorem, and noting that $\int g(\omega)d\mu(\omega) = P'(A)$ finishes the proof. $\square$

We now turn to the proof of Theorem 1.4. First, an obvious approximation argument allows to reduce to the case where $\mathscr{F}$ is finite. An equally obvious argument allows then to reduce to the case where Ω is finite (so that we can use Theorem 4.1).

On Ω^n we consider the function

$$Z(x) = \sup_{f \in \mathscr{F}} \sum_{i \leqq n} f(x_i) \ .$$

We consider a number a and the set $A = \{y; Z(y) \leqq a\}$.

Consider $f \in \mathscr{F}$ such that $Z(x) = \sum_{i \leqq n} f(x_i)$. Then

$$\begin{aligned} Z(x) - a &\leqq \int \sum_{i \leqq n} (f(x_i) - f(y))d\nu(y) \\ &= \sum_{i \leqq n} \int_\Omega (f(x_i) - f(y_i))\, d_i\,(y_i)d\mu(y_i) \end{aligned}$$

where d_i is defined as in (4.1).

Lemma 4.8. *If* u, $v \geqq 0$, $u \leqq 1$, *we have*

$$uv \leqq u^2 + 2\psi(v) \ . \tag{4.25}$$

Proof. We have $uv - u^2 \leqq v^2/4$, so this is true if $v \leqq 2$, since then $\psi(v) = \tau v^2$, and $\tau \geqq 1/8$. But if $v \geqq 2$, $2\psi(v) \geqq 2v \log 2 \geqq v$. $\square$

Consider a number $\delta \geqq 2U$ that will be chosen later. Then, using (4.25) for $u = \delta^{-1}|f(x_i) - f(\omega)|$, we get

$$\begin{aligned} &\int_\Omega (f(x_i) - f(\omega))\, d_i(\omega)d\mu(\omega) \\ &\leqq \delta \int \frac{|f(x_i) - f(\omega)|}{\delta}\, d_i(\omega)d\mu(\omega) \\ &\leqq \delta^{-1} \int_\Omega (f(x_i) - f(\omega))^2\, d\mu(\omega) + 2\delta \int \psi(d_i)d\mu \ . \end{aligned}$$

We observe that

$$n \int f^2(\omega)d\mu(\omega) \leqq V$$

so that

$$Z(x) \leqq a + 2\delta^{-1}\left(V + \sum_{i \leqq n} f^2(x_i)\right) + 2\delta m(A,x)$$

and thus, for each numbers $u, w > 0$,

$$\begin{aligned} &P(Z(u) \geqq a + 2\delta^{-1}(V + w) + 2\delta u) \\ &\leqq \frac{1}{P(A)} \exp\left(-\frac{u}{K}\right) + P\left(\sup_{f \in \mathscr{F}} \sum_{i \leqq n} f(x_i)^2 \geqq w\right) \end{aligned} \tag{4.26}$$

where we have used (4.4).

Lemma 4.9. *We have*

$$P\left(\sup_{f\in\mathscr{F}}\sum_{i\leqq n} f(x_i)^2 \geqq 4V + kU^2\right) \leqq 4\cdot 2^{-k} . \tag{4.27}$$

Proof. This is a straightforward use of the 2 points inequality, as in [T3] Sect. 13.

In (4.26), we take $w = 4V + kU^2$, and using (4.27) we get, taking k of order u

$$P(Z \geqq a + 2\delta^{-1}(5V + uU^2) + 2\delta u) \leqq \frac{5}{P(A)}\exp\left(-\frac{u}{K}\right) .$$

We now take $\delta = \min(2U, \sqrt{V/U})$ to get

$$P(Z \geqq a + K\max(uU, \sqrt{uV})) \leqq \frac{5}{P(A)}\exp\left(-\frac{u}{K}\right)$$

so that, for $v > 0$

$$P(Z \geqq a + v) \leqq \frac{5}{P(A)}\exp\left(-\frac{1}{K}\min\left(\frac{v^2}{V}, \frac{v}{U}\right)\right)$$

Taking $a = M - v$, we get

$$\frac{1}{2} \leqq \frac{5}{P(Z \leqq M - v)}\exp\left(-\frac{1}{K}\min\left(\frac{v^2}{V}, \frac{v}{U}\right)\right) .$$

Combining both estimates, we get

$$P(|Z - M| \geqq v) \leqq K\exp\left(-\frac{1}{K}\min\left(\frac{v^2}{V}, \frac{v}{U}\right)\right) . \tag{4.28}$$

A standard computation shows that this implies

$$E(|Z - M|) \leqq K(U + \sqrt{V})$$

which plugged into (4.28) gives, after elementary estimates

$$P(|Z - EZ| \geqq v) \leqq K\exp\left(-\frac{1}{K}\min\left(\frac{v^2}{V}, \frac{v}{U}\right)\right) . \tag{4.29}$$

This is not quite what we need, although this does imply (1.13) for $t \leqq 2V/U$. But for $v \geqq V/U$, we want an exponent behaving like $\frac{v}{V}\log\frac{vU}{V}$ rather than v/V. This will be obtained by a truncation argument. We consider a parameter $\rho < U$, that will be determined later. For $x \in \Omega^n$, we set

$$W(x) = \sup_{f\in\mathscr{F}}\sum_{i\leqq n} |f(x_i)|_{\{|f(x_i)|\geqq\rho\}} .$$

Thus

$$W(x) \leqq \frac{1}{\rho}\sup_{f\in\mathscr{F}}\sum_{i\leqq n} f(x_i)^2$$

and thus $EW \leqq V/\rho$.

