PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 0, Fasc. 0 (0000), pp. 000–000
doi:10.19195/0208-4147.0.0.0

DOOB'S ESTIMATE FOR COHERENT RANDOM VARIABLES AND MAXIMAL OPERATORS ON TREES

BY

STANISŁAW CICHOMSKI (WARSAW) AND ADAM OSĘKOWSKI (WARSAW)

Abstract. Let ξ be an integrable random variable defined on $(\Omega, \mathcal{F}, \mathbb{P})$. Fix $k \in \mathbb{Z}_+$ and let $\{\mathcal{G}_i^j\}_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant k}$ be a reference family of sub-σ-fields of $\mathcal{F}$, such that $\{\mathcal{G}_i^j\}_{1 \leqslant i \leqslant n}$ is a filtration for each $j \in \{1, 2, \ldots, k\}$. In this article we explain the underlying connection between the analysis of the maximal functions of the corresponding coherent vector and basic combinatorial properties of the uncentered Hardy–Littlewood maximal operator. Following a classical approach of Grafakos, Kinnunen and Montgomery-Smith, we establish an appropriate version of the celebrated Doob's maximal estimate.

2010 AMS Mathematics Subject Classification: Primary: 60E15, 60G42.

Key words and phrases: Coherent distribution; maximal operator; martingale; best constants.

1. INTRODUCTION

The inspiration for the results obtained in this paper comes from the recent developments in the theory of coherent distributions. To introduce the necessary notions, suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is an arbitrary nonatomic probability space. Following [3], we say that a random vector $X = (X_1, X_2, \ldots, X_n)$ is coherent, if there exist a random variable ξ taking values in $\{0, 1\}$ and a sequence $\mathcal{G} = (\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_n)$ of sub-σ-algebras of $\mathcal{F}$ such that $X_k = \mathbb{E}(\xi|\mathcal{G}_k)$ for all $k = 1, 2, \ldots, n$. The motivation for this definition lies from economics, where coherent distributions are used to model the behavior of agents with partially overlapping information sources [1], [10]. From the mathematical point of view, such random vectors enjoy many interesting structural properties; for some latest theoretical advances on this subject, see e.g. [2], [6], [7]. In this article, we will be interested in the universal sharp norm comparison of ξ and the maximal function of X. We will drop the assumption $\mathbb{P}(\xi \in \{0, 1\}) = 1$ and work with arbitrary integrable random variables. For such a ξ and a sequence $\mathcal{G}$, the associated maximal function is given by $M_{\mathcal{G}}\xi = \sup_j |\mathbb{E}(\xi|\mathcal{G}_j)|$. The starting point is the classical result of Doob, which asserts that

$$\|M_{\mathcal{G}}\xi\|_p \leqslant \frac{p}{p-1}\|\xi\|_p, \qquad 1 < p \leqslant \infty, \tag{1.1}$$

in the case when $\mathcal{G}$ is a filtration, i.e., we have the nesting condition $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \ldots \subseteq \mathcal{G}_n$. Furthermore, for each p the number $p/(p-1)$ is the best universal constant (i.e., not depending on the length of $\mathcal{G}$) allowed in the estimate. The main goal of this paper is to consider (1.1) for more general families of σ-algebras: we will assume that $\mathcal{G}$ can be decomposed into the union of filtrations. Specifically, we let $\mathcal{G}$ be of the form

$$\mathcal{G} := \{\mathcal{G}_i^j\}_{\substack{1\leqslant i\leqslant n,\\ 1\leqslant j\leqslant k}},$$

and require the inclusions $\mathcal{G}_1^j \subseteq \mathcal{G}_2^j \subseteq \ldots \subseteq \mathcal{G}_n^j$ for each j. No relation between σ-algebras $\mathcal{G}_i^j$ with different j is imposed. Thus, our investigation can be seen as a natural halfway state between the study of general coherent distributions and classical martingales. Furthermore, this subject enters into the still vague framework of martingales indexed by partially ordered sets. For a general introduction to this theory see [12], for related Doob's type inequalities see [4], [5], [11], [13]. Our reasoning will reveal an unexpected connection between the analysis of $\max_{i,j} |\mathbb{E}(\xi|\mathcal{G}_i^j)|$ and basic combinatorial properties of the uncentered Hardy–Littlewood maximal operator on tree-shaped domains. Due to this interdependence, we will be able to extend the classical approach introduced in [8], [9] and derive an appropriate sharp version of (1.1).

