A note on the maximal inequalities for VC classes

Rafał Latała

Abstract

We investigate generalizations of Levy and Levy-Octaviani maximal inequalities. A general conjecture is stated and proved in several particular cases.

Introduction. The famous inequality due to Levy states that for any a.s. convergent series $\sum_{i=1}^{\infty} X_i$ of independent symmetric r.v. with values in some separable Banach space and $t > 0$ we have

$$P(\max_n \|\sum_{i=1}^{n} X_i\| \geq t) \leq 2P(\|\sum_{i=1}^{\infty} X_i\| \geq t). \tag{1}$$

The generalization of Levy inequality to a nonsymmetric case is frequently called Levy-Octaviani inequality. It states that for any a.s. convergent series $\sum_{i=1}^{\infty} X_i$ of independent Banach-space valued r.v. and $t > 0$

$$P(\max_n \|\sum_{i=1}^{n} X_i\| \geq 3t) \leq 3\max_n P(\|\sum_{i=1}^{n} X_i\| \geq t). \tag{2}$$

Both Levy and Levy-Octaviani inequalities have numerous applications (e.g. see [KW]). Roughly speaking they enable often to reduce an almost sure statement to a statement in probability (like for example in Itô-Nisio theorem).

[1]Partially supported by KBN Grant (Poland) 2 P301 022 07

1991 Mathematics Subject Classification. 60G50, 60E15

Key words and phrases: sums of independent random variables, maximal inequality, VC class

However sometimes one has to consider more complicated sets of indices and ways of converging of sums of random variables. Therefore it would be very useful to have suitable versions of maximal inequalities (1) and (2) in more general setting. The purpose of this article is to propose some version of the maximal inequality and to collect known facts and conjectures about it.

Some part of this paper consists of well known facts, which are already part of the folklore. Theorem 1 and parts of Proposition 1 were communicated to the author by S. Kwapień. However we were unable to find suitable references in the existing literature (Theorem 1 is stated in [Kr], but only with an idea of the proof). Therefore, for the completeness, we decided to include these statements in our paper together with the proofs.

Notation. We will denote by (ε_i) the Bernoulli sequence, i.e. a sequence of i.i.d. symmetric r.v. taking on values ± 1. A sequence of independent standard $\mathcal{N}(0,1)$ Gaussian random variables will be denoted by (g_i).

If (T,d) is a compact metric space and $\varepsilon > 0$ then $N(T,d,\varepsilon)$ will denote the minimal number of closed balls of radius ε that covers T.

Proposition 1 *Let $\mathcal{C}$ be a class of subsets of I and $(F, \|.\|)$ be a fixed separable Banach space. Then the following conditions are equivalent*

a) *Exists $K_1 < \infty$ such that for any sequence (X_i) of independent symmetric r.v. with values in F satisfying $\#\{i : X_i \neq 0\} < \infty$ a.s.*

$$\forall_{t>0}\ P(\max_{C\in\mathcal{C}} \|\sum_{i\in C} X_i\| \geq K_1 t) \leq K_1 P(\|\sum_{i\in I} X_i\| \geq t).$$

b) *Exists $K_2 < \infty$ such that for any sequence (X_i) of independent symmetric r.v. with values in F satisfying $\#\{i : X_i \neq 0\} < \infty$ a.s.*

$$E\max_{C\in\mathcal{C}} \|\sum_{i\in C} X_i\| \leq K_2 E\|\sum_{i\in I} X_i\|.$$

c) *Exists $K_3 < \infty$ such that for any sequence (v_i) of vectors in F satisfying $\#\{i : v_i \neq 0\} < \infty$*

$$\forall_{t>0}\ P(\max_{C\in\mathcal{C}} \|\sum_{i\in C} v_i\varepsilon_i\| \geq K_3 t) \leq K_3 P(\|\sum_{i\in I} v_i\varepsilon_i\| \geq t).$$

