A NOTE ON TALAGRAND'S CONVEX HULL CONCENTRATION INEQUALITY

DAVID POLLARD

ABSTRACT. The paper reexamines an argument by Talagrand that leads to a remarkable exponential tail bound for the concentration of probability near a set. The main novelty is the replacement of a mysterious calculus inequality by an application of Jensen's inequality.

1. INTRODUCTION

Let $\mathcal{X}$ be a set equipped with a sigma-field $\mathcal{A}$. For each vector $w = (w_1, \dots, w_n)$ in $\mathbb{R}_+^n$, the weighted Hamming distance between two vectors $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$, in $\mathcal{X}^n$ is defined as

$$d_w(x,y) := \sum\nolimits_{i \le n} w_i h_i(x,y) \qquad \text{where } h_i(x,y) = \begin{cases} 1 & \text{if } x_i \neq y_i \\ 0 & \text{otherwise.} \end{cases}$$

For a subset A of $\mathcal{X}^n$ and $x \in \mathcal{X}^n$, the distances $d_w(x, A)$ and $D(x, A)$ are defined by

$$d_w(x) := \inf\{y \in A : d_w(x,y)\} \qquad \text{and} \qquad D(x,A) := \sup\nolimits_{w \in \mathcal{W}} d_w(X, A).$$

where the supremum is taken over all weights in the set

$$\mathcal{W} := \{(w_1, \dots, w_n) : w_i \ge 0 \text{ for each } i \text{ and } |w|^2 := \sum\nolimits_{i \le n} w_i^2 \le 1\}.$$

Talagrand (1995, Section 4.1) proved a remarkable concentration inequality for random elements $X = (X_1, \dots, X_n)$ of $\mathcal{X}^n$ with independent coordinates and subsets $A \in \mathcal{A}^n$:

$$\text{(1)} \qquad \mathbb{P}\{X \in A\}\mathbb{P}\{D(X,A) \ge t\} \le \exp(-t^2/4) \qquad \text{for all } t \ge 0.$$

As Talagrand showed, this inequality has many applications to problems in combinatorial optimization and other areas. See Talagrand (1996b), Steele (1997, Chapter 6), and McDiarmid (1998, Section 4) for further examples.

Date: 13 January 2007.

2000 *Mathematics Subject Classification*. Primary 62E20. Secondary: 60F05, 62G08, 62G20.

Key words and phrases. Concentration of measure; convex hull; convexity.

There has been a strong push in the literature to establish concentration and deviation inequalities by "more intuitive" methods, such as those based on the tensorization, as in Ledoux (1996), Boucheron, Lugosi, and Massart (2000), Massart (2003), and Lugosi (2003). I suspect the search for alternative approaches has been driven by the miraculous roles played by some elementary inequalities in Talagrand's proofs.

Talagrand (1995) used an induction on n to establish his result. He invoked a slightly mysterious inequality in the inductive step,

$$\inf_{0\le\theta\le 1} u^{-\theta}\exp\left(\frac{(1-\theta)^2}{4}\right) \le 2-u \qquad \text{for } 0<u<1,$$

which he borrowed from Johnson and Schechtman (1991)—see Talagrand's discussion following his Lemma 4.1.3 for an explanation of how those authors generalized the earlier result by Talagrand (1988). There is similar mystery in Talagrand's Lemma 4.2.1, which (in my notation) asserts that

$$\sup_{0\le\theta\le 1} u^{-\theta/c}\exp\big(\psi_c(1-\theta)\big) \le \frac{1+c-u}{c} \qquad \text{for } 0<u<1,$$

where ψ_c is defined in equation (6) below. (My $\psi_c(u)$ equals Talagrand's $\xi(\alpha,u)$ with $\alpha = 1/c$.) Talagrand (1991) had used this inequality to further generalize the result of Johnson and Schechtman, giving the concentration inequality listed as Theorem 4.2.4 in his 1995 paper. It was my attempts to understand how he arrived at his $\xi(\alpha,u)$ function that led me to the conccavity argument that I present in the present note.