THEOREM 1.1. *Let $1 < p < \infty$ be a given parameter and assume that $\mathcal{G} = \{\mathcal{G}_i^j\}_{1\leqslant i\leqslant n, 1\leqslant j\leqslant k}$ is the union of filtrations as above. Then for any random variable $\xi \in L^p$ we have the estimate*

$$\|M_{\mathcal{G}}\xi\|_p \leqslant C_{p,k}\|\xi\|_p, \tag{1.2}$$

where $C_{p,k}$ is the unique root of the equation

$$(p-1)C_{p,k}^p - pC_{p,k}^{p-1} - (k-1) = 0. \tag{1.3}$$

For fixed $1 < p < \infty$ and $k \geqslant 1$, the constant $C_{p,k}$ is the best possible: given $\varepsilon > 0$, there is an integer n, a family $\mathcal{G}$ as above and a positive random variable $\xi \in L^p$ for which

$$\|M_{\mathcal{G}}\xi\|_p > (C_{p,k} - \varepsilon)\|\xi\|_p. \tag{1.4}$$

That is, the constant $C_{p,k}$ is the best universal constant allowed in (1.2), where the universality is the non-dependence on n, the length of the filtrations building $\mathcal{G}$. We would like to point out that the constant $C_{p,k}$ is still optimal if we restrict ourselves to random variables ξ taking values in $[0,1]$. This follows by a simple approximation argument: given a positive almost extremal variable ξ (i.e., satisfying (1.4)), we replace it with $\min\{\xi, L\}$, where L is a positive constant. If L is sufficiently large, then this new variable still satisfies (1.4), and hence so does $\min\{\xi, L\}/L$, by homogeneity. It remains to note that the latter variable takes values in $[0, 1]$.

Interestingly, in the case $\xi \in \{0, 1\}$, which originates in the coherent context, the optimal constant is smaller: here is the precise formulation.

THEOREM 1.2. *Let* $\mathcal{G} = \{\mathcal{G}_i^j\}_{1\leqslant i\leqslant n, 1\leqslant j\leqslant k}$ *be the union of filtrations as above and let* $1 < p < \infty$. *Then for any random variable* ξ *with values in* $\{0, 1\}$ *we have*

$$\|M_{\mathcal{G}}\xi\|_p \leqslant \left(1 + \frac{k}{p-1}\right)^{1/p} \|\xi\|_p. \tag{1.5}$$

The constant is the best possible for each k *and each* p.

We turn our attention to the analytic contents of the paper. Let k be a fixed positive integer. Consider the set $\mathcal{R}_k = \bigcup_{j=1}^k H_j$, where H_j is the line segment on the complex plane, with endpoints 0 and $e^{2\pi ij/k}$, $j = 1, 2, \ldots, k$. That is, $\mathcal{R}_k$ is a tree-shaped domain being the union of k rays $H_1, H_2, \ldots, H_k$, each having length one. We equip $\mathcal{R}_k$ with the standard British railway metric and the normalized one-dimensional Lebesgue measure λ_k. Then we can introduce the concept of the decreasing rearrangement on $\mathcal{R}_k$. Namely, for an arbitrary random variable ξ on $(\Omega, \mathcal{F}, \mathbb{P})$, we define first its distribution function $d_\xi : [0, \infty) \to [0, 1]$ by $d_\xi(s) = \mathbb{P}(|\xi| > s)$. Then the associated k-decreasing rearrangement $\xi^*_{(k)} : \mathcal{R}_k \to [0, \infty)$ is given by

$$\xi^*_{(k)}(e^{2\pi ij/k}t) = \inf\{s > 0 \,:\, d_\xi(s) \leqslant t\}, \qquad j = 1, 2, \ldots, k.$$

Equivalently, $\xi^*_{(k)}$ can be defined by taking the standard decreasing rearrangement ξ^* on $[0, 1]$ and copying it on each ray H_j, in accordance with the natural order induced by the distance from 0. Thus, we immediately see that $|\xi|$ and $\xi^*_{(k)}$ have the same distributions (as random variables on Ω and $\mathcal{R}_k$, respectively). Furthermore, $\xi^*_{(k)}$ is radially decreasing, i.e., $\xi^*_{(k)}(x) = \xi^*_{(k)}(|x|)$ decreases as $|x|$ grows.

Finally, we introduce the uncentered Hardy-Littlewood maximal function $\mathcal{M}_{(k)}$ in the above setup. This operator acts on integrable functions f on $\mathcal{R}_k$ by the usual formula

$$\mathcal{M}_{(k)}f(x) = \sup \frac{1}{\lambda_k(B)} \int_B |f| \mathrm{d}\lambda_k, \qquad x \in \mathcal{R}_k,$$

where the supremum is taken over all open balls $B \subseteq \mathcal{R}_k$ which contain x. We will identify the L^p norm of this object.

THEOREM 1.3. *For any* $1 < p < \infty$ *and any* $k \geqslant 2$ *we have* $\|\mathcal{M}_{(k)}\|_{L^p \to L^p} = C_{p,k}$, *where* $C_{p,k}$ *is given in* (1.3).