We now use the q-points inequality (1.16), as in [T3], Sect. 13 to get, for $t \geqq KEW$

$$P(W \geqq t) \leqq K\exp\left(-\frac{t}{KU}\log\frac{t}{KEW}\right) \tag{4.30}$$
$$\leqq K\exp\left(-\frac{t}{KU}\log\frac{t\rho}{KV}\right) .$$

This indicates that it will be a good idea to take $\rho = \sqrt{UV/t}$, so that we keep the logarithmic term above. We will then consider the class

$$\mathscr{F}_\rho = \{f 1_{\{|f| \leqq \rho\}}\}$$

to which we will apply (4.30). Defining Z_ρ in the obvious manner, we have

$$|Z - Z_\rho| \leqq W .$$

It follows in particular that

$$|EZ - EZ_\rho| \leqq EW \leqq \frac{V}{\rho} = \sqrt{\frac{tV}{U}} \tag{4.31}$$

Also,

$$P(Z \geqq EZ + t) \leqq P\left(W \geqq \frac{t}{2}\right) + P\left(Z_\rho \geqq EZ + \frac{t}{2}\right) .$$

We use (4.31) for the first term; for the second term,

$$P\left(Z_\rho \geqq EZ + \frac{t}{2}\right) \leqq P\left(Z_\rho \geqq EZ_\rho + \frac{t}{4}\right)$$

provided $t \geqq 4\sqrt{tV/U}$ i.e. $t \geqq 16V/U$. To bound this last term, we use (4.30), which implies

$$P\left(Z_\rho \geqq EZ_\rho + \frac{t}{4}\right) \leqq K\exp\left(-\frac{1}{K}\min\left(\frac{t^2}{V}, \frac{t}{\rho}\right)\right)$$
$$\leqq K\exp\left(-\frac{1}{K}\min\left(\frac{t^2}{V}, \frac{t^{3/2}}{\sqrt{UV}}\right)\right)$$
$$= K\exp\left(-\frac{1}{K}\frac{t^{3/2}}{\sqrt{UV}}\right) .$$

The result follows.

5. *q*-points inequalities

Consider an integer $q \geqq 2$. For points $y_1, \ldots, y_q, x$ of Ω^n we set

$$U(y_1, \ldots, y_q, x) = \sum_{i \leqq n} \prod_{\ell \leqq q} h_i(x, y_\ell) .$$

Given positive measures $\nu_1, \ldots, \nu_q$ on Ω^n, we consider

$$G(\nu_1, \ldots, \nu_q, x) = \int (a(q)+1)^{U(y_1,\ldots,y_q,x)}\, d\nu_1(y_1)\cdots d\nu_q(y_q) \tag{5.1}$$

where the number $a(q)$ is defined as follows:

$$a(q) = \frac{(1+qt_q)^q}{(1+q)^{q-1}}$$

where t_q is the largest root of the equation

$$(1-t)(1+tq)^q = (1+t(q+1))^{q-1} \ .$$

This seemingly strange definition is motivated by Lemma 5.3 below.

Given a set A in Ω^q, we define $G_q(A,x)$ as the infimum of $G(\nu, \ldots, \nu, x)$ over all choices of atomic probabilities ν such that $\nu(A) = 1$.

Theorem 5.1. *Given a set $A \subset \Omega^n$, we have*

$$\int G_q(A,x)dP(x) \leqq \frac{1}{P(A)^q} \ . \tag{5.2}$$

In this statement, one could also investigate what happens when we allow a right-hand side $P(A)^\alpha (\alpha > 0)$. This, however, would make the calculus part of the proof rather unpalatable and did not bring much improvement in the case of (1.16). On the other hand, the calculus is easier if $q = 2$, and is taken care of by Lemma 2.3. In this case one can recover Theorem 2.1 through (2.23) by the proof we will present. Thus, in one sense, Theorem 2.1 is special case of Theorem 5.1. On the other hand, the extra parameter q in Theorem 5.1 makes things look more complicated than they really are. Thus we have felt that presenting a direct proof of Theorem 2.1 was worth a small amount of repetition.

To better understand the quantity $a(q)$, we note the following:

Lemma 5.2. a) $a(2) = \frac{1}{3}$
b) $\lim_{q\to\infty} \frac{a(q)}{q} \geqq e^{-e}$

Proof. a) is elementary, since $t_2 = 0$. To see b, we note that

$$\left(\frac{1+t(q+1)}{1+tq}\right)^{q-1} = \left(1 + \frac{t}{1+tq}\right)^{q-1} \leqq \exp(q-1)\frac{t}{1+tq} \leqq e$$

Thus $t_q \geqq r_q$, where

$$(1-r_q)(1+qr_q) = e$$

Clearly, $r_q \to 1$, so that $(1-r_q)q \to e$. Now

$$a(q) \geqq \frac{(1+qr_q)^q}{(1+q)^{q-1}} = (1+qr_q)\left(1+\frac{(r_q-1)q}{q+1}\right)^{q-1}$$

and this implies the result. □

To relate Theorem 5.1 with (1.13), we note that

$$G_q(A, x) \geqq (a(q) + 1)^{H_q(A,x)} .$$

In view of Lemma 5.2, (5.1) is a kind of improvement over (1.13).

At the root of Theorem 5.1 is the following Calculus lemma.