It is my purpose to modify Talagrand's proof so that the inductive step becomes a simple application of the Hölder inequality (essentially as in the original proof) and the Jensen inequality. Most of my methods are minor variations on the methods in the papers just cited; my only claim of originality is for the recognition that the mysterious inequalities can be replaced by more familiar appeals to concavity. See also the Remarks at the end of this Section.

The distance $D(x,A)$ has another representation, as a minimization over a convex subset of $[0,1]$. Write $h(x,y)$ for the point of $\{0,1\}^n$ with ith coordinate $h_i(x,y)$. For each fixed x, the function $h(x,\cdot)$ maps A onto a subset $h(x,A) := \{h(x,y) : y\in A\}$ of $\{0,1\}^n$. The convex hull, $\text{co}\,(h(x,A))$, of $h(x,A)$ in $[0,1]^n$ is compact, and

$$D(x,A) = \inf\{|\xi| : \xi \in \text{co}\,(h(x,A))\}.$$

Each point ξ of $\text{co}\,(h(x,A))$ can be written as $\int h(x,y)\,\nu(dy)$ for a ν in the set $\mathcal{P}(A)$ of all Borel probability measures for which $\nu(A) = 1$. That is,

$\xi_i = \nu\{y \in A : y_i \neq x_i\}$. Thus

$$D(x,A)^2 = \inf_{\nu \in \mathcal{P}(A)} \sum\nolimits_{i \leq n} \left(\nu\{y \in A : y_i \neq x_i\}\right)^2 . \tag{2}$$

Talagrand actually proved inequality (1) by showing that

$$\mathbb{P}\{X \in A\}\mathbb{P}\exp\left(\tfrac{1}{4}D(X,A)^2\right) \leq 1. \tag{3}$$

He also established an even stronger result, in which the $D(X,A)^2/4$ in (3) is replaced by a more complicated distance function.

For each convex, increasing function ψ with $\psi(0) = 0 = \psi'(0)$ define

$$F_\psi(x,A) := \inf_{\nu \in \mathcal{P}(A)} \sum\nolimits_{i \leq n} \psi\left(\nu\{y \in A : y_i \neq x_i\}\right), \tag{4}$$

For each $c > 0$, Talagrand (1995, Section 4.2) showed that

$$(\mathbb{P}\{X \in A\})^c \mathbb{P}\exp\left(F_{\psi_c}(X,A)\right) \leq 1, \tag{5}$$

where

(6)

$$\begin{aligned} \psi_c(\theta) &:= c^{-1}\left((1-\theta)\log(1-\theta) - (1-\theta+c)\log\left(\frac{(1-\theta)+c}{1+c}\right)\right) \\ &= \sum\nolimits_{k\geq 2} \frac{\theta^k}{k}\left(\frac{R_c + R_c^2 + \cdots + R_c^{k-1}}{(k-1)}\right) \qquad \text{with } R_c := \frac{1}{c+1}. \\ &\geq \frac{\theta^2}{2+2c} \end{aligned}$$

As you will see in Section 3, this strange function is actually the largest solution to a differential inequality,

$$\psi''(1-\theta) \leq 1/(\theta^2 + \theta c) \qquad \text{for } 0 < \theta < 1.$$

Inequality (5) improves on (3) because $D(x,A)^2/4 \leq F_{\psi_1}(x,A)$.

Following the lead of Talagrand (1995, Section 4.4), we can ask for general conditions on the convex ψ under which an analog of (5) holds with some other decreasing function of $\mathbb{P}\{X \in A\}$ as an upper bound. The following slight modification of Talagrand's theorems gives a sufficient condition in a form that serves to emphasize the role played by Jensen's inequlity.