The case $k = 2$ was established by Grafakos and Montgomery-Smith [9]. Our contribution is the analysis for $k \geqslant 3$. Furthermore, we will link the context of coherent distributions with the analytic setup above, intertwining the contents of Theorems 1.1 and 1.3.

THEOREM 1.4. *Let $k, n \geqslant 1$ be fixed integers. Suppose further that ξ is an integrable random variable and assume that $\mathcal{G} = \{\mathcal{G}_i^j\}_{1\leqslant i\leqslant n, 1\leqslant j\leqslant k}$ is a union of filtrations as above. Then the maximal function $M_{\mathcal{G}}\xi$ satisfies the majorization*

$$(M_{\mathcal{G}}\xi)^*_{(k)} \leqslant \mathcal{M}_{(k)}(\xi^*_{(k)}) \qquad \lambda_k\text{-almost everywhere on } \mathcal{R}_k. \tag{1.6}$$

The remaining part of the paper is split into two sections. In Section 2 we establish Theorem 1.4. In the last part of the paper, we establish the L^p bound $\|\mathcal{M}_{(k)}\|_{L^p\to L^p} \leqslant C_{p,k}$, which allows us to deduce (1.2) immediately. Furthermore, we show there the sharpness of the latter inequality, thus completing the proofs of all aforementioned results.

From now on, the parameter k will be kept fixed; to simplify the notation, we will skip the index and write ξ^*, $\mathcal{M}$ instead of $\xi^*_{(k)}$ and $\mathcal{M}_{(k)}$, respectively.

2. PROOF OF THEOREM 1.4

We will need the following property of the Hardy-Littlewood maximal operator.

LEMMA 2.1. *Suppose that ξ is an integrable random variable. Then for any $s > 0$ such that $\lambda_k(\mathcal{M}\xi^* > s) < 1$ we have*

$$\begin{aligned} s\big((k-1)\lambda_k(\xi^* > s) + \lambda_k(\mathcal{M}\xi^* > s)\big) \\ = (k-1)\int_{\{\xi^*>s\}} \xi^* d\lambda_k + \int_{\{\mathcal{M}\xi^*>s\}} \xi^* d\lambda_k. \end{aligned}$$

Proof. If $s \geqslant \|\xi\|_\infty$, then the assertion is evident (both sides are zero), so from now on we assume that $s < \|\xi\|_\infty$. The function $\mathcal{M}\xi^*$ is radially decreasing along the rays of $\mathcal{R}_k$. Furthermore, it is continuous, which follows directly from Lebesgue's dominated convergence theorem. Thus there exists $u \in \mathcal{R}_k$, lying on the ray H_1, for which $s = \mathcal{M}\xi^*(u)$. It is easy to identify the ball B for which the supremum defining $\mathcal{M}\xi^*(u)$ is attained: u must be one of its boundary points, and the intersection $B \cap H_j$ for $j \neq 1$ must be the part of H_j on which we have $f > s$. It remains to note that the equality

$$s = \mathcal{M}\xi^*(u) = \frac{1}{\lambda_k(B)}\int_B \xi^* \mathrm{d}\lambda_k$$

is equivalent to the claim. Indeed, we have $\lambda_k(B) = \frac{k-1}{k}\lambda_k(\xi^* > s) + \frac{1}{k}\lambda_k(\mathcal{M}\xi^* > s)$, with a similar identity for $\int_B \xi^* \mathrm{d}\lambda_k$. ∎

P r o o f o f T h e o r e m 1.4. It is enough to show the tail inequality

$$\mathbb{P}(M_{\mathcal{G}}\xi > s) \leqslant \lambda_k(\mathcal{M}\xi^* > s) \tag{2.1}$$

for all s. Now we consider two separate steps.

Step 1. Reductions. Let us first exclude the trivial cases: from now on, we will assume that $\lambda_k(\mathcal{M}\xi^* > s) < 1$ and $s < \|\xi\|_\infty$. Indeed, if $\lambda_k(\mathcal{M}\xi^* > s) = 1$, then there is nothing to prove, while for $s \geqslant \|\xi\|_\infty$ both sides of (2.1) are zero. Adding the full σ-algebras $\mathcal{G}_{n+1}^j = \mathcal{F}$, $j = 1, 2, \ldots, k$ to the collection $\mathcal{G}$ if necessary, we may and do assume that

$$\max_i |\mathbb{E}(\xi|\mathcal{G}_i^j)| \geqslant |\xi| \qquad \text{almost surely for all } j. \tag{2.2}$$

In particular, this gives $M_{\mathcal{G}}\xi \geqslant |\xi|$ with probability 1.