d) *Exists $K_4 < \infty$ such that for any sequence (v_i) of vectors in F satisfying $\#\{i : v_i \neq 0\} < \infty$*

$$E \max_{C\in\mathcal{C}} \|\sum_{i\in C} v_i\varepsilon_i\| \leq K_4 E\|\sum_{i\in I} v_i\varepsilon_i\|.$$

e) *Exists $K_5 < \infty$ such that for any sequence (X_i) of independent r.v. with values in F satisfying $\#\{i : X_i \neq 0\} < \infty$ a.s.*

$$\forall_{t>0}\ P(\max_{C\in\mathcal{C}} \|\sum_{i\in C} X_i\| \geq K_5 t) \leq K_5 \max_{C\in\mathcal{C}\cup\{I\}} P(\|\sum_{i\in C} X_i\| \geq t).$$

f) *Exists $K_6 < \infty$ such that for any sequence (X_i) of independent r.v. with values in F satisfying $\#\{i : X_i \neq 0\} < \infty$ a.s.*

$$E \max_{C\in\mathcal{C}} \|\sum_{i\in C} X_i\| \leq K_6 \max_{C\in\mathcal{C}\cup\{I\}} E\|\sum_{i\in C} X_i\|.$$

Proof. Implications a)⇒c), b)⇒d) and c)⇒d) are obvious. By Fubini Theorem easily follows that c)⇒a) and d)⇒b). Moreover for symmetric r.v. and $C \subset I$ we have $P(\|\sum_{i\in C} X_i\| \geq t) \leq 2P(\|\sum_{i\in I} X_i\| \geq t)$ and $E\|\sum_{i\in C} X_i\| \leq E\|\sum_{i\in I} X_i\|$, so e)⇒a) and f)⇒b). Thus to prove Proposition 1 it is enough to show that d)⇒c), a)⇒e) and b)⇒f).

d)⇒c). By the results of [DM] it follows that there exists absolute constant $K < \infty$ such that for any sequence of vectors w_i in some Banach space E we have

$$\begin{aligned} &P(\|\sum \varepsilon_i w_i\| \geq K(E\|\sum \varepsilon_i w_i\| + t)) \\ &\qquad \leq K \max\{P(w^*(\sum \varepsilon_i w_i) \geq t) : w^* \in \mathrm{Ext}(B_{E^*})\}, \end{aligned} \tag{3}$$

where $\mathrm{Ext}(B_{E^*})$ denotes the set of extremal points in the unit ball of the dual space E^*.

Let us notice that $\max_{C\in\mathcal{C}} \|\sum_{i\in C} \varepsilon_i v_i\| = \|\sum_{i\in I} \varepsilon_i w_i\|_E$ for a suitable choice of $w_i \in E := l^\infty(\mathcal{C}; F)$. Hence (3) implies that

$$P(\max_{C\in\mathcal{C}} \|\sum_{i\in C} \varepsilon_i v_i\| \geq K(E \max_{C\in\mathcal{C}} \|\sum_{i\in C} \varepsilon_i v_i\| + t)) \leq KP(\|\sum_{i\in I} \varepsilon_i v_i\| \geq t). \tag{4}$$

We will show that c) holds for $K_3 = \max(8, (2K_4+1)K)$. If $t \geq \frac{1}{2}E\|\sum_{i\in I}\varepsilon_i v_i\|$ then by d) $E\max_{C\in\mathcal{C}}\|\sum_{i\in C}\varepsilon_i v_i\| + t \leq (2K_4+1)t$ and c) follows by (4). For $t \leq \frac{1}{2}E\|\sum_{i\in I}\varepsilon_i v_i\|$, by Paley-Zygmund inequality (see [Ka], p.8) we get

$$P(\|\sum_{i\in I}\varepsilon_i v_i\| \geq t) \geq \frac{1}{4}\frac{(E\|\sum_{i\in I}\varepsilon_i v_i\|)^2}{E\|\sum_{i\in I}\varepsilon_i v_i\|^2} \geq \frac{1}{8}$$

and the inequality in c) is obvious.