Theorem 1. *Suppose γ is a decreasing function with $\gamma(0) = \infty$ and ψ is a convex function. Define $G_\psi(\eta,\theta) := \psi(1-\theta) + \theta\eta$ and $G_\psi(\eta) := \inf_{0\leq\theta\leq 1} G_\psi(\eta,\theta)$ for $\eta \in \mathbb{R}^+$. Suppose*

(i) $r \mapsto \exp\left(G_\psi(\gamma(r) - \gamma(r_0))\right)$ *is concave on* $[0, r_0]$, *for each* $r_0 \leq 1$

(ii) $(1-p)e^{\psi(1)} + p \leq e^{\gamma(p)}$ *for* $0 \leq p \leq 1$.

Then $\mathbb{P}\exp\left(F_\psi(X,A)\right) \leq \exp\left(\gamma\left(\mathbb{P}\{X \in A\}\right)\right)$ for every $A \in \mathcal{A}^n$ and every random element X of $\mathcal{X}^n$ with independent components.

The next lemma, a more general version of which is proved in Section 3, leads to a simple sufficient condition for the concavity assumption (i) of Theorem 1 to hold.

Lemma 2 (Concavity lemma)**.** *Suppose* $\psi : [0,1] \to \mathbb{R}^+$ *is convex and increasing, with* $\psi(0) = 0 = \psi'(0)$ *and* $\psi''(\theta) > 0$ *for* $0 < \theta < 1$. *Suppose* $\xi : [0, r_0] \to \mathbb{R}^+ \cup \{\infty\}$ *is continuous and twice differentiable on* $(0, r_0)$. *Suppose also that there exists some finite constant* c *for which* $\xi''(r) \leq c\xi'(r)^2$ *for* $0 < r < r_0$. *If*

$$\psi''(1-\theta) \leq 1/(\theta^2 + \theta c) \qquad \text{for } 0 < \theta < 1$$

then the function $r \mapsto \exp\big(G_\psi(\xi(r))\big)$ *is concave on* $[0, r_0]$.

The Lemma will be applied with $\xi(r) = \gamma(r) - \gamma(r_0)$ for $0 \leq r \leq r_0$. As shown in Section 3, the conditions of the Lemma hold for $\psi(\theta) = \theta^2/4$ with $\gamma(r) = \log(1/r)$ and also for the ψ_c from (6) with $\gamma(r) = c^{-1}\log(1/r)$.

Remarks.

(i) If $\gamma(0)$ were finite, the inequality asserted by Theorem 1 could not hold for all nonempty A and all X. For example, if each X_i had a nonatomic distribution and A were a singleton set we would have $F_\psi(X, A) = n\psi(1)$ almost surely. The quantity $\mathbb{P}\exp\big(F_\psi(X, A)\big)$ would exceed $\exp(\gamma(0))$ for large enough n. It it to avoid this difficulty that we need $\gamma(0) = \infty$.

(ii) Assumption (ii) of the Theorem, which is essentially an assumption that the asserted inequality holds for $n = 1$, is easy to check if γ is a convex function with $\gamma(1) \geq 0$. For then the function $B(p) := \exp(\gamma(p))$ is convex with $B(1) \geq 1$ and $B'(1) = \gamma'(1)e^{\gamma(1)}$. We have

$$B(p) \geq (1-p)e^{\psi(1)} + p \qquad \text{for all } p \text{ in } [0,1]$$

if $B'(1) \leq 1 - e^{\psi(1)}$.

(iii) A helpful referee has noted that both my specific examples are already covered by Talagrand's results. He (or she) asked whether there are other (ψ, γ) pairs that lead to other useful concentration inequalities. A good question, but I do not yet have any convincing examples. Actually, I had originally thought that my methods would extend to the limiting case where c tends to zero, leading to an answer to the question posed on page 128 by Talagrand (1995). Unfortunately my proof ran afoul of the requirement $\gamma(0) = \infty$. I suspect more progress might be made by replacing the strong assumption on ψ'' from Lemma 2 by something closer to the sufficient conditions presented in Section 3.