Step 2. Proof of theorem. Fix an arbitrary $s > 0$ and write

$$\mathbb{P}(M_{\mathcal{G}}\xi > s) = \mathbb{P}(A_1 \cup A_2 \cup \ldots \cup A_k),$$

where $A_j = \{\max_i |\mathbb{E}(\xi|\mathcal{G}_i^j)| > s\}$, $j = 1, 2, \ldots, k$. Let us distinguish the additional event $A_0 = \{|\xi| > s\}$ and observe that $A_0 \subseteq A_j$ for each j, in the light of (2.2). Note that if $\tilde{A}_j$ is an arbitrary event satisfying $A_0 \subseteq \tilde{A}_j \subseteq A_j$, then we have

$$s\mathbb{P}(\tilde{A}_j) - \int_{\tilde{A}_j} |\xi| \mathrm{d}\mathbb{P} = \int_{\tilde{A}_j} (s - |\xi|)\mathrm{d}\mathbb{P} \leqslant \int_{A_j} (s - |\xi|)\mathrm{d}\mathbb{P} \leqslant 0, \tag{2.3}$$

where the latter bound follows from Doob's weak-type bound for martingale maximal function. Next, we write

$$\begin{aligned} &\mathbb{P}(A_1 \cup A_2 \cup \ldots \cup A_k) \\ &= \mathbb{P}(A_0 \cup A_1 \cup A_2 \cup \ldots \cup A_k) \\ &= \mathbb{P}(A_0) + \mathbb{P}(A_1 \setminus A_0) + \mathbb{P}(A_2 \setminus (A_1 \cup A_0)) \\ &\quad + \ldots + \mathbb{P}(A_n \setminus (A_{n-1} \cup A_{n-2} \cup \ldots \cup A_0)). \end{aligned}$$

Set $\tilde{A}_j = A_0 \cup \big(A_j \setminus (A_{j-1} \cup A_{j-2} \cup \ldots \cup A_0)\big)$, apply (2.3) and add the estimates over j. Combining the result with the above formula for $\mathbb{P}(A_1 \cup A_2 \cup \ldots \cup A_k)$, we obtain

$$s\big[\mathbb{P}(A_1 \cup A_2 \cup \ldots \cup A_k) + (k-1)\mathbb{P}(A_0)\big] = s\sum_{j=1}^k \mathbb{P}(\tilde{A}_j) \leqslant \sum_{j=1}^k \int_{\tilde{A}_j} |\xi| \mathrm{d}\mathbb{P},$$

or equivalently,

$$s\big[\mathbb{P}(M_{\mathcal{G}}\xi > s) + (k-1)\mathbb{P}(A_0)\big] \leqslant \int_{\{M_{\mathcal{G}}\xi > s\}} |\xi| \mathrm{d}\mathbb{P} + (k-1)\int_{A_0} |\xi| \mathrm{d}\mathbb{P}.$$

Since $|\xi|$ and ξ^* are equidistributed, we have $\mathbb{P}(A_0) = \lambda_k(\xi^* > s)$ and $\int_{A_0} |\xi| \mathrm{d}\mathbb{P} = \int_{\{\xi^* > s\}} \xi^* \mathrm{d}\lambda_k$. Plugging this above and applying Lemma 2.1, we get

$$\int\limits_{\{M_{\mathcal{G}}\xi > s\}} (s - |\xi|)\mathrm{d}\mathbb{P} \leqslant \int\limits_{\{M\xi^* > s\}} (s - \xi^*)\mathrm{d}\lambda_k,$$

or, subtracting the equality $\int_{\{|\xi| > s\}} (s - |\xi|)\mathrm{d}\mathbb{P} = \int_{\{\xi^* > s\}} (s - \xi^*)\mathrm{d}\lambda_k$,

$$\int\limits_{\{M_{\mathcal{G}}\xi > s\}} (s - |\xi|)_+ \mathrm{d}\mathbb{P} \leqslant \int\limits_{\{\mathcal{M}\xi^* > s\}} (s - \xi^*)_+ \mathrm{d}\lambda_k = \int\limits_{\mathcal{R}_k} \chi_{\{\mathcal{M}\xi^* > s\}} (s - \xi^*)_+ \mathrm{d}\lambda_k.$$

However, the nonnegative functions $\chi_{\{\mathcal{M}\xi^* > s\}}$ and $(s - \xi^*)_+$ have the reversed monotonicity along the rays: the first of them is non-increasing, while the second is non-decreasing. Since $(s - |\xi|)_+$ and $(s - \xi^*)_+$ have the same distribution, (2.1) follows. ■

3. L^P ESTIMATES

We turn our attention to Theorems 1.1 and 1.3. Let us start with the L^p bound for the uncentered maximal operator; the key ingredient of the proof is the following weak-type estimate.

PROPOSITION 3.1. *For an arbitrary integrable function f on $\mathcal{R}_k$ and any $s > 0$ we have*

(3.1)

$$s\lambda_k(\mathcal{M}f > s) + s(k-1)\lambda_k(|f| > s) \leqslant \int\limits_{\{\mathcal{M}f > s\}} |f| d\lambda_k + (k-1) \int\limits_{\{|f| > s\}} |f| d\lambda_k.$$

P r o o f. It is convenient to split the reasoning into two steps.

