STUDIA MATHEMATICA

Online First version

The suprema of selector processes with the application to positive infinitely divisible processes

by

WITOLD BEDNORZ, RAFAŁ MARTYNEK and RAFAŁ MELLER

Abstract. We provide an alternative proof of the recent result of Park and Pham [Ann. of Math. 199 (2024)] on the expected suprema of positive selector and empirical processes, which we find useful in further generalizations we discuss. As an application, we obtain a result concerning positive infinitely divisible processes.

1. Introduction. Recent results (see [13, 1]) show that the size of the expected suprema of both empirical and infinitely divisible processes can be explained by replacing the initial index set by two new index sets. Their Minkowski's sum contains the initial index. The size od the process defined on the first, new index set is characterized by Talagrand's gamma functionals, while the process defined on the other index set is a positive process. An understanding of the size of positive processes was lacking and some conjectures were made by Talagrand [12, 13]. Some of them have recently been discussed and proved in [10].

The key new concept was the notion of a fragment, which was used in the proof of the Kahn–Kalai conjecture [10]. We have adopted this idea and managed to obtain a much simpler proof than the original one in [10]. Our proof is also more amenable to further generalizations.

The primary goal which motivates an alternative proof of Park and Pham's result is to provide a small cover for processes $(Y_t)_{t\in T}$ for $T \subset \mathbb{R}_+^d$ and $Y_t = \sum_{i=1}^d t_i Y_i$, where Y_i's are non-negative, independent, but not necessarily identically distributed. The case of i.i.d. variables is already a highly non-trivial result presented in [2] (see also the end of Section 2 below for the discussion of a result which goes beyond the i.i.d. case and is also presented in [2]). From this point of view, selector processes should be seen as a crucial, initial case one has to understand. Interestingly, in the recent paper

2020 *Mathematics Subject Classification*: Primary 60G15; Secondary 60G17.

Key words and phrases: expected supremum, positive process, p-smallness.

Received 20 May 2025; revised 20 January 2026.

Published online 19 September 2026.

DOI: 10.4064/sm250520-27-2

by Dubroff, Kahn and Park [3] our Key Lemma (Lemma 3.4 below) is used to establish a result concerning the so-called expectation threshold and the fractional expectation threshold. Moreover, they call our result the quantitative strengthening of the result in [10, Theorem 1.5] and use it for a very elegant computation in the direction of one of Talagrand's conjectures [12, Conjecture 6.3]

The overall goal is to relate the size of a stochastic process to the geometry of the index set. This leads to a very efficient method of chaining and, in turn, to a characterization in terms of gamma functionals. The most famous result in this area is the Majorizing Measure Theorem (see e.g. [13, Theorem 2.10.1]) for Gaussian processes. One of the geometric consequences of this result is that the construction of the partition of the index set also provides, in a sense, "witnesses" to the size of the supremum. Let us briefly explain this concept (cf. [13, discussion after Theorem 2.12.2 and Section 13.2]). Let $(G_t)_{t\in T}$ be a centered Gaussian process on some index set T. By the Majorizing Measure Theorem one can find a jointly Gaussian sequence $(u_k)_{k\geqslant 1}$ such that for some universal constant $L>0$,

$$\Big\{\sup_{t\in T}|G_t| \geqslant L\,\mathbf{E}\sup_{t\in T}|G_t|\Big\} \subset \bigcup_{k\geqslant 1}\{u_k\geqslant 1\}$$

and

$$\sum_{k\geqslant 1}\mathbf{P}(u_k\geqslant 1)\leqslant \frac{1}{2}.$$

Let us emphasize that the above statement is weaker than the Majorizing Measure Theorem in the sense that it follows relatively easily from it, but the other direction is not at all clear. The basis of the chaining method is a result on Sudakov type bounds, which guarantees that if the points in the index set are well separated, then one can control the expected supremum of the process from below. Since we do not have such a result in the context of selector processes, chaining seems infeasible in this case. Therefore, we want to characterize $\sup_{t\in T} X_t$ for certain nonnegative stochastic processes in the sense of the above representation.

As mentioned before, the main motivation for simplifying the proof of Theorem 2.1 is to find new methods of bounding $\mathbf{E}\sup_{t\in T} X_t$, where $(X_t)_{t\in T}$ is a nonnegative process. Due to the use of multisets in [10] in the case of empirical processes, it is difficult to imagine how this method can serve in different settings. The main difference between our approach and the original proof in [10] is the method of counting of so-called bad sets. In particular, we introduce a kind of a threshold function (see Definition 3.7 and the beginning of the proof of Lemma 5.14), which turns out to be an appropriate tool for generalizations. As an example, we give a result for permutationally invariant processes in the Appendix. Although the further generalizations in [2] are much more complicated than the proof for selector processes, an appropriate version of the threshold function plays a crucial role there.