2. Proof of Theorem 1

Argue by induction on n. As a way of keeping the notation straight, replace the subscript on $F_\psi(x, B)$ by an n when the argument B is a subset of $\mathcal{X}^n$. Also, work with the product measure $\mathbb{Q} = \otimes_{i\le n} Q_i$ for the distribution of X and $\mathbb{Q}_{-n} = \otimes_{i<n} Q_i$ for the distribution of $(X_1, \dots, X_{n-1})$. The assertion of the Theorem then becomes

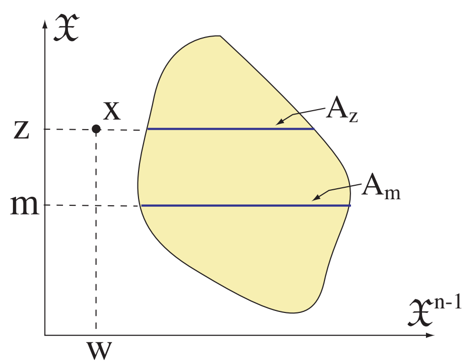
$$\mathbb{Q}\exp\big(F_n(x, A)\big) \le \exp(\gamma(\mathbb{Q}A))$$

For $n = 1$ and $B \in \mathcal{A}$ we have $F_1(x, B) = \psi(1)\{x \notin B\} + 0\{x \in B\}$ so that $Q_1 \exp\big(F_1(x, B)\big) \le (1-p)e^{\psi(1)} + p$, where $p = Q_1 B$. Assumption (ii) then gives the desired $\exp(\gamma(p))$ bound.

Now suppose that $n > 1$ and that the inductive hypothesis is valid for dimensions strictly smaller than n. Write $\mathbb{Q}$ as $\mathbb{Q}_{-n} \otimes Q_n$. To simplify notation, write w for $x_{-n} := (x_1, \dots, x_{n-1})$ and z for x_n. Define the cross section $A_z := \{w \in \mathcal{X}^{n-1} : (w, z) \in A\}$ and write R_z for $\mathbb{Q}_{-n} A_z$. Define $r_0 := \sup_{z\in\mathcal{X}} R_z$. Notice that $r_0 \ge Q_n^z R_z = \mathbb{Q}A$.

The key to the proof is a recursive bound for F_n: for each $x = (w, z)$ with $A_z \ne \emptyset$, each m with $A_m \ne \emptyset$, and all $\theta = 1 - \bar\theta \in [0, 1]$,

$$\psi(\bar\theta) + F_n(x, A) \le \theta F_{n-1}(w, A_z) + \bar\theta F_{n-1}(w, A_m). \tag{7}$$



To establish inequality (7), suppose μ_z is a probability measure concentrated on A_z and μ_m is a probability measure concentrated on A_m. For a θ in $[0, 1]$, define $\nu = \theta\mu_z \otimes \delta_z + \bar\theta\mu_m \otimes \delta_m$, a probability measure concentrated on the subset $(A_z \times \{z\}) \cup (A_m \times \{m\})$ of A. Notice that, for $i < n$,

$$\nu\{y \in A : y_i \ne x_i\} = \theta\mu_z\{w \in A_z : y_i \ne x_i\} + \bar\theta\mu_m\{w \in A_m : y_i \ne x_i\}$$

so that, by convexity of ψ,

$$\psi\left(\nu\{y_i \ne x_i\}\right) \le \theta\psi(\mu_z\{w \in A_z : y_i \ne x_i\}) + \bar\theta\psi(\mu_m\{w \in A_m : y_i \ne x_i\});$$

and (remembering that $x_n = z$),

$$\nu\{y \in A : y_n \neq x_n\} = \begin{cases} \bar\theta & \text{if } z \neq m \\ 0 & \text{otherwise} \end{cases} \leq \bar\theta.$$