a)⇒e) and b)⇒f). Let X_i' be an independent copy of X_i, then the variables $X_i - X_i'$ are symmetric, $E\|\sum_{i\in I}(X_i - X_i')\| \leq 2E\|\sum_{i\in I}X_i\|$ and $P(\|\sum_{i\in I}(X_i - X_i')\| \geq 2t) \leq 2P(\|\sum_{i\in I}X_i\| \geq t)$. Thus both implications are simple consequences of the following lemma

Lemma 1 *If* $\max_{C\in\mathcal{C}} P(\|\sum_{i\in C}X_i\| \geq t/2) \leq 1/2$ *then*

$$P(\max_{C\in\mathcal{C}}\|\sum_{i\in C}X_i\| \geq t) \leq 2P(\max_{C\in\mathcal{C}}\|\sum_{i\in C}(X_i - X_i')\| \geq \frac{t}{2}).$$

Proof of Lemma 1. Suppose that $\mathcal{C} = \{C_1, C_2, \ldots\}$ and for simplifying the notation let $Y_k = \sum_{i\in C_k}X_i, Y_k' = \sum_{i\in C_k}X_i'$ for $k = 1, 2, \ldots$. We have

$$P(\|Y_k\| \geq t, \|Y_i\| < t \text{ for } i < k, \max_k\|Y_k - Y_k'\| \leq \frac{t}{2})$$

$$\leq P(\|Y_k\| \geq t, \|Y_i\| < t \text{ for } i < k, \|Y_k'\| \geq \frac{t}{2})$$

$$\leq P(\|Y_k\| \geq t, \|Y_i\| < t \text{ for } i < k)\max_k P(\|Y_k\| \geq \frac{t}{2}).$$

Hence summing the above inequalities over k we get

$$P(\max_k\|Y_k\| \geq t) \leq P(\max_k\|Y_k - Y_k'\| \geq \frac{t}{2}) + P(\max_k\|Y_k\| \geq t)\max_k P(\|Y_k\| \geq \frac{t}{2})$$

and Lemma 1 follows.

Definition 1 In the sequel we will say that the class $\mathcal{C}$ of subsets of I *satisfies the maximal inequality in* F if any of conditions a)-f) of Proposition 1 holds true. If this is true for any separable Banach space F we will say that $\mathcal{C}$ *satisfies the maximal inequality* or that it is the *MI-class.*

It is therefore of interest to solve the following

Main Problem. Determine all classes $\mathcal{C}$ that satisfy the maximal inequality.

It has turned out that the following definition plays the crucial role for this problem

Definition 2. We say that a class $\mathcal{C}$ of subsets of I *shatters* the set $A \subset I$ if

$$\{A \cap C : C \in \mathcal{C}\} = 2^A.$$

A class $\mathcal{C}$ is called a *Vapnik-Chervonenkis class* (or in short a *VC class*) *of order* n if it does not shatter any set of cardinality $n+1$ and it shatters some set of cardinality n. A class will be called a *VC class* if it is a VC class of some finite order.

For some properties and examples of VC classes see e.g. [D1, D2, SY].

Proposition 2 *If $\mathcal{C}$ satifies the maximal inequality in some Banach space F then $\mathcal{C}$ is a VC class.*

Proof. Obviously it is enough to prove Proposition for $F = I\!R$. Suppose that $\mathcal{C}$ shatters the set $A \subset I$ of cardinality n. Let

$$v_i = \begin{cases} 1 & \text{for} \quad i \in A \\ 0 & \text{for} \quad i \in I \setminus A \end{cases}.$$

Then

$$E|\sum \varepsilon_i v_i| \le (E|\sum \varepsilon_i v_i|^2)^{1/2} = \sqrt{n}$$

and

$$E \max_{C \in \mathcal{C}} |\sum_{i \in C} \varepsilon_i v_i| = E \max_{B \subset A} |\sum_{i \in B} \varepsilon_i|$$

$$= E \max(\#\{i \in A : \varepsilon_i = 1\}, \#\{i \in A : \varepsilon_i = -1\}) \ge \frac{n}{2}.$$

Therefore if condition d) of Proposition 1 is satisfied then $\mathcal{C}$ does not shatter any set of cardinality $> 4K_4^2$.