2. Results and motivation. The letters K, C stand for universal constants. We adopt the convention that K may differ at each occurrence. As already mentioned, the first goal is to present a simplified proof of the result in [10] concerning a positive selector process. Consider a sequence $(\delta_i)_{i=1}^n$ of independent Bernoulli random variables with

$$\mathbf{P}(\delta_i = 1) = p = 1 - \mathbf{P}(\delta_i = 0)$$

for some $p \in (0,1)$ and a set T of sequences $t = (t_i)_{i=1}^n$ with $t_i \geqslant 0$. We assume that T is finite. Let

$$\delta(T) := \mathbf{E} \sup_{t \in T} \sum_{i=1}^n t_i \delta_i .$$

We use the notation $[n] := \{1, \ldots, n\}$. To describe the size of $\delta(T)$, we want to construct "explicit witnesses" of the event

$$\Big\{ \sup_{t \in T} \sum_{i=1}^n t_i \delta_i \geqslant K \delta(T) \Big\}$$

whose measure is small. To formalize this goal let us recall the notions of up-sets and p-smallness. Let $I \subseteq [n]$. Then the *up-set generated by* I is defined as

$$\langle I \rangle := \{J \subseteq [n] : J \supseteq I\},$$

and we say that a collection $\mathcal{F}$ of subsets of $[n]$ is *p-small* if there exists a collection $\mathcal{G}$ of subsets of $[n]$ such that

$$\mathcal{F} \subseteq \langle \mathcal{G} \rangle := \bigcup_{I \in \mathcal{G}} \langle I \rangle \quad \text{and} \sum_{I \in \mathcal{G}} p^{|I|} \leqslant \frac{1}{2}.$$

We refer to $\mathcal{G}$ as a *cover* of $\mathcal{F}$. We prove the following result describing the size of $\delta(T)$.

THEOREM 2.1. *The family*

$$\Big\{ I \subset [n] : \sup_{t \in T} \sum_{i \in I} t_i \geqslant K \delta(T) \Big\} \tag{1}$$

is p-small. One may take $K = 221$.

REMARK 2.2. Clearly $\sup_{t \in T} \sum_{i=1}^n t_i \delta_i \leqslant \sup_{t \in T} \sum_{i=1}^n t_i \leqslant \frac{1}{p} \delta(T)$, so the family in (1) is empty for $K > 1/p$. This shows that Theorem 2.1 is interesting only when p is small.

We point out that the proof of Theorem 2.1 is existential, we do not construct this family explicitly. This is a common situation in this field. Usually, constructive results are derived only in some special cases. The best example is Talagrand's Majorizing Measure Theorem for which an explicit construction of the majorizing measure is known only for simplexes and ellipses (and a few exotic examples). The main new consequence of Theorem 2.1 is the analogous result for positive infinitely divisible processes (see [13, Definition 12.3.4]). Notice that following [13, Chapter 12] we take Rosiński's series representation of infinitely divisible processes as their definition which uses a Poisson point process (see e.g. [13, Section 12.1]).

Consider a finite set T and a measurable space $\mathbb{R}^T$, provided with the cylindrical σ-algebra

$$\sigma(\{x \in \mathbb{R}^T : x(t) \in A\} : t \in T,\, A \in \mathcal{B}(\mathbb{R})),$$

i.e. the σ-algebra generated by the coordinate functions. Let ν be a σ-finite measure on $\mathbb{R}^T$ with $\nu(0) = 0$. Assume that for any $t \in T$,

$$\int_{\mathbb{R}^T} |x(t)| \wedge 1 \, d\nu(x) < \infty,$$

where $\wedge$ is the minimum of two values. The main additional assumption we make about the measure ν is that it has no atoms. Consider a Poisson point process $(Y_i)_{i\geqslant 1}$ on $\mathbb{R}^T$ with intensity measure ν. Then we define *a positive infinitely divisible process with Lévy measure* ν as

$$|X|_t := \sum_{i\geqslant 1} |Y_i(t)|.$$

The standard simplification is to consider T as a space of functions on $\mathbb{R}^T$ by associating to $t \in T$ the function $x \mapsto x(t)$ on $\mathbb{R}^T$, so that we can write $t(Y_i)$ instead of $Y_i(t)$. Thus, we have

$$\mathbf{E} \sup_{t\in T} |X|_t = \mathbf{E} \sup_{t\in T} \sum_{i\geqslant 1} |t(Y_i)|.$$