Thus

$$F_n(x, A) \quad \leq \psi(\bar\theta) + \theta \textstyle\sum_{i<n} \psi\left(\mu_z\{y_i \neq x_i\}\right) + \bar\theta \textstyle\sum_{i<n} \psi\left(\mu_m\{y_i \neq x_i\}\right).$$

The two sums over the first $n-1$ coordinates are like those that appear in the definitions of $F_{n-1}(w, A_z)$ and $F_{n-1}(w, A_z)$. Indeed, taking an infimum over all $\mu_z \in \mathcal{P}(A_z)$ and $\mu_m \in \mathcal{P}(A_m)$ we get the expression on the right-hand side of (7).

Take exponentials of both sides of (7) then integrate out with respect to $\mathbb{Q}_{-n}$ over the w component. For $0 < \theta < 1$ invoke the Hölder inquality, $\mathbb{Q}_{-n} U^\theta V^{\bar\theta} \leq \left(\mathbb{Q}_{-n} U\right)^\theta \left(\mathbb{Q}_{-n} V\right)^{\bar\theta}$, with $U = \exp(F_{n-1}(w, A_z))$ and $V = \exp(F_{n-1}(w, A_m))$, for a fixed m. For each z with $A_z \neq \emptyset$ we get

$$\begin{aligned} &\mathbb{Q}_{-n} \exp\left(F_n((w, z), A)\right) \\ &\leq e^{\psi(\bar\theta)} \left(\mathbb{Q}_{-n} \exp\left(F_{n-1}(w, A_z)\right)\right)^\theta \left(\mathbb{Q}_{-n} \exp\left(F_{n-1}(w, A_m)\right)\right)^{\bar\theta}. \end{aligned} \tag{8}$$

The inequality also hold in the extreme cases where $\theta = 0$ or $\theta = 1$, by continuity. The inductive hypothesis bounds the last product by

$$\exp\left(\psi(\bar\theta) + \theta\gamma(R_z) + \bar\theta\gamma(R_m)\right) = \exp\left(\gamma(R_m) + G(\gamma(R_z) - \gamma(R_m), \theta)\right).$$

The exponent is a decreasing function of R_m. Take an infimum over m, to replace $\gamma(R_m)$ by $\gamma(r_0)$. Then take an infimum over θ to get

$$\begin{aligned} \mathbb{Q}_{-n} \exp \quad &\left(F_n((w, z), A)\right) \leq \exp\left(\gamma(r_0) + G(\xi(R_z))\right) \\ &\text{where } \xi(r) := \gamma(R_z) - \gamma(r_0) \text{ for } 0 \leq r \leq r_0. \end{aligned} \tag{9}$$

If the cross section A_z is empty, the set $\mathcal{P}(A_z)$ is empty. The argument leading from (7) to (9) still works if we fix θ equal to zero throughout, giving the bound

$$\mathbb{Q}^w_{-n} \exp\left(F_n(x, A)\right) \leq \exp\left(\gamma(r_0) + \psi(1)\right) \qquad \text{if } A_z = \emptyset.$$

Thus the inequality (9) also holds with $R_z = 0$ when $A_z = \emptyset$, because $\xi(0) = \gamma(0) - \gamma(r_0) = \infty$ and $G(\infty) = \psi(1)$.

By Assumption (i), the function $r \mapsto \exp\left(G(\xi(r))\right)$ is concave on $[0, r_0]$. Integrate both sides of (9) with respect to Q_n to average out over the z variable. Then invoke Jensen's inequality and the fact that $Q_n R_z = \mathbb{Q}A$, to deduce that

$$\mathbb{Q} \exp\left(F_n(x, A)\right) \leq \exp\left(\gamma(r_0) + G\left(\gamma(\mathbb{Q}A) - \gamma(r_0)\right)\right).$$

Finally, use the inequality $G(\eta) \leq \eta$ to bound the last expression by $\exp(\gamma(\mathbb{Q}A))$, thereby completing the inductive step.