Theorem 1 *The class $\mathcal{C}$ of subsets of I satisfies the maximal inequality in $I\!R$ if and only if $\mathcal{C}$ is a VC class.*

In the proof of this theorem we will use the following two results of Dudley (see [LT], Theorems 11.1 and 14.12)

Theorem A *Let $\psi_2(x) = e^{x^2} - 1$ and (X_t) be a random process on (T, d) such that*

$$E\psi_2(|X_t - X_s|/d(t,s)) \le 1 \text{ for any } t, s \in T.$$

Then

$$E \sup_{s,t\in T} |X_t - X_s| \le 12 \int_0^\infty \ln^{1/2} N(T, d, \varepsilon) d\varepsilon.$$

Theorem B *Let Q be a probability measure on I and $d_Q(A, B) = (Q(A \div B))^{1/2}$ for $A, B \subset I$. Then for any VC class $\mathcal{C}$ of order $\le n$ and $\varepsilon \in (0, 1)$*

$$\ln N(\mathcal{C}, d_Q, \varepsilon) \le K_B n(1 - \ln \varepsilon),$$

where K_B is an absolute constant.

Proof of Theorem 1. One implication follows by Proposition 2. To prove the second, assume that $\mathcal{C}$ is a VC class of order $\le n$ and we will prove the condition d) of Proposition 1. We may also assume that $\emptyset \in \mathcal{C}$. Let v_i be fixed real numbers with $\sum v_i^2 = 1$ and $X_A = \sum_{i\in A} \varepsilon_i v_i$ for $A \subset I$. Let us also define the probability measure Q on I by the formula $Q(A) = \sum_{i\in A} v_i^2$ and a distance d on $\mathcal{C}$ by $d(A, B) = (Q(A \div B))^{1/2}$. Then $N(\mathcal{C}, d, \varepsilon) = 1$ for $\varepsilon > 1$. By the properties of Rademacher sums (see [LT], sect.4.1) there exists universal constant K such that $\|X_A\|_{\psi_2} \le K(\sum_{i\in A} v_i^2)^{1/2}$, so $E\psi_2((X_A - X_B)/Kd(A, B)) \le 1$. Therefore by Theorem A and B

$$E \sup_{C\in\mathcal{C}} |\sum_{i\in C} \varepsilon_i v_i| \le E \sup_{C,C'\in\mathcal{C}} |X_C - X_{C'}| \le 12 \int_0^\infty \ln^{1/2} N(\mathcal{C}, d, \varepsilon/K) d\varepsilon$$

$$\le 12\sqrt{K_B}\sqrt{n} \int_0^K (1 - \ln \varepsilon + \ln K)^{1/2} = \tilde{K}\sqrt{n}.$$

The Theorem 1 follows if we notice that $E|\sum_{i\in I} \varepsilon_i v_i| \ge (\sum v_i^2)^{1/2}/\sqrt{2}$ by Khinchine inequality.

Theorem 1 and Proposition 2 suggest that the following conjecture is reasonable

Conjecture. A class $\mathcal{C}$ satisfies the maximal inequality if and only if $\mathcal{C}$ is a VC class.

Using Theorem 1 and Talagrands majorizing measure theorem, L. Krawczyk proved in [Kr] that if $\mathcal{C}$ is a VC class then conditions (a) and (b) holds if we additionaly assume that X_i are Gaussian vectors. This was slightly generalized in [L] to the following Theorem.