Step 1. Special balls in $\mathcal{R}_k$. Let us consider the level set $E = \{x \in \mathcal{R}_k : \mathcal{M}f > s\}$. Then for each $x \in E$ there is an open ball $B_x \subseteq \mathcal{R}_k$ which contains x and satisfies $\lambda_k(B_x)^{-1} \int_{B_x} |f| \mathrm{d}\lambda_k > s$. This inequality implies that $B_x \subseteq E$ and hence $\bigcup_{x \in E} B_x = E$. By the Lindelöf's theorem, we may pick a countable subcollection $(B_{x_n})_{n=1}^\infty$ such that $\bigcup_{n=1}^\infty B_{x_n} = E$. With no loss of generality, we may assume that B_{x_i} is not a subset of B_{x_j} for $i \neq j$. We fix an integer N and restrict ourselves to the finite family $\mathcal{B} = (B_{x_n})_{n=1}^N$. The idea is to pick a subcollection $\mathcal{B}'$ of $\mathcal{B}$ which does not overlap too much. To this end, we will choose appropriate balls from each separate ray of $\mathcal{R}_k$, exploiting the natural order induced by the distance from 0. For simplicity, we will only describe the procedure for the k-th ray (i.e., for the interval $[0, 1]$), the argument for other rays is the same, up to rotation.

First, we pick a ball from $\mathcal{B}$ which contains zero and call it J_0 (if no ball in $\mathcal{B}$ contains zero, we let $J_0 = \emptyset$; if there are several balls with this property, we take the ball whose intersection with $[0, 1]$ has the biggest measure). Next we apply the following inductive procedure.

1° Suppose that we have successfully defined J_n. Consider the family of all intervals $J \in \mathcal{B}$ which intersect J_n and satisfy $\sup J > \sup J_n$. If this family is nonempty, choose the interval with largest left-endpoint (if this object is not unique, pick the one with the biggest measure) and denote it by J_{n+1}.

2° If the family in 1° is empty, then consider all intervals $J \in \mathcal{B}$ with $\inf J \geqslant \sup J_n$. If this family is nonempty, choose an element with the smallest left-endpoint (again, if this object is not unique, pick the one with the biggest measure) and denote it by J_{n+1}.

3° Go to 1°.

Since the family $\mathcal{B}$ is finite, the procedure stops after a number of sets (in 1° and 2°, there are no balls to choose from) and returns a family J_0^j, J_1^j, J_2^j, …, $J_{m_j}^j$ of balls. Observe that by the very construction, J_0^j, J_2^j, J_4^j, … are pairwise disjoint and the same is true for J_1^j, J_3^j, J_5^j, …. Letting

$$\mathcal{B}' = \{J_\ell^j \,:\, 1 \leqslant \ell \leqslant m_j,\, j = 1,\, 2,\, \ldots,\, k\},$$

we easily check that

$$\bigcup_{B \in \mathcal{B}} B = \bigcup_{B \in \mathcal{B}'} B. \tag{3.2}$$

Next, by the disjointness properties of the sequences J_i^j, note that a family $\mathcal{B}'$ has the following property: each point $x \in \mathcal{R}_k$ belongs to at most $k+1$ elements of $\mathcal{B}'$. Moreover, we can actually improve this last bound by 1. Now, say that there is a point $x_0 \in \mathcal{R}_k$ which belongs to exactly $k+1$ elements of $\mathcal{B}'$ and let us assume that x_0 belongs to the the k-th ray H_k. By the extremality of J_0^k we must have $(J_0^i \cap [0,1]) \subset (J_0^k \cap [0,1])$ for all $i = 1, 2, \ldots, k-1$, and hence

$$x_0 \in \bigcap_{j=1}^k J_0^j \cap J_1^k.$$

Thus, we simply remove J_0^k from the family $\mathcal{B}'$. Such a modification does not affect the validity of (3.2) and proves our assertion.

Step 2. Calculation. Since $\mathcal{B}' \subseteq \mathcal{B}$, each element B of $\mathcal{B}'$ satisfies

$$s\lambda_k(B) \leqslant \int_B |f| \mathrm{d}\lambda_k.$$

Summing over all $B \in \mathcal{B}'$, we thus obtain

$$s\left[\lambda\left(\bigcup_{B\in\mathcal{B}'} B\right) + \sum_{j=2}^k \lambda_k(A_j)\right] \leqslant \int_{\bigcup_{B\in\mathcal{B}'} B} |f|\mathrm{d}\lambda_k + \sum_{j=2}^k \int_{A_j} |f|\mathrm{d}\lambda_k,$$