Remark. Note that it is important to integrate with respect to Q_n before using the bound on G: the upper bound $\exp(-\gamma(R_z))$ is a convex function of R_z, not concave.

3. Proof of the Concavity Lemma

I will establish a more detailed set of results than asserted by Lemma 2. Invoke the monotonicity and continuity of ψ' to define $g(\eta)$ as the solution to $\psi'\big(1-g(\eta)\big) = \eta$ if $0 \le \eta < \psi'(1)$ and $g(\eta) = 0$ if $\psi'(1) \le \eta$. Then the following assertions are true. (I drop the ψ subscripts for notational simplicity.)

(i)

$$G(\eta) = \begin{cases} \psi\big(1-g(\eta)\big) + \eta g(\eta) & \text{for } 0 \le \eta < \psi'(1) \\ \psi(1) & \text{for } \psi'(1) \le \eta \end{cases}$$

(ii) G is increasing and concave, with a continuous, decreasing first derivative g. In particular, $G(0) = 0$ and $G'(0) = g(0) = 1$.

(iii) $G''(\eta) = g'(\eta) = -\left[\psi''\big(1-g(\eta)\big)\right]^{-1}$ for $0 < \eta < \psi'(1)$.

(iv) $G(\eta) \le \eta$ for all $\eta \in \mathbb{R}^+$.

(v) Suppose $\xi : J \to \mathbb{R}^+$ is a convex function defined on a subinterval J of the real line, with $\xi' \neq 0$ on the interior of J. Suppose

$$\frac{1}{\psi''(1-\xi_r)} \ge g(\xi_r)^2 + g(\xi_r)\xi''(r)/\xi'(r)^2,$$

for all r in the interior of J for which $\xi_r := \xi(r) \in (0,1)$. Then $r \mapsto \exp\big(G(\xi(r))\big)$ is a concave function on J.

Proof of (i) through (iv). The fact that G is concave and increasing follows from its definition as an infimum of increasing linear functions of η. (It would also follow from the fact that $G'(\eta) = g(\eta)$, which is nonnegative and decreasing.) Replacement of the infimum over $0 \le \theta \le 1$ by the value at $\theta = 1$ gives the inequality $G(\eta) \le \eta$.

If $\eta \ge \psi'(1)$, the derivative $-\psi'(1-\theta)+\eta$ is nonnegative on $[0,1]$, which ensures that the infimum is achieved at $\theta = 1$.

If $0 < \eta < \psi'(1)$, the infimum is achieved at the zero of the derivative, $\theta = g(\eta)$. Differentiation of the defining equality $\psi'\big(1-g(\eta)\big) = \eta$ then gives the expression for $g'(\eta)$. Similarly

$$G'(\eta) = -\psi'\big(1-g(\eta)\big)g'(\eta) + \eta g'(\eta) + g(\eta) = g(\eta).$$

The infimum that defines $G(0)$ is achieved at $g(0) = 1$, which gives $G(0) = \psi(0) = 0$. Continuity of g at then gives $G'(0) = g(0) = 1$.

Proof of (v). Note that the function $L(r) := \exp\big(G(\xi(r))\big)$ is continuous on J and takes the value $e^{\psi(1)}$ for all r at which $\xi(r) \ge \psi'(1)$. The second derivative $L''(r)$ exists except possibly at points r for which $\xi(r) = \psi'(1)$. In particular, $L''(r) = 0$ when $\xi(r) > \psi'(1)$ and

$$L''(r) = \left(g'(\xi_r)(\xi_r')^2 + g(\xi_r)\xi_r'' + g(\xi_r)^2(\xi_r')^2\right) L(r) \qquad \text{for } 0 < \xi_r < \psi'(1).$$

From (iii) and the positivity of L, the last expression is ≤ 0 if and only if

$$-\frac{(\xi_r')^2}{\psi''(1 - g(\xi_r))} + g(\xi_r)\xi_r'' + g(\xi_r)^2(\xi_r')^2 \leq 0.$$

Divide through by $(\xi_r')^2$ then rearrange to get the asserted inequality for ψ''. Lemma 2 follows as a special case of (i) through (iv).