Theorem 2 *Let $(X_i)_{i\in I}$ be a sequence of symmetric real random variables with logarithmically concave tails i.e. such that the functions $N_i(t) = -\ln P(|X_i| > t)$ are convex on $[0,\infty)$ and such that*

$$\forall_{t>0}\ N_i(2t) \leq AN_i(t)$$

for some constant $A < \infty$. Then for any VC class $\mathcal{C}$ of subsets of I of order $\leq n$ there exists a constant K, which depends only on A and n such that for any sequence of vectors v_i in some Banach space for which the sum $\sum v_iX_i$ is a.e. convergent, the following inequality holds

$$E\sup_{C\in\mathcal{C}} \|\sum_{i\in C} v_iX_i\| \leq KE\|\sum_{i\in I} v_iX_i\|.$$

Remark. Using concentration properties of logconcave measures one may prove in the similar way as in the proof of implication d)⇒c) of Proposition 1 that under the assumptions of Theorem 2

$$P(\sup_{C\in\mathcal{C}} \|\sum_{i\in C} v_iX_i\| \geq \tilde{K}t) \leq \tilde{K}P(\|\sum_{i\in I} v_iX_i\| \geq t)$$

for any $t > 0$, where $\tilde{K}$ is a constant depending only on A and n.

Corollary 1 *Let F be a separable Banach space with finite cotype. Then every VC class $\mathcal{C}$ satisfies the maximal inequality in F.*

Proof. Let $v_i \in F$ be as in condition (d). Then since F has finite cotype

$$E\|\sum_{i\in I} v_ig_i\| \leq AE\|\sum_{i\in I} v_i\varepsilon_i\|,$$

where A is a constant depending only on F. By the contraction principle

$$E\max_{C\in\mathcal{C}}\|\sum_{i\in C} v_i\varepsilon_i\| \leq \sqrt{\frac{\pi}{2}} E\max_{C\in\mathcal{C}}\|\sum_{i\in C} v_i g_i\|$$

and condition (d) immediately follows by the result of Krawczyk.

Remark. Proofs of the result of Krawczyk and Theorem 2 are based on general theorems about geometric conditions equivalent to the boundedness of processes $(\sum t_i X_i)_{t\in T}$. For Rademacher processes an important conjecture (for some partial results see [T2]) states that if for some $T \subset l^2$, $E\sup_{t\in T}\sum\varepsilon_i t_i < \infty$ then $T \subset U + KB_1$ for some $K < \infty$, where B_1 denotes a ball in l^1 and U is such that $E\sup_{t\in U}\sum t_i g_i < \infty$. It is not hard to check that the above conjecture immediately implies our conjecture about VC classes.

Definition 3 Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be two families of subset of I. Then we may define the following families

$$\mathcal{C}_1^c = \{I \setminus C_1 : C_1 \in \mathcal{C}_1\}$$

$$\mathcal{C}_1 \wedge \mathcal{C}_2 = \{C_1 \cap C_2 : C_1 \in \mathcal{C}_1, C_2 \in \mathcal{C}_2\}$$

and

$$\mathcal{C}_1 \vee \mathcal{C}_2 = \{C_1 \cup C_2 : C_1 \in \mathcal{C}_1, C_2 \in \mathcal{C}_2\}.$$

Proposition 3 *Suppose that $\mathcal{C}_1$ and $\mathcal{C}_2$ are MI-classes. Then also the families $\mathcal{C}_1^c$, $\mathcal{C}_1 \wedge \mathcal{C}_2$ and $\mathcal{C}_1 \vee \mathcal{C}_2$ satisfy the maximal inequality.*