THEOREM 2.3. *Let $|X|_t$ be a positive, infinitely divisible process with a nonatomic Lévy measure ν, indexed by a finite set T. Then there exists a family $\mathcal{F}$ of pairs (g, u), where each $g : \mathbb{R}^T \to \mathbb{R}_+$ is measurable* (*with respect to the cylindrical σ-algebra*) *and $u \geqslant 0$ such that*

$$\Big\{\sup_{t\in T} |X|_t \geqslant K\, \mathbf{E} \sup_{t\in T} |X|_t\Big\} \subset \bigcup_{(g,u)\in\mathcal{F}} \{|X|_g \geqslant u\}, \tag{2}$$

and

$$\sum_{(g,u)\in\mathcal{F}} \mathbf{P}(|X|_g \geqslant u) \leqslant \frac{1}{2}. \tag{3}$$

Obviously, by the Markov inequality, the probability of the event described in (2) is less than $1/K$. But the goal is to find "explicit witnesses" that this event is small.

The third setting that we work with is the one of positive empirical processes. Although we must assume that the underlying measure μ is nonatomic, we find it instructive to see that the result can be obtained without the use of multisets, in contrast to [10]. We believe this makes the application of the result for the selectors more transparent. For the result in full generality we refer to [10].

THEOREM 2.4. *Let (E,ρ) be some metric space, and $d>0$ be a fixed integer. Suppose μ is a probability measure on E with no atoms. Let $X_1,\ldots,X_d$ be independent random variables distributed according to μ. Let T be a finite class of nonnegative Borel measurable functions on E. Then there exists a family $\mathcal{F}$ of pairs (g,u) where each $g : E\to\mathbb{R}_+$ is Borel measurable and $u\geqslant 0$ such that*

$$\Big\{\sup_{t\in T}\sum_{i=1}^{d} t(X_i)\geqslant K\,\mathbf{E}\sup_{t\in T}\sum_{i=1}^{d} t(X_i)\Big\}\subset\bigcup_{(g,u)\in\mathcal{F}}\Big\{\frac{1}{d}\sum_{i=1}^{d} g(X_i)\geqslant u\Big\}\tag{4}$$

and

$$\sum_{(g,u)\in\mathcal{F}}\mathbf{P}\Big(\frac{1}{d}\sum_{i=1}^{d} g(X_i)\geqslant u\Big)\leqslant\frac{1}{2}.\tag{5}$$

The metric structure is irrelevant for the formulation of the theorem. However, it is crucial for our proof. Roughly speaking, we need to decompose a set $A\subset E$ into many subsets of almost equal measure (see Step 3 of the proof). It is not clear whether such a construction is possible in nonmetric spaces.

As we show in [2], the assumption that the measures have no atoms is not necessary in both Theorems 2.3 and 2.4. We also suspect that the assumption that T is finite in Theorems 2.3 or 2.4 is redundant, but it is unclear how to extend the results to infinite index sets.

We finish this section with a short discussion of a possible direction for future work. In [2] we manage, by using the techniques developed in this paper, to provide a small cover for

$$\Big\{\sup_{f\in T}\sum_{i=1}^{d} f_i(X_i)\geqslant K\,\mathbf{E}\sup_{f\in T}\sum_{i=1}^{d} f_i(X_i)\Big\},\tag{6}$$

where $X_1,\ldots,X_d$ are nonnegative i.i.d. random variables and T is a finite set of d-tuples of nonnegative functions. We believe that this result can be of use in some other important problems. Let us briefly discuss one of them. The problem of finding tight bounds on the operator norm of random matrices is of importance in probability theory and functional analysis. The most important case of nonhomogeneous Gaussian matrices has been solved by R. van Handel, R. Latała and P. Youssef [7]. But the equally important question about nonhomogeneous Rademacher matrices is still open. The task is to find a two-sided bound for the quantities

$$\mathbf{E}\sup_{x,y\in B_2^n}\sum_{ij}a_{ij}x_iy_j\varepsilon_{ij}=\mathbf{E}\sup_{x\in B_2^n}\sqrt{\sum_i\Big(\sum_j a_{ij}x_j\varepsilon_{ij}\Big)^2}\approx\sqrt{\mathbf{E}\sup_{x\in B_2^n}\sum_i\Big(\sum_j a_{ij}x_j\varepsilon_{ij}\Big)^2},\tag{7}$$

where a_{ij} are real numbers and ε_{ij} are independent, symmetric ± 1 random variables. The quantity $\sup_{x\in B_2^n}\sum_i(\sum_j a_{ij}x_j\varepsilon_{ij})^2$ can be easily expressed as $\sup_{f\in T}\sum_{i=1}^d f_i(X_i)$. The existence of a small cover of the set (6) may result in a better understanding of the problem of estimating the operator norm of Rademacher matrices. In particular, it may lead to a counterpart of Theorem 3.1 in this case, potentially providing new results. For a broader discussion of estimates of (7), we refer the reader to [5, 6]; see also [9] for the case of more general random variables.