Lemma 5.3. *If* $0 \leqq x \leqq 1$, *we have*

$$\inf_{0 \leqq \lambda \leqq 1} \left[a(q)\lambda^q \left(1 + (1 - \lambda) \left(\frac{1}{x} - 1 \right) \right)^q \right] \leqq q + 1 - qx .$$

Proof. Setting $x = 1 - t$, we look for the best value of a for which, for all $0 \leqq t < 1$,

$$\inf_{0 \leqq \mu \leqq 1} \left[a(1 - \mu)^q + \left(1 + \mu \frac{t}{1 - t} \right)^q \right] \leqq 1 + qt . \tag{5.3}$$

Setting $b = t/(1 - t)$, $s = (b/a)^{1/q-1}$ the infimum is obtained for $\mu = (1 - s)/(1 + bs)$ if $s \leqq 1$ and $\mu = 0$ if $s \geqq 1$. The case $s \geqq 1$ occurs exactly if $t/(1 - t) \geqq a$ i.e. $t \geqq a/(1 + a)$. To check (5.3) for $t \geqq a/(1 + a)$, it suffices to check it for $t = a/(1 + a)$, in which case $\mu = (1 - s)/(1 + bs)$. Thereby it suffices to show that the value of

$$a(1 - \mu)^q + \left(1 + \mu \frac{t}{1 - t} \right)^q$$

for $\mu = (1 - s)/(1 + bs)$ is a most $1 + qt$. Algebra shows that this value is $(b + 1)^q/(bs + 1)^{q_1}$. Further algebra shows that it suffices that

$$\forall 0 \leqq t \leqq 1, \quad 1 \leqq g(t) = (1 + tq)((1 - t)^{\frac{q}{q-1}} + ct^{\frac{q}{q-1}})^{q-1}$$

where $c = a^{-1/(q-1)}$. Algebra shows that the derivative of g has the sign of

$$k(t) = c(1 + t(q + 1)) - t^{\frac{q-2}{q-1}}(1 - t)^{\frac{1}{q-1}}(q + 1) .$$

Thus, it is positive, then negative, then positive, and the minimum of g is obtained at the largest root of the equation $k(t) = 0$. The value of $a(q)$ is chosen so that this minimum is 1. $\square$

We now prove Theorem 5.1. There will be two inductions over n in the proof. We make the convention that we denote by a prime all the quantities computed relatively to Ω^{n+1}, and without a prime those computed relatively to Ω^n.

First, for each non-empty subset A of Ω^{n+1} we construct an atomic probability measure $\nu_{A,x}$ supported by A, as follows. If $n = 0$, we set $\nu_{A,x} = \delta_x$ if $x \in A$; If $x \notin A$, we set $\nu_{A,x} = \delta_{x_0}$ where x_0 is any point of A. If $n \geqq 1$, we consider

$$B = \{z \in \Omega^n \ \exists \omega \in \Omega, (z, \omega) \in A\} \tag{5.4}$$

$$A_\omega = \{z \in \Omega^n; \ (z, \omega) \in A\} . \tag{5.5}$$

Thus, if $(x, \omega) \in \Omega^{n+1}$, we know by induction how to construct $\eta_0 = v_{B,x}$ and $\eta_1 = v_{A_\omega, x}$. We consider two atomic probabilities γ_0, γ_1 on A such that the projection of γ_0 on Ω^n is η_0, and the probability γ_1 is supported by $\Omega^n \times \{\omega\}$, and its projection on Ω^n is η_1.

We then consider λ that minimizes

$$a(q)\lambda^q + \left(1 + (1-\lambda)\left(\frac{1}{g(\omega)} - 1\right)\right)^q ,$$

where $g(\omega) = P(A_\omega)/P(B)$.

Thus, by Lemma 5.3 we have

$$a(q)\lambda^q + \left(1 - (1-\lambda)\left(\frac{1}{g(\omega)} - 1\right)\right)^q \leqq q + 1 - qg(\omega) . \tag{5.6}$$

We set $v_{A,(x,\omega)} = \lambda\gamma_0 + (1-\lambda)\gamma_1$. Next we show by induction over n that, given sets $A_1, \dots, A_q \subset \Omega^{n+1}$, we have

$$\int G'(x,\omega) dP'(x,\omega) \leqq \frac{1}{\prod_{\ell \leqq q} P'(A_\ell)}$$

where

$$G'(x,\omega) = G'(v_1, \dots, v_q, (x,\omega)) ,$$

$v_\ell = v_{A_\ell,(x,\omega)}$. Specializing this statement to the case when $A_1 = \cdots = A_q$ proves (5.2). For $\ell \leqq q$, we define $A_{\omega,\ell}, B_\ell$ as in (5.4), (5.5); more generally, we keep the notations of the construction of $v_{A,x}$, giving the obvious meaning to the subscript ℓ.

By construction for each ℓ we can write $v_\ell = v_{0,\ell} + v_{1,\ell}$ (where $v_{0,\ell} = \lambda_\ell\gamma_{0,\ell}, v_{1,\ell} = (1-\lambda_\ell)\gamma_{1,\ell}$). To bound $G'(x,\omega)$ we expend the product $\bigotimes_{\ell \leqq q} v_\ell$ as

$$\sum_\varepsilon \bigotimes_{\ell \leqq q} v_{\varepsilon_i,\ell}$$

where the summation is over all sequences ε of zeroes and ones. Thus

$$G'(x,\omega) = \sum_\varepsilon G'(v_{\varepsilon_1,1}, \dots, v_{\varepsilon_q,q}, (x,\omega)) .$$