Proof. Since $\|\sum_{i\in I\setminus C} X_i\| \leq \|\sum_{i\in I} X_i\| + \|\sum_{i\in C} X_i\|$ by the triangle inequality, we immediately get that $\mathcal{C}_1^c$ is a MI-class. Moreover $\mathcal{C}_1 \vee \mathcal{C}_2 = (\mathcal{C}_1^c \wedge \mathcal{C}_2^c)^c$, so it is enough to prove that $\mathcal{C}_1 \wedge \mathcal{C}_2$ satisfies the maximal inequality. We will check the condition (b). Let $(F, \|.\|)$ be a given Banach space and $\tilde{F} = l^\infty(\mathcal{C}, F)$. Suppose that $\mathcal{C}_1$ and $\mathcal{C}_2$ satisfy (b) in F and $\tilde{F}$ respectively with constants K and $\tilde{K}$. Let $\tilde{X}_i$ be independent r.v. with values in $\tilde{F}$ defined by the formula

$$\tilde{X}_i(C) = \begin{cases} X_i & \text{for } i \in C \\ 0 & \text{for } i \notin C. \end{cases}$$

Then for $A \subset I$, we have

$$\|\sum_{i\in A}\tilde{X}_i\|_{\tilde{F}} = \sup_{C_1\in\mathcal{C}_1}\|\sum_{i\in A\cap C_1}X_i\|.$$

Thus

$$E\max_{C\in\mathcal{C}_1\wedge\mathcal{C}_2}\|\sum_{i\in C}X_i\| = E\max_{C_2\in\mathcal{C}_2}\|\sum_{i\in C_2}\tilde{X}_i\|_{\tilde{F}} \le \tilde{K}E\|\sum_{i\in I}\tilde{X}_i\|_{\tilde{F}}$$
$$= \tilde{K}E\max_{C_1\in\mathcal{C}_1}\|\sum_{i\in C_1}X_i\| \le K\tilde{K}E\|\sum_{i\in I}X_i\|.$$

Proposition 4 *Every VC class of order 1 satisfies the maximal inequality.*

Proof. Following the notation of [S] we will call the family $\mathcal{F}$ of subsets of I a chain if it is linearly ordered by the inclusion, i.e. for each $A, B \in \mathcal{F}$ either $A \subset B$ or $B \subset A$. Families of the form $\mathcal{F}_1 \wedge \mathcal{F}_2$, where $\mathcal{F}_1$ and $\mathcal{F}_2$ are chains, will be called 2-chains. Smoktunowicz in [S] proved that if $\mathcal{C}$ is a VC class of order 1 then $\mathcal{C} \subset \mathcal{G}_1 \vee \mathcal{G}_2^c$ for some 2-chains $\mathcal{G}_1$ and $\mathcal{G}_2$. Since every chain is a MI-class by Levy inequality, Proposition 4 follows by Proposition 3.

By Propositions 3 and 4 we immediately get the following

Corollary 2 *Suppose that $\mathcal{C}'$ is a family of subsets of I that is obtained from some VC classes of order 1 by finitely many operations ^{c}, $\vee$ and $\wedge$. Then any subfamily $\mathcal{C} \subset \mathcal{C}'$ satisfies the maximal inequality.*

Unfortunately even very simple VC classes are not of the form described in the above Corollary. A. Smoktunowicz in [S] showed that the family of all lines in Z^2 does not have such form. As follows from the below Corollary 4 the lattest family is an MI-class, so Corollary 2 does not describe all families that satisfy the maximal inequality.

Definition 4 In the last part of the paper we will consider the classes $\mathcal{C}$ of subsets of I, which satisfy the additional condition

$$\forall_{A,B\in\mathcal{C}}\ A \neq B \Rightarrow \#(A\cap B) \le 1. \tag{5}$$

Such classes will be called *1-disjoint.*

We will also denote for a fixed sequence of vectors v_i and $A \subset I$ by X_A the variable $\sum_{i\in A} v_i\varepsilon_i$.