The organization of the paper is the following. In the next three sections, we prove, respectively, Theorem 2.1, the result on infinitely divisible processes and on empirical processes. Two of them are present in [10] while the one concerning positive infinitely divisible processes is our main application of the result about selector processes. In the Appendix we present the result for processes invariant under permutations of the index set, which aims to show how the main argument can be adjusted for further applications.

3. Positive selector processes. Theorem 2.1 can be reformulated as follows (see the proof of Theorem 5.3 below or [10, Section 2.1]).

THEOREM 3.1. *Let $\mathcal{F}$ be a family of subsets of $[n]$ which is not p-small. Suppose that for each $I\in\mathcal{F}$ we are given coefficients $(\mu_I(i))_{i\in I}$ with $\mu_I(i)\geqslant 0$ and*

$$\mu_I(I):=\sum_{i\in I}\mu_I(i)=1.$$

Then

$$\mathbf{E}\sup_{I\in\mathcal{F}}\sum_{i\in I}\mu_I(i)\delta_i\geqslant\frac{1}{220}.$$

Theorem 3.1 is obvious if constant $\frac{1}{220}$ is replaced by p, where $p=\mathbf{P}(\delta_i=1)$ (even without the essential assumption that $\mathcal{F}$ is not p-small). Thus it is interesting only for small values of p. (cf. Remark 2.2).

REMARK 3.2. Obviously, the empty set does not contribute to the supremum over $I\in\mathcal{F}$, so we may assume that $\emptyset\notin\mathcal{F}$. For any such family $\mathcal{F}$ of subsets of $[n]$ we have $\mathcal{F}\subseteq\bigcup_{i\leqslant n}\langle\{i\}\rangle$. Thus, Theorem 3.1 is trivial when

$np < 1/2$ (in this case any family $\mathcal{F}$ is p-small). Thus, with no loss in generality, we may assume that $np \geqslant 1/2$.

Our goal now is to prove Theorem 3.1. To this end, we introduce a pivotal definition.

DEFINITION 3.3. We say that a set $X \subseteq [n]$ is *c-bad* $(c > 0)$ if

$$\sup_{I \in \mathcal{F}} \mu_I(X \cap I) < c. \tag{8}$$

By saying that $X \subseteq [n]$ is *random* we mean that

$$X = \{i \in [n] : \delta'_i = 1\}, \tag{9}$$

where $(\delta'_i)_{i \leqslant n}$ are independent and identically distributed Bernoulli variables with arbitrary probability of success. Notice that the distribution of X conditioned on its cardinality being m is uniform over m-element subsets of $[n]$.

We deduce Theorem 3.1 from the following.

LEMMA 3.4 (Key Lemma). *Let $\mathcal{F}$ be a family of subsets of $[n]$ which is not p-small. For any $m \leqslant n$ we have*

$$\mathbf{P}\big(X \textit{ is } 1/2\textit{-bad} \,\big|\, |X| = m\big) \leqslant \sum_{t=1}^{n} \left(4\frac{np}{m}\right)^t,$$

where a bad set is defined in (8)*.*

Proof of Theorem 3.1. By Remark 3.2 we can assume that $np \geqslant 1/2$. Thus we may further assume that Cnp is an integer and $9 \leqslant C \leqslant 11$. Observe that for each i we can represent $\delta_i = \delta'_i \delta''_i$, where $(\delta'_i)_{i=1}^n, (\delta''_i)_{i=1}^n$ are independent Bernoulli random variables with $\mathbf{P}(\delta'_i = 1) = Cp$, $\mathbf{P}(\delta''_i = 1) = 1/C$. By Jensen's inequality,

$$\begin{aligned} \mathbf{E} \sup_{I \in \mathcal{F}} \sum_{i \in I} \mu_I(i) \delta_i &= \mathbf{E} \sup_{I \in \mathcal{F}} \sum_{i \in I} \mu_I(i) \delta'_i \delta''_i \\ &\geqslant \mathbf{E} \sup_{I \in \mathcal{F}} \sum_{i \in I} \mu_I(i) \delta'_i \mathbf{E} \delta''_i \geqslant \frac{1}{11} \mathbf{E} \sup_{I \in \mathcal{F}} \sum_{i \in I} \mu_I(i) \delta'_i. \end{aligned} \tag{10}$$