The term $\prod_{\ell \leqq q} h_{n+1}((x,\omega), y_\ell)$ plays no part in these terms, except in the case $\varepsilon_1 = \cdots = \varepsilon_q = 0$, where it contributes a factor $1 + a(q)$. Thus

$$G'(x,\omega) = a(q)G(\theta_{0,1}, \dots, \theta_{0,q}, x) + \sum_\varepsilon G(\theta_{\varepsilon_1,1}, \dots, \theta_{\varepsilon_1,q}, x)$$

where $\theta_{0,\ell} = \lambda_\ell\eta_{0,\ell}, \theta_{1,\ell} = (1-\lambda_\ell)\eta_{1,\ell}$. We now integrate over x, and use the induction hypothesis to get

$$\int G'(x,\omega) dP(x) \leqq a(q) \prod_{\ell \leqq q} \frac{\lambda_\ell}{P(B_\ell)} + \prod_{\ell \leqq q} \left(\frac{\lambda_\ell}{P(B_\ell)} + \frac{1-\lambda_\ell}{P(A_{\ell,\omega})}\right) .$$

Setting $g_\ell(\omega) = P(A_{\ell,\omega})/P(B_\ell)$, we get

$$\begin{aligned}
&\int G'(x,\omega)dP(x)\\
&\leqq \frac{1}{\prod_{\ell\leqq q} P(B_\ell)}\left(a(q)\prod_{\ell\leqq q}\lambda_\ell + \prod_{\ell\leqq q}\left(1+(1-\lambda_\ell)\left(\frac{1}{g_\ell(\omega)}-1\right)\right)\right)\\
&\leqq \prod_{\ell\leqq q}\left(\frac{1}{P(B_\ell)^q}\left(a(q)\lambda^q + \left(1+(1-\lambda_\ell)\left(\frac{1}{g_\ell(\omega)}-1\right)\right)^q\right)^{1/q}\right) ,
\end{aligned}$$

where the second inequality uses Hölder's inequality in the two point space. We now use (5.6) and we get a bound

$$\prod_{\ell\leqq q}\left(\frac{1}{P(B_\ell)^q}(q+1-qg_\ell(\omega))\right)^{1/q} .$$

We now integrate in ω, using Hölder's inequality, to get

$$\int G'(x,\omega)dP(x)\,d\mu(\omega) \leqq \prod_{\ell\leqq q}\left(\frac{1}{P(B_\ell)^q}\left(q+1-q\frac{P'(A_\ell)}{P(B_\ell)}\right)\right)^{1/q}$$

where we have used the fact that

$$\int g_\ell(\omega)d\mu(\omega) = \frac{P'(A_\ell)}{P(B_\ell)}$$

by Fubini Theorem. But now, for $x \leqq 1$,

$$q+1-qx \leqq x^{-q} .$$

The proof is finished. □

To state the second result of this section, we introduce, for $y_1,\ldots,y_q,x$ in Ω^n the number

$$V(y_1,\ldots,y_q,x) = \mathrm{card}\left\{i \leqq n;\ \sum_{\ell\leqq q} h_i(x,y_\ell) \geqq 2\right\} .$$

Given sets $A_1,\ldots,A_q$ we set

$$V(A_1,\ldots,A_q,x) = \inf\{V(y_1,\ldots,y_q,x);\, y_1\in A_1,\ldots,y_q\in A_q\} . \tag{5.7}$$

We consider the number τ such $e^{\tau/2}+e^{-\tau}=2$.

Theorem 5.4. *For all subsets $A_1,\ldots,A_q$ of Ω^n we have*

$$\int \exp\frac{\tau}{q}V(A_1,\ldots,A_q,x)dP(x) \leqq \frac{1}{(\prod_{\ell\leqq q}P(A_\ell))^{1/q}} .$$

As in Theorem 5.1, we consider different sets to allow the induction to proceed. However, we will be interested in the following consequence

$$P(V_q(A,\cdot\,) > u) \leqq \frac{1}{P(A)} \exp\left(-\frac{u\tau}{q}\right) . \tag{5.8}$$

where

$$V_q(A,x) = V(A,\ldots,A,x) \ ,$$

(q occurrences of A).

An essential feature of (5.8) is that, as q increases, the coefficient $1/P(A)$ does not blow up. The relevance of this will be made clear in the proof of Proposition 5.6. An interesting feature of Theorem 5.4 is that it does not seem possible in this statement to find the points $y_1,\ldots,y_q$ that witness (the order of) $V(A_1,\ldots,A_q,x)$, by selecting them independently according to a given measure v. We sketch the argument for this claim. Consider a probability measure v on A. Consider, for $i \leqq q$, the set $A_i = \{y \in A; h_i(x,y) = 1\}$. Then

$$P(\exists i,m \leqq q, y_i, y_m \in A_\ell) \geqq \frac{1}{K} \min(1, q^2 v(A_\ell)^2)$$

so that

$$\int \frac{1}{q} V(y_1,\ldots,y_q,x) dv(y_1),\ldots,dv(y_q) \geqq \frac{1}{Kq} \sum_{\ell \leqq n} \min(1, q^2 v(A_\ell)^2) \ .$$

Study of the “canonical example” show that it is not possible that

$$\frac{1}{q} \int \inf_{v} \sum_{\ell \leqq n} \min(1, q^2 v(A_\ell)) dP(x)$$

remain bounded independently of q,n when $P(A) \leqq 1/2$.

The proof of Theorem 5.4 will rely upon the following:

Proposition 5.5. *Given $q \geqq 2$, and numbers $g_1 \leqq \cdots \leqq g_q \leqq 1$, we have*

$$\min\left(e^{\tau/q}, \left(\prod_{j=2}^{q} g_j\right)^{-1/q}\right) + \frac{1}{q}\sum_{j=1}^{q} g_j \leqq 2 \, . \tag{5.9}$$

It is of course essential that the term g_1 not be included in the product.