where A_j is the collection of all $x \in \mathcal{R}_k$ which belong to exactly j elements of $\mathcal{B}'$. This is equivalent to

$$\begin{aligned} s\lambda\left(\bigcup_{B\in\mathcal{B}} B\right) &\leqslant \int\limits_{\bigcup_{B\in\mathcal{B}} B} |f|\mathrm{d}\lambda_k + \sum_{j=2}^{k} \int\limits_{A_j} (|f|-s)\mathrm{d}\lambda_k \\ &\leqslant \int\limits_{\bigcup_{B\in\mathcal{B}} B} |f|\mathrm{d}\lambda_k + \sum_{j=2}^{k} \int\limits_{A_j} (|f|-s)_+\mathrm{d}\lambda_k \\ &\leqslant \int\limits_{\bigcup_{B\in\mathcal{B}} B} |f|\mathrm{d}\lambda_k + (k-1) \int\limits_{\bigcup_{j=2}^{k} A_j} (|f|-s)_+\mathrm{d}\lambda_k \\ &\leqslant \int\limits_{\bigcup_{B\in\mathcal{B}} B} |f|\mathrm{d}\lambda_k + (k-1) \int\limits_{\mathcal{R}_k} (|f|-s)_+\mathrm{d}\lambda_k. \end{aligned}$$

Now recall that the family $\mathcal{B}$ depended on N. Letting this parameter to infinity and using Lebesgue's monotone convergence theorem, we obtain

$$s\lambda(E) \leqslant \int\limits_{E} |f|\mathrm{d}\lambda_k + (k-1) \int\limits_{\mathcal{R}_k} (|f|-s)_+\mathrm{d}\lambda_k.$$

This is precisely the claim. ■

Now, using the standard integration argument, we obtain the L^p estimate for the uncentered maximal operator on $\mathcal{R}_k$.

P r o o f o f (1.2). By Fubini's theorem, we have

$$\begin{aligned} &\int\limits_{\mathcal{R}_k} \big(\mathcal{M}f\big)^p\mathrm{d}\lambda_k + (k-1) \int\limits_{\mathcal{R}_k} |f|^p\mathrm{d}\lambda_k \\ &= p\int\limits_0^\infty s^{p-1}\left[\lambda_k(\mathcal{M}f > s) + (k-1)\lambda_k(|f|>s)\right]\mathrm{d}s, \end{aligned}$$

which by (3.1) does not exceed

$$\begin{aligned} &p\int\limits_0^\infty s^{p-2}\left[\int\limits_{\{\mathcal{M}f>s\}} |f|\mathrm{d}\lambda_k + (k-1)\int\limits_{\{|f|>s\}} |f|\mathrm{d}\lambda_k\right]\mathrm{d}s \\ &= \frac{p}{p-1}\int\limits_{\mathcal{R}_k} \big((\mathcal{M}f)^{p-1}|f| + (k-1)|f|^p\big)\mathrm{d}\lambda_k. \end{aligned}$$

Here in the last passage we have used Fubini's theorem again. This gives the bound

$$\int\limits_{\mathcal{R}_k} \big(\mathcal{M}f\big)^p\mathrm{d}\lambda_k \leqslant \frac{p}{p-1}\int\limits_{\mathcal{R}_k} \big(\mathcal{M}f\big)^{p-1}|f|\mathrm{d}\lambda_k + \frac{k-1}{p-1}\int\limits_{\mathcal{R}_k} |f|^p\mathrm{d}\lambda_k.$$

However, by Hölder's inequality, we have

$$\int_{\mathcal{R}_k} (\mathcal{M}f)^{p-1}|f|\mathrm{d}\lambda_k \leqslant \left(\int_{\mathcal{R}_k} (\mathcal{M}f)^p \mathrm{d}\lambda_k\right)^{(p-1)/p} \left(\int_{\mathcal{R}_k} |f|^p \mathrm{d}\lambda_k\right)^{1/p},$$

which combined with the previous estimate yields

$$(p-1)\Big(\frac{\|\mathcal{M}f\|_{L^p(\mathcal{R}_k)}}{\|f\|_{L^p(\mathcal{R}_k)}}\Big)^p - p\Big(\frac{\|\mathcal{M}f\|_{L^p(\mathcal{R}_k)}}{\|f\|_{L^p(\mathcal{R}_k)}}\Big)^{p-1} - (k-1) \leqslant 0.$$

It remains to note that the function $s \mapsto (p-1)s^p - ps^{p-1} - (k-1)$ is increasing on $[1,\infty)$ and $C_{p,k}$ is its unique root. This establishes the desired L^p bound $\|\mathcal{M}f\|_{L^p(\mathcal{R}_k)} \leqslant C_{p,k}\|f\|_{L^p(\mathcal{R}_k)}$. ∎

Combining the L^p estimate we have just proved with the inequality (1.6), we immediately obtain (1.2), Doob's inequality for the coherent random variables. It remains to prove the optimality of the constant $C_{p,k}$ in the latter estimate. Having proved this sharpness, we immediately deduce the optimality of the constant for the uncentered maximal operator.