Lemma 2 *If $M \geq 2\max_i \|v_i\|$, class $\mathcal{C}$ is 1-disjoint and $A_1, \ldots, A_n \in \mathcal{C}$ are such that for some $t > 0$*

$$P(\|X_{A_k}\| \geq M) \geq t \quad \text{for } k = 1, \ldots, n,$$

then there exist disjoint subsets $B_1, \ldots, B_m \subset I$ such that $m \geq \sqrt[3]{n}$ and

$$P(\|X_{B_k}\| \geq M/2) \geq t/2 \quad \text{for } k = 1, \ldots, m.$$

Proof. In this proof we will say that the set A_k is *good* if

$$\forall_{C \subset A_k} \ \#C \leq \sqrt[3]{n} \Rightarrow P(\|X_C\| \geq M/2) \leq t/2. \tag{6}$$

Let us notice that the last condition also implies by the triangle inequality that $P(\|X_{A_k \setminus C}\| \geq M/2) \geq t/2$. We will consider 3 cases

Case 1. Among $A_1, \ldots, A_n$ there are $m \geq \sqrt[3]{n}$ good sets, say $A_1, \ldots, A_m$. Without loss of generality we may assume that $m < \sqrt[3]{n} + 1$. If we put $B_1 = A_1$ and $B_i = A_i \setminus (\bigcup_{j<i} A_j)$ for $1 < i \leq m$ we get by (5) that $\#(A_i \setminus B_i) \leq i - 1 \leq \sqrt[3]{n}$. Hence we get the thesis in this case by the definition of good sets.

Case 2. There exists $i \in I$ such that $\#\{k : i \in A_k\} \geq \sqrt[3]{n}$. Without loss of generality we may assume that $i \in A_1 \cap \ldots \cap A_m$ with $m \geq \sqrt[3]{n}$. We put in this case $B_k = A_k \setminus \{i\}$ and notice that $\|X_{B_k}\| \geq \|X_{A_k}\| - \|v_i\| \geq \|X_{A_k}\| - M/2$. Sets B_k are disjoint by the property (5).

Case 3. There are less then $\sqrt[3]{n}$ good sets A_k and $\#\{k : i \in A_k\} < \sqrt[3]{n}$ for all $i \in I$. We have more then $n - \sqrt[3]{n}$ not good sets, let A_{i_1} be one of them. We may then find $B_1 \subset A_{i_1}$ with $\#B_1 \leq \sqrt[3]{n}$ and $P(\|X_{B_1}\| \geq M/2) \geq t/2$. At most $\sqrt[3]{n}\#B_1$ sets A_i have nonempty intersection with B_1. So we have more then $n - \sqrt[3]{n} - \sqrt[3]{n^2}$ not good sets disjoint with B_1, let A_{i_2} be one of them. Then we may find $B_2 \subset A_{i_2}$ with $\#B_2 \leq \sqrt[3]{n}$ and $P(\|X_{B_2}\| \geq M/2) \geq t/2$. Continuing in this way completes the proof.

Corollary 3 *Suppose that $M \geq 8E\|X_I\|$ and class $\mathcal{C}$ is 1-disjoint, then*

$$\sum_{A \in \mathcal{C}} (P(\|X_A\| \geq M))^4 \leq 2^{14} P(\|X_I\| \geq M/2).$$

Proof. Suppose that there exist $A_1, \ldots, A_n \in \mathcal{C}$ such that $P(\|X_{A_k}\| \geq M) \geq t$ for all k. Then by Lemma 1 we may find disjoint subsets $B_1, \ldots, B_m \subset I$ with $m \geq \sqrt[3]{n}$ and $P(\|X_{B_k}\| \geq M/2) \geq t/2$. But by Levy inequality

$$P(\max \|X_{B_k}\| \geq M/2) \leq 2P(\|X_I\| \geq M/2) \leq 1/2,$$

so

$$t\sqrt[3]{n}/2 \leq \sum P(\|X_{B_k}\| \geq M/2) \leq 4P(\|X_I\| \geq M/2)$$

and $\sqrt[3]{n} \leq 8P(\|X_I\| \geq M/2)/t$. Therefore we obtain

$$\sum_{A\in\mathcal{C}}(P(\|X_A\| \geq M))^4 \leq 16\sum_{n=1}^{\infty} 2^{-4n}\#\{A \in \mathcal{C} : P(\|X_A\| \geq M) \geq 2^{-n}\}$$

$$\leq 2^{13}P(\|X_I\| \geq M/2)\sum_{n=1}^{\infty} 2^{-4n}2^{3n} \leq 2^{14}P(\|X_I\| \geq M/2).$$