Lemma 5.6. *For $2 \leqq p \leqq q$, we have*

$$qe^{\tau/q} + pe^{-\tau/(p-1)} \leqq p + q \, . \tag{5.10}$$

Proof. For $x \geqq 1$, consider the function

$$f(x) = x(e^{1/x} - 1) \, .$$

Then

$$f'(x) = e^{1/x}\left(1 - \frac{1}{x}\right) - 1 \leqq 0$$

because $e^u(1-u) \leqq 1$ for $u \leqq 1$. Thereby to prove (5.10) it suffices to show that it holds for $q = p$, i.e.

$$e^{\tau/p} + e^{-\tau/(p-1)} \leqq 2\,.$$

Thus, it suffices to show that for $0 \leqq x \leqq 1/2$ we have

$$e^{\tau x} + e^{\tau x/1-x} \leqq 2\,. \tag{5.11}$$

Consider the function $f(x) = e^{\tau x} + e^{\tau x/1-x}$. Calculus shows that $f'(x)$ has the sign of

$$h(x) = \tau\left(x - \frac{x^2}{2}\right) + (1-x)\log(1-x)\,.$$

Now, $h'(x) = (\tau - 1) - \tau x - \log(1-x)$. Since $h'' \geqq 0$, h' increases. Thus, h' is first negative, then positive. Thus h is first negative, then positive. Thus f decreases, then increases, and thus to prove (5.11) it suffices to prove it for $x = 1/2$, where it is true by choice of τ. □

We now turn to the proof of Proposition 5.5. By increasing some g_i's if necessary we see that it suffices to assume that $\prod_{j=2}^q g_j \geqq e^{-\tau}$.

We observe that, given $a \geqq b$, when ab is kept constant, $a + b$ increases if a increases. It should then be clear that to prove (5.7) one can assume

$$g_1 = \cdots = g_p < g_{p+1} \leqq g_{p+2} = \cdots = g_q = 1$$

where $p + 1 \leqq q$. Thus we have to show

$$(g_1^{p-1} g_{p+1})^{-1/q} + \frac{1}{q}(pg_1 + g_{p+1}) \leqq 1 + \frac{p+1}{q} \tag{5.12}$$

under the conditions $g_1 \leqq g_{p+1} \leqq 1$, $g_1^{p-1} g_{p+1} \geqq e^{-\tau}$.

To prove that a function $Ax + Bx^{-1/q}$ is at most V in an interval $[a, b]$, it suffices (when $A, B \geqq 0$) to show that this is true for $x = a$ or $x = b$. Thus it suffices to prove (5.12) when $g_{p+1} = 1$, $g_{p+1} = g_p$, $g_{p+1} = g_1^{-(p-1)} e^{-\tau}$.

Case 1. $g_{p+1} = 1$. Then (5.12) becomes

$$g_1^{-\frac{p-1}{q}} + \frac{p}{q} g_1 \leqq 1 + \frac{p}{q}\,. \tag{5.13}$$

Since $g_1^{p-1} \geqq e^{-\tau}$, it suffices to check (5.13) when $g_1 = 1$ (obvious) and $g_1 = e^{-\tau/(p-1)}$, where it follows from (5.10).

Case 2. $g_{p+1} = g_p$. The analysis is identical to the previous one, except that one should change p into $p + 1$.

Case 3. $g_{p+1} = g_1^{-(p-1)} e^{-\tau}$. Then we have to show that

$$e^{\tau/q} + \frac{1}{q}\left[pg_1 + \frac{e^{-\tau}}{g_1^{p-1}}\right] \leqq 1 + \frac{p+1}{q}\,. \tag{5.14}$$

Since $g_p \leqq g_{p+1} \leqq 1$, and since $g_1^{p-1} g_{p+1} \geqq e^{-\tau}$, we have $g_1^{p-1} \leqq e^{-\tau}$, $g_1^p \geqq e^{-\tau}$, so that is suffices to prove (5.14) in the case $g_1 = e^{-\tau/(p-1)}$, $g_1 = e^{-\tau/p}$, where it again follows from (5.10).

We now prove Theorem 5.4. The proof is by induction over n, and we perform only the induction from n to $n+1$. Consider sets $A_1, \ldots, A_q$ on Ω^{n+1}; for $\omega \in \Omega$, $\ell \leqq q$ consider $A_{\ell,\omega}$ as in (5.6); consider B_ℓ is in (5.5).

The key to the proof is the following observation, where the "prime" has the same meaning as in the proof of Theorem 5.1.

$$V'(A_1, \ldots, A_q, (x, \omega)) \leqq 1 + V(B_1, \ldots, B_q, x) \tag{5.15}$$

$$V'(A_1, \ldots, A_q, (x, \omega)) \leqq V(C_1, \ldots, C_q, x) \tag{5.16}$$

where for each $\ell \leqq q, C_\ell$ is $A_{\ell,\omega}$, except for at most one exceptional index $\ell(q) \leqq q$, for which $C_{\ell(q)}$ is $B_{\ell(q)}$.