Proof of sharpness of $C_{p,k}$. Let $1 < p < \infty$ and $k \in \{1,2,\ldots\}$ be fixed. Consider the probability space $\mathcal{R}_k$ with its Borel subsets and normalized one-dimensional Lebesgue's measure λ_k. Fix an auxiliary constant $r \in (0, p^{-1})$ and consider the random variable $\xi(x) = |x|^{-r}$: then the estimate $r < p^{-1}$ guarantees that this variable belongs to L^p. To define the filtrations, let $\lambda_{r,k}$ be the unique root of the equation

$$\lambda(1-r) - (k-1)r\lambda^{(r-1)/r} - 1 = 0, \qquad 1 \leqslant \lambda < \infty. \tag{3.3}$$

The existence and uniqueness of $\lambda_{r,k}$ is direct consequence of the fact that the left-hand side, considered as a function of λ, is strictly increasing, negative at $\lambda = 1$ and positive for large λ. Now, for any $j \in \{1,2,\ldots,k\}$, introduce the closed ball B_j which has the center $e^{2\pi ij/k}(1-\lambda_{r,k}^{-1/r})/2$ and radius $(1+\lambda_{r,k}^{-1/r})/2$. This ball covers the whole ray H_j and some portion of the remaining rays: $|B_j \cap H_i| = \lambda_{r,k}^{-1/r}$ for $i \neq j$. Therefore if x lies on the j-th ray of $\mathcal{R}_k$, then the rescaled ball $|x|B_j = \{|x|y \in \mathcal{R}_k : y \in B_j\}$ satisfies

$$\frac{1}{\lambda_k(|x|B_j)} \int_{|x|B_j} \xi \mathrm{d}\lambda_k = \frac{\int_0^{|x|} \omega^{-r}\mathrm{d}\omega + (k-1)\int_0^{\lambda_{r,k}^{-1/r}|x|} \omega^{-r}\mathrm{d}\omega}{|x| + (k-1)\lambda_{r,k}^{-1/r}|x|} = \lambda_{r,k}\cdot\xi(x),$$

by (3.3). Since both sides are homogeneous of order $-r$ (as a function of x), one can actually show a bit more: for any $\varepsilon > 0$ there is $\delta \in (0,1)$ such that if $y \in H_j$ satisfies $\delta < |y/x| \leqslant 1$, then

$$\frac{1}{\lambda_k(|x|B_j)} \int\limits_{|x|B_j} \xi \mathrm{d}\lambda_k \geqslant (\lambda_{r,k} - \varepsilon) \cdot \xi(y). \tag{3.4}$$

Fix ε, δ with the above property and pick a large integer N. For any $n = 0$, 1, 2, ..., N, let $\mathcal{G}_n^j$ be the σ-algebra generated by the balls B_j, δB_j, $\delta^2 B_j$, ..., $\delta^{n-1} B_j$. It follows directly from (3.4) that

$$M_{\mathcal{G}}\xi \geqslant (\lambda_{r,k} - \varepsilon)\xi \qquad \text{almost surely on } \mathcal{R}_k \setminus \delta^N B_j.$$

But $\xi \in L^p$, as we have already discussed above. Since ε and N were taken arbitrarily, the best constant allowed in the estimate (1.2) is at least $\lambda_{r,k}$. It remains to note that if we let $r \to p^{-1}$, then $\lambda_{r,k}$ converges to the constant $C_{p,k}$: in the limit, the equation (3.3) becomes (1.3). This proves the desired sharpness. ■

Finally, we handle the sharp version of Doob's estimate in the coherent setting.

Proof of Theorem 1.2. Put $\mathbb{P}(\xi = 1) = q$. Then for $t \in [0,1]$ we have the identity $\xi^*(e^{2\pi ij/k}t) = \mathbb{1}(t \leqslant q)$ and therefore

$$\mathcal{M}\xi^*(e^{2\pi ij/k}t) = \begin{cases} 1 & \text{if } t \leqslant q, \\ \dfrac{kq}{(k-1)q+t} & \text{if } t > q, \end{cases}$$

for all $j = 1, 2, \dots, k$. By Theorem 1.4, we can write

$$\begin{aligned} \frac{\|M_{\mathcal{G}}\xi\|_p^p}{\|\xi\|_p^p} \leqslant \frac{\|\mathcal{M}\xi^*\|_p^p}{\|\xi\|_p^p} &= \left[q + \int\limits_q^1 \left(\frac{kq}{(k-1)q+t} \right)^p \mathrm{d}t \right] \frac{1}{q} \\ &= 1 + \int\limits_1^{1/q} \left(\frac{k}{k-1+s} \right)^p \mathrm{d}s \\ &\leqslant 1 + \int\limits_1^{\infty} \left(\frac{k}{k-1+s} \right)^p \mathrm{d}s = 1 + \frac{k}{p-1}, \end{aligned}$$