Corollary 4 *There exists a universal constant C such that for any 1-disjoint class $\mathcal{C}$*

$$\sum_{A\in\mathcal{C}} E\|X_A\|I_{\{\|X_A\|\geq CE\|X_I\|\}} \leq CE\|X_I\|.$$

In particular

$$E\max_{A\in\mathcal{C}}\|X_A\| \leq 2CE\|X_I\|,$$

so the maximal inequality holds for any VC class satisfying (5).

Proof. By the properties of Rademacher sums (cf [Ka]) we have

$$P(\|X_A\| \geq 4M) \leq C_1(P(\|X_A\| \geq M))^4$$

for some constant $C_1 < \infty$. Therefore by the previous Corollary we obtain for $M \geq 8E\|X_I\|$

$$\sum_{A\in\mathcal{C}} P(\|X_A\| \geq 4M) \leq 2^{14}C_1P(\|X_I\| \geq M/2).$$

Corollary follows by integration the above inequality with respect to M.

Remark. All the above results remain true (with a change of constants) if we substitute (5) by the more general condition

$$\forall_{A,B\in\mathcal{C}}\ A \neq B \Rightarrow \#(A\cap B) \leq m,$$

where m is a fixed positive integer.

Acknowledgments. The main part of this research was carried out when the author was working in the Institute of Mathematics Polish Academy of Science. The author would like to thank Prof. Stanisław Kwapień for intruducing him into the subject and many useful discussion and suggestions.

References

[DM] S. J. Dilworth and S. J. Montgomery-Smith, *The distribution of vector-valued Rademacher series*, Ann. Prob. 21 (1993), 2046-2052.

[D1] R. M. Dudley, *A course on empirical processes*, École d'été de probabilités de Saint-Flour, XII–1982, 1-142, Lecture Notes in Math. 1097, Springer, Berlin-New York, 1984

[D2] R. M. Dudley, *The structure of some Vapnik-Červonenkis classes*, Proceedings of the Berkeley conference in honor of Jerzy Neyman and Jack Kiefer (Berkeley, Calif. 1983), 495-508, Wadsworth, Belmont, Calif., 1985

[Ka] J.-P. Kahane, *Some random series of functions*, Cambridge University Press, 1985, 2nd ed.

[Kr] L. Krawczyk *Maximal Inequality for Gaussian Vectors*, Bull. Acad. Polon. Sci. Math. 44 (1996), 157-160.

[KW] S. Kwapień, W. A. Woyczyński, *Random Series and Stochastic Integrals: Single and Multiple*, Birkhauser, Boston 1992.

[L] R. Latała *Sudakov minoration principle and supremum of some processes,* Geom. and Funct. Anal. 7 (1997), 936-953.

[LT] M. Ledoux, M. Talagrand, *Probability in Banach Spaces*, Springer, Berlin Heidelberg 1991.

[S] A. Smoktunowicz, *A remark on Vapnik-Chervonienkis classes*, Colloq. Math. 74 (1997), 93-98.

[SY] G. Stengle, J. E. Yukich *Some new Vapnik-Chervonenkis classes*, Ann. Statist. 17 (1989), 1441-1446.

[T1] M. Talagrand, *Regularity of Gaussian processes*, Acta Math. 159 (1987), 99-149.

[T2] M. Talagrand, *Construction of majorizing measures, Bernoulli processes and cotype*, Geom. and Funct. Anal. 4 (1994), 660-717.

Institute of Mathematics
Warsaw University
Banacha 2
02-097 Warszawa, Poland
E-mail:rlatala@mimuw.edu.pl