Using these bounds and induction, we see that

$$\begin{aligned}\int \exp \frac{\tau}{q} V'(A_1, \ldots, A_q, (x, \omega)) dP(x) \\ \leqq \min\left(\frac{e^{\tau/q}}{\prod_{\ell \leqq q} P(B_\ell)^{1/q}}, \frac{1}{\prod_{\ell \leqq q} P(C_\ell)^{1/q}} \right).\end{aligned} \tag{5.17}$$

We write $g_\ell(\omega) = P(A_{\ell,\omega})/P(B_\ell)$. Taking the best possible index $\ell(q)$ we get a bound

$$\frac{1}{\prod_{\ell \leqq q} P(B_\ell)^{1/q}} \min\left(e^\tau, \frac{\min_{\ell \leqq q} (g_\ell(\omega))^{1/q}}{\prod_{\ell \leqq q} (g_\ell(\omega))^{1/q}} \right)$$

and, by (5.9), this is at most

$$\frac{1}{\prod_{\ell \leqq q} P(B_\ell)^{1/q}} \left(2 - \frac{1}{q} \sum_{\ell=1}^{q} g_\ell(\omega) \right).$$

Integration of (5.17) in ω, use of Fubini theorem and of the observation $\int g_\ell(\omega) d\mu(\omega) = P'(A_\ell)/P(B_\ell)$ yield

$$\int \exp \frac{\tau}{q} V'(A_1, \ldots, A_q, z)\, dP'(z) \leqq \frac{1}{\prod_{\ell \leqq q} P(B_\ell)^{1/q}} \left(2 - \frac{1}{q} \sum_{\ell \leqq q} \frac{P'(A_\ell)}{P(B_\ell)} \right). \tag{5.18}$$

Now, using the inequality $\log x \leqq x - 1$, we see that if $r_\ell \leqq 1$

$$2 - \frac{1}{q} \sum_{\ell \leqq q} r_\ell \leqq 1 + \sum_{\ell \leqq q} \frac{1}{q}(1 - r_\ell) \leqq 1 + \frac{1}{q} \log\left(\frac{1}{\prod_{\ell \leqq q} r_\ell} \right) \leqq \frac{1}{(\prod_{\ell \leqq q} r_\ell)^{1/q}}.$$

Combining with (5.18) this concludes the proof. □

We now give an application of Theorem 5.4. Although the result obtained is weaker than (1.13), the simple proof seems nonetheless of interest. We consider a class of functions $\mathscr{F}$ from Ω to $[0,1]$. For a subset F of Ω we set

$\theta(F) = \sup_{f \in \mathscr{F}} \sum_{\omega \in F} f(\omega)$. Thus, for each $p \geqq 1$, each finite sets $F, F_1, \ldots, F_p$, each numbers $(\alpha_i)_{i \leqq p}, \alpha_i \geqq 0$, we have

$$1_F \leqq \sum_{i \leqq p} \alpha_i 1_{F_i} \Rightarrow \theta(F) \leqq \sum_{i \leqq p} \alpha_i \theta(F_i)\,. \tag{5.19}$$

We also have

$$\forall \omega \in \Omega, \quad \theta(\{\omega\}) \leqq 1\,. \tag{5.20}$$

Consider independent random variables $(X_i)_{i \leqq n}$ valued in Ω. In order to simplify notations, we suppose that their common law μ has no atoms, so that the points $(X_i)_{i \leqq n}$ are all distinct a.s., and there is no ambiguity about the meaning of $Z = \theta(X_1, \ldots, X_n)$; we assume this function to be measurable.

We denote by M a median of Z.

Proposition 5.6. *For each $t \leqq 2M$ we have*

$$P(|Z - M| \geqq t) \leqq K \exp -\frac{t^2}{KM}\,.$$

Proof. We consider the function $Z(x) = \theta(\{x_1, \ldots, x_n\})$ on Ω^n. Considering $a > 0$, we set $A = \{y; Z(y) \leqq a\}$.

Lemma 5.7. *For each $x \in \Omega^n$, we have*

$$Z(x) \leqq \frac{q}{q-1} a + V_q(A, x)\,.$$

Proof. Consider $y_1, \ldots, y_q$ in A such that

$$V_q(A, x) = \operatorname{card} I$$

where

$$I = \left\{ i \leqq n; \sum_{\ell \leqq q} h_i(x, y_\ell) \geqq 2 \right\}\,.$$

Thus

$$i \notin I \Rightarrow \sum_{\ell \leqq q} h_i(x, y_\ell) \leqq 1\,.$$

Thus, if we define

$$F_\ell = \{x_i; h_i(x, y_\ell) = 0\}$$

we have

$$\forall i \notin I, \quad 1_{\{x_i\}} \leqq \frac{q}{q-1} \sum_{\ell \leqq q} 1_{F_\ell}$$

and if

$$F = \{x_i; i \notin I\}\,,$$

we have

$$1_F \leqq \frac{q}{q-1} \sum_{\ell \leqq q} 1_{F_\ell}$$

so that, by (5.19),

$$\theta(F) \leqq \frac{q}{q-1} \sum_{\ell \leqq q} \theta(F_i)\,.$$

Now, by (5.19), $\theta(F_i) \leqq \theta(\{y_1, \ldots, y_\ell\}) \leqq a$. Moreover,

$$\theta(\{x_1, \ldots, x_n\}) \leqq \theta(F) + \operatorname{card} I$$

by (5.18) and (5.19). This completes the proof. □

Combining Lemma 5.7 with (5.8), we get

$$P\left(Z \geqq \frac{q}{q-1}a + u\right) \geqq \frac{1}{P(A)} \exp\left(-\frac{u\tau}{q}\right).$$

We first take $a = M$, so that

$$P\left(Z \geqq M + \frac{1}{q-1}M + u\right) \leqq 2\exp\left(-\frac{u\tau}{q}\right)$$

and thus, taking q the smallest such that $M(q-1) \leqq u$, we get

$$P(Z \geqq M + 2u) \leqq 2\exp\left(-\frac{u^2}{KM}\right).$$

Now, if $aq/(q-1) + u \leqq M$, then (5.20) implies

$$P(Z \leqq a) \leqq 2\exp\left(-\frac{u\tau}{q}\right).$$

Taking $a = M - 2u$, and q as above yield the result. □

Remark. When t is small, we need to take large values of q. It is hence essential to have $P(A)^{-1}$ rather than $P(A)^{-q}$ in (5.8).