Special cases. If $\sup_r \xi''(r)/\xi'(r)^2 \leq c$, with c a positive constant, the inequality from part (v) will certainly hold if

$$\psi''(1 - \theta) \leq (\theta^2 + c\theta)^{-1} \qquad \text{for all } 0 < \theta < 1. \tag{10}$$

This differential inequality can be solved, subject to the constraints $0 = \psi(0) = \psi'(0)$, by two integrations. Indeed,

$$\psi'(1-\theta) = \int_\theta^1 \psi''(1-t)\,dt \leq \int_\theta^1 \frac{dt}{t^2 + ct} = c^{-1}\left(-\log\theta + \log\left(\frac{\theta + c}{1 + c}\right)\right)$$

and, with ψ_c defined by (6),

$$\psi(1-\theta) = \int_\theta^1 \psi'(1-t)\,dt \leq c^{-1}\int_\theta^1 -\log t + \log\left(\frac{t + c}{1 + c}\right)\,dt = \psi_c(1-\theta).$$

Note that $\psi_c(1 - \theta)$ is the solution to the differential equation

$$\psi_c''(1 - \theta) = \frac{1}{\theta^2 + c\theta} \qquad \text{for all } 0 < \theta < 1, \text{ with } \psi_c(0) = \psi_c'(0) = 0.$$

It is the largest solution to (10).

REFERENCES

Boucheron, S., G. Lugosi, and P. Massart (2000). A sharp concentration inequality with applications. *Random Structures and Algorithms 16*, 277–292.

Johnson, W. B. and G. Schechtman (1991). Remarks on Talagrand's deviation inequality for Rademacher functions. *Springer Lecture Notes in Mathematics 1470*, 72–77.

Ledoux, M. (1996). On Talagrand's deviation inequalities for product measures. *ESAIM: Probability and Statistics 1*, 63–87.

Lugosi, G. (2003). Concentration-of-measure inequalities. Notes from the Summer School on Machine Learning, Australian National University. Available at http://www.econ.upf.es/~lugosi/.

Massart, P. (2003, July). Saint-Flour Lecture Notes 2003: Concentration Inequalities and Model Selection. Available at http://www.math.u-psud.fr/˜massart/.

McDiarmid, C. (1998). Concentration. In M. Habib, C. McDiarmid, J. Ramirez-Alfonsen, and B. Reed (Eds.), *Probabilistic Methods for Algorithmic Discrete Mathematics*, pp. 195–248. Springer-Verlag.

Steele, J. M. (1997). *Probability Theory and Combinatorial Optimization*. SIAM.

Talagrand, M. (1988). An isoperimetric theorem on the cube and the Kintchine-Kahane inequalities. *Proceedings of the American Mathematical Society 104*, 905–909.

Talagrand, M. (1991). A new isoperimetric inequality for product measure and the tails of sums of independent random variables. *Geometric and Functional Analysis 1*, 211–223.

Talagrand, M. (1995). Concentration of measure and isoperimetric inequalities in product spaces. *Publications Mathématiques de l'I.H.E.S. 81*, 73–205.

Talagrand, M. (1996a). New concentration inequalities in product spaces. *Inventiones mathematicae 126*, 505–563.

Talagrand, M. (1996b). A new look at independence. *Annals of Probability 24*, 1–34.

STATISTICS DEPARTMENT, YALE UNIVERSITY, BOX 208290 YALE STATION, NEW HAVEN, CT 06520-8290.

E-mail address: `david.pollard@yale.edu`

URL: `http://www.stat.yale.edu/~pollard/`