which gives the desired bound. To see that the estimate is sharp, we construct an example for which all the inequalities above become almost-equalities. Precisely, consider the probability space $\mathcal{R}_k$ with its Borel subsets and normalized one-dimensional Lebesgue's measure λ_k and fix an arbitrary $\varepsilon > 0$. Introduce the random variable $\xi(x) = \mathbb{1}(|x| < q)$, where $q \in (0,1)$ satisfies

$$\int\limits_1^{1/q} \left(\frac{k}{k-1+s} \right)^p \mathrm{d}s + \varepsilon = \int\limits_1^{\infty} \left(\frac{k}{k-1+s} \right)^p \mathrm{d}s.$$

For fixed $1 \leqslant j \leqslant k$ and $0 \leqslant n \leqslant N$, distinguish the point $x_n = (N-n)/(2N)$ and let B_n^j be the ball centered at $e^{2\pi ij/k}x_n$ and of radius $x_n + q$. Finally, consider the filtration $(\mathcal{G}_n^j)_{0\leqslant n\leqslant N} = (\sigma(B_0^j, B_1^j, B_2^j, \ldots, B_n^j))_{0\leqslant n\leqslant N}$. Arguing as above, one easily checks that the maximal function $M_{\mathcal{G}}\xi$ can be made arbitrarily close, in L^∞ norm, to $\mathcal{M}\xi^*$, by picking N sufficiently large. Thus one can guarantee that $\|M_{\mathcal{G}}\xi\|_p^p/\|\xi\|_p^p + \varepsilon > \|\mathcal{M}\xi^*\|_p^p/\|\xi\|_p^p$, and hence we obtain

$$\frac{\|M_{\mathcal{G}}\xi\|_p^p}{\|\xi\|_p^p} > 1 + \frac{k}{p-1} - 2\varepsilon.$$

Since ε was chosen arbitrarily, the sharpness follows. ■

REFERENCES

[1] I. Arieli, Y. Babichenko, F. Sandomirskiy and O. Tamuz, *Feasible Joint Posterior Beliefs* J. Polit. Econ. 129 (2021), pp. 2546–2594.

[2] K. Burdzy and S. Pal, *Can Coherent Predictions be Contradictory?* Adv. in Appl. Probab. 53 (2021), pp. 133–161.

[3] K. Burdzy and J. Pitman, *Bounds on the probability of radically different opinions,* Electron. Commun. Probab. 25 (2020), pp. 1–12.

[4] R. Cairoli and J.B. Walsh, *Stochastic integrals in the plane* Acta Math. 134 (1975), pp. 111 – 183.

[5] T.C. Christofides and R.J. Serfling, *Maximal Inequalities for Multidimensionally Indexed Submartingale Arrays* Ann. Probab. 18 (1990), pp. 630–641.

[6] S. Cichomski and A. Osękowski, *The maximal difference among expert's opinions* Electron. J. Probab. 26 (2021), pp. 1–17.

[7] S. Cichomski and F. Petrov, *A combinatorial proof of the Burdzy–Pitman conjecture* Electron. Commun. Probab. 28 (2023), pp. 1–7.

[8] L. Grafakos and J. Kinnunen, *Sharp inequalities for maximal functions associated with general measures,* Proc. Roy. Soc. Edinburgh Sect. A. 128 (1998), pp. 717–723.

[9] L. Grafakos and S. Montgomery-Smith, *Best constants for uncentered maximal functions,* Bull. London Math. Soc. 29 (1997), pp. 60–64.

[10] K. He, F. Sandomirskiy and O. Tamuz, *Private Private Information* (preprint) available at arXiv:2112.14356 [econ.TH] (2022), pp. 1–70.

[11] T.J. He and Y.L. Hou, *Some Inequalities for Tree Martingales* Acta Math. Appl. Sin. Engl. Ser. 21 (2005), pp. 671–682.

[12] B.G. Ivanoff and E. Merzbach, *Set–indexed Martingales* Chapman and Hall, 2000.

[13] E. Wong and M. Zakai, *Weak Martingales and Stochastic Integrals in the Plane*Ann. Probability 4 (1976), pp. 570 – 586.

University of Warsaw
Faculty of Mathematics
Banacha 2, 02-097
Warsaw, Poland
E-mail: s.cichomski@uw.edu.pl

University of Warsaw
Faculty of Mathematics
Banacha 2, 02-097
Warsaw, Poland
E-mail: ados@mimuw.edu.pl

Received on 16.06.2023;
revised version on xx.xx.xxxx