6. Large sets

The theory of concentration of measure is closely related to the topic of classical isoperimetry. Roughly speaking, it offers weaker statements in a more general setting.

The exact extent of the connections remains to be understood. One such connection will be described in the present section. To provide perspective, we will discuss a typical situation where isoperimetry is fundamental, the case of $\mathbb{R}^n$, provided with the Euclidean distance d and the canonical gaussian measure γ_n. For a subset A of $\mathbb{R}^n$, we denote $d(x,A)$ the distance of x to $\mathbb{R}^n$. It is a weak (but nontrivial) consequence of gaussian isoperimetry that, if $\varepsilon = 1 - P(A)$ is small we have

$$\int d^2(x,A)\,d\gamma_n(x) \leqq \frac{K\varepsilon}{\log 1/\varepsilon} \tag{6.1}$$

independently of n, A. To convince oneself that (6.1) is already of interest, one can observe that it contains the fact that $\gamma_n(\{d(\cdot, A) \geqq K/\sqrt{\log 1/\varepsilon}\}) \leqq \varepsilon/2$.

On the other hand, it is known that the isoperimetric properties of $\{0,1\}^n$ resemble those of Gauss space when the notion of boundary of a set is suitably defined. The purpose of the present section is to prove an inequality of the same nature as (6.1), but where the notion of "distance" is replaced by the functional $e(A, x)$ of Sect. 2. For simplicity we consider only the case $\Omega = \{0,1\}$, μ uniform.

Theorem 6.2. *We can choose $\beta > 0$ such that the function $e(A, x)$ of Sect. 2 satisfies*

$$\int_{\{0,1\}^n} (e(A,x) - 1) dP(x) \leqq \frac{(1 - P(A))}{P(A) \log e/(1 - P(A))} \tag{6.2}$$

for all n and all subsets A of $\{0,1\}^n$.

To make the connection with (6.1), one should observe that $e(A,x) - 1 \geqq 0$, and is zero if $x \in A$.

Certainly the first task is to understand the properties of the function

$$\varphi(x) = \frac{x}{(1-x)\log(e/x)} .$$

Lemma 6.1. *There is a number M such that, whenever $0 \leqq b_0 \leqq b_1 \leqq 1$ we have*

$$\varphi(b_0) + \varphi(b_1) - 2\varphi\left(\frac{b_0 + b_1}{2}\right) \leqq M \frac{(\varphi(b_1) - \varphi(b_0))^2}{2b_1 + \varphi(b_0)} .$$

Proof. We first observe that φ' increases, so that φ is convex.

Consider first the case where $\varphi(b_0) \leqq \varphi(b_1)/2$. In this case, by convexity

$$\varphi(b_0) + \varphi(b_1) - 2\varphi\left(\frac{b_0 + b_1}{2}\right) \leqq \varphi(b_1) - 2\varphi\left(\frac{b_1}{2}\right)$$

while

$$\varphi(b_1) - \varphi(b_0) \geqq \frac{1}{2}\varphi(b_1) .$$

Thereby, it suffices to show that if $0 < x < 1$,

$$\varphi(x) - 2\varphi\left(\frac{x}{2}\right) \leqq \frac{K\varphi(x)^2}{x + \varphi(x)}$$

a fact left to the reader.

Assume now that $\varphi(b_0) \geqq \varphi(b_1)/2$. Then we write

$$\varphi(b_0) + \varphi(b_1) - 2\varphi\left(\frac{b_0 + b_1}{2}\right) \leqq \frac{(b_1 - b_0)^2}{2} \sup\{\varphi''(z); b_0 \leqq z \leqq b_1\} .$$

On the other hand, by convexity we have

$$\varphi(b_1) - \varphi(b_0) \geqq (b_1 - b_0)\varphi'(b_0)\,.$$

To conclude the reader will prove the following two elementary facts: First, if $\varphi(x) \geqq \varphi(y)/2$, then $y \leqq Kx$ (it is useful to distinguish whether, say $x \leqq 1/2$ or $x \geqq 1/2$). Second, we have $\varphi''(x)(x + \varphi(x)) \leqq K\varphi'(x)^2$. □

We will select the number β later; we recall that $\gamma = 1 + \beta^2$. Since $e(A,x) = 1$ if $x \in A$, to prove (6.2) it suffices to prove that if B is one set that contains the complement of A, then

$$\int_B e(A,x)dP(x) \leqq P(B) + \varphi(1 - P(A))\,. \tag{6.3}$$

Observe also that if (6.3) holds for one such set B, it holds for each such set B.

The proof of (6.3) will again be by induction upon n. Consider $A \subset \{0,1\}^{n+1}$, and for $j \in \{0,1\}$, write $A_j = \{x \in \{0,1\}^n; (x,j) \in A\}$, and set $a_j = P(A_j)$. For specificity we assume $a_1 \leqq a_0$. Proceeding as in the proof of Theorem 2.1, we see that for each $0 \leqq \lambda \leqq 1$ and each $x \in \{0,1\}^n$ we have

$$\begin{aligned} e(A,(x,1)) \leqq\ & (1-\lambda)^2 e(A_1,x) + 2\lambda(1-\lambda)(e(A_1,x)e(A_0,x))^{1/2} \\ & + \lambda^2\gamma e(A_0,x)\,. \end{aligned} \tag{6.4}$$

Also, we have $e(A,(x,0)) \leqq e(A_0,x)$. Consider now $C = \{0,1\}^n \backslash (A_0 \cup A_1)$ and $B = C \times \{0,1\}$. We will prove (6.3) for this choice of B. Using (6.4), Cauchy–Schwarz, we have

$$\begin{aligned} \int_B e(A,y)dP'(y) \leqq\ & \frac{1}{2}(1-\lambda)^2 \int_C e(A_1,x)dP(x) \\ & + \lambda(1-\lambda)\left(\int_C e(A_0,x)dP(x) \int_C e(A_1,x)dp(x)\right)^{1/2} \\ & + \frac{\gamma\lambda^2}{2}\int_C e(A_0,x)dP(x)\,. \end{aligned}$$

We set $c = P(C)$, $b_j = 1 - a_j$, $d_j = c + \varphi(b_j)$, and we use the induction hypothesis. We are then reduced to show that

$$\begin{aligned} & d_0 + \inf_{0 \leqq \lambda \leqq 1}((1-\lambda)^2 d_1 + 2\lambda(1-\lambda)\sqrt{d_0 d_1} + \gamma\lambda^2 d_0) \\ & \leqq 2\left(c + \varphi\left(\frac{b_0 + b_1}{2}\right)\right)\,. \end{aligned} \tag{6.5}$$

We note that $d_1 \geqq d_0$ (since $b_0 \leqq b_1$). As in the proof of Lemma 2.3, we have to consider only the case $d_1 \leqq \gamma^2 d_0$, where the inf in (6.4) is obtained for

$$\lambda = \frac{d_1 - \sqrt{d_1 d_0}}{d_1 - 2\sqrt{d_0 d_1} + \gamma d_0}$$

which satisfies $0 \leqq \lambda \leqq 1$. Then (6.4) reduces to

$$d_0 + d_1 - \frac{(d_1 - \sqrt{d_0 d_1})^2}{d_1 - 2\sqrt{d_0 d_1} + \gamma d_0} + \gamma d_0 \leqq 2\left(c + \varphi\left(\frac{b_0 + b_1}{2}\right)\right).$$

Now

$$(d_1 - \sqrt{d_0 d_1})^2 = d_1 \frac{(d_1 - d_0)^2}{(\sqrt{d_0} + \sqrt{d_1})^2} \geqq \frac{1}{4}(d_1 - d_0)^2$$

$$d_1 + \gamma d_0 - 2\sqrt{d_1 d_0} = (\gamma - 1)d_0 + (\sqrt{d_1} - \sqrt{d_0})^2 \leqq \gamma(\gamma - 1)d_0$$

since $d_1 \leqq \gamma^2 d_0$. Thus it suffices to show that

$$d_0 + d_1 - \frac{(d_0 - d_1)^2}{4\gamma(\gamma - 1)d_0} \leqq 2c + 2\varphi\left(\frac{b_0 + b_1}{2}\right)$$

or, equivalently, since $c \leqq 2b_1$,

$$\varphi(b_0) + \varphi(b_1) - 2\varphi\left(\frac{b_0 + b_1}{2}\right) \leqq \frac{(\varphi(b_1) - \varphi(b_0))^2}{4\gamma(\gamma - 1)(2b_1 + \varphi(b_0))}.$$

We then realize that it suffices to take β close enough to 0 that $(4\gamma(\gamma - 1))^{-1}$ is larger than the number M of Lemma 6.1 (and that the theorem holds for $n = 1$). □

Acknowledgements. I am grateful to A. Dembo for pointing out that a previous version of Theorem 5.1 was stated in an unclever way.

References

[B] A. Bonami: Etude des coefficients de Fourier des fonctions de $L_p(G)$, Ann. Inst. Fourier **20**, 335–402 (1970)

[B-M1] L. Bergé, P. Massart: Minimum contrast estimations on sieves (to appear)

[B-M2] L. Bergé, P. Massart: From Model Selection to Adoptive estimation, Festschrift in honor of LeCam's 70th Birthday, Ed. N. Pollard, G. Young, Springer

[D] A. Dembo: Information inequalities and concentration of measure, Preprint, 1995

[D-Z] A. Dembo, O. Zeitouni: Transportation approach to some concentration inequalities in product spaces, Preprint, 1996

[L1] M. Ledoux: Inégalités isopérimétriques en Analyse et en probabilités, Seminaire Bourkaki, Juin 1993, Astérisque **216**, 343–375 (1993)

[L2] M. Ledoux: Isoperimetry and Gaussian Analysis, Ecole d'été de Probabilités de Saint-Flour 1994, Lecture Notes in Math, Springer Verlag (to appear)

[L3] M. Ledoux: On Talagrand's deviation inequalities for product measures

[L-T] M. Ledoux, M. Talagrand: Probability in Banach Spaces, Springer Verlag, 1991

[M-S] V. Milman, G. Schetchman: Asymptotic theory of Finite dimensional normed spaces, Lecture Notes in Math 1200, Springer Verlag, 1985

[P] G. Pisier: Probabilistic methods in the geometry of Banach Spaces, Probability and analysis, Varema (Italy) 1985, Lecture Notes in Mathematics, Vol. 1206, Springer Verlag, 167–241 (1986)

[T1] M. Talagrand: A simple proof of the majorizing measure theorem, Geometric and Functional analysis **2**, 119–125 (1992)

[T2] M. Talagrand: Sharper bounds for gaussian and empirical processes, Ann. Probab. **22**, 28–76 (1994)

[T3] M. Talagrand: Concentration of measure and isoperimetric inequalities in product spaces, Publications I.H.E.S., **81**, 73–205 (1995)

[T4] M. Talagrand: A new look at independence, Ann. Probab. **24**, 1–34 (1996)

[T5] M. Talagrand: Are all sets of positive measure essentially convex? Operator theory, advances and application, **77**, Birkhäuser 295–310 (1995)