Lemma 3.4 applied for $m \geqslant Cnp$ gives

$$\begin{aligned} \mathbf{E}\Big(\sup_{I \in \mathcal{F}} \sum_{i \in I} \mu_I(i) \delta'_i \,\Big|\, \sum_{i=1}^n \delta'_i = m\Big) &\geqslant \frac{1}{2} \mathbf{P}\Big(X \text{ is not } 1/2\text{-bad} \,\Big|\, \sum_{i=1}^n \delta'_i = m\Big) \\ \geqslant \frac{1}{2}\left(1 - \sum_{t=1}^n \left(4\frac{np}{m}\right)^t\right) &\geqslant \frac{1}{2}\left(1 - \frac{np}{m} \frac{4}{1 - \frac{4np}{m}}\right) \geqslant \frac{1}{2}\left(1 - \frac{4}{C-4}\right) \geqslant \frac{1}{10}. \end{aligned} \tag{11}$$

So, by conditioning,

$$\mathbf{E}\sup_{I\in\mathcal{F}}\sum_{i\in I}\mu_I(i)\delta_i'\geqslant\sum_{m\geqslant Cpn}\mathbf{E}\Big(\sup_{I\in\mathcal{F}}\sum_{i\in I}\mu_I(i)\delta_i'\Big|\sum_{i=1}^n\delta_i'=m\Big)\mathbf{P}\Big(\sum_{i=1}^n\delta_i'=m\Big) \tag{12}$$
$$\geqslant\frac{1}{10}\mathbf{P}\Big(\sum_{i=1}^n\delta_i'\geqslant Cpn\Big)\geqslant\frac{1}{20},$$

where in the last inequality we have used [8] (we recall that Cnp is an integer). The assertion follows by (10) and (12). ∎

Remark 3.5. Theorem 2.1 implies that the event

$$\Big\{\sup_{t\in T}\sum_{i=1}^n t_i\delta_i\geqslant 221C\delta(T)\Big\}$$

is Cp-small. To see this, let δ_i',δ_i'' be as in the proof of Theorem 3.1. Denote $\delta'(T):=\mathbf{E}\sup_{t\in T}\sum_{i=1}^n t_i\delta_i'$. By Theorem 2.1 the event

$$\Big\{\sup_{t\in T}\sum_{i=1}^n t_i\delta_i'\geqslant 221\delta'(T)\Big\}$$

is Cp-small and trivially

$$\Big\{\sup_{t\in T}\sum_{i=1}^n t_i\delta_i'\geqslant 221\delta'(T)\Big\}\supset\Big\{\sup_{t\in T}\sum_{i=1}^n t_i\delta_i\geqslant 221\delta'(T)\Big\}.$$

The inequality (10) implies that $\delta'(T)\leqslant C\delta(T)$. Thus, we obtain

$$\Big\{\sup_{t\in T}\sum_{i=1}^n t_i\delta_i\geqslant 221C\delta(T)\Big\}\subset\Big\{\sup_{t\in T}\sum_{i=1}^n t_i\delta_i'\geqslant 221\delta'(T)\Big\},$$

and the result follows.

3.1. Proof of the Key Lemma. The overall strategy for the proof is to construct a cover of $\mathcal{F}$ based on the class of bad sets with fixed cardinality and show that its measure is appropriately small. To achieve this goal we need sets closely related to so-called fragments (cf. [10, Definition 2.3] or [11], [4] for more context on the notion of fragment). Let I_j be the subset of I consisting of j elements with the largest values of coefficients, i.e.

$$I_j=\Big\{i_l\in I,\,1\leqslant l\leqslant j:\mu_I(i_1)\geqslant\cdots\geqslant\mu_I(i_j)\geqslant\sup_{i\in I\setminus\{i_1,\ldots,i_j\}}\mu_I(i)\Big\},$$

where by convention, $I_0=\emptyset$ and $I_j=I$ for $j\geqslant|I|$. Also, $I_j^c=I\setminus I_j$.

Lemma 3.6. *Let $X\subseteq[n]$ be c-bad (recall (8)) for some $c>0$ Then for any $I\in\mathcal{F}$ there exists a number $j(I,X)$ such that*

$$|I_{j(I,X)}\cap X|<c|I_{j(I,X)}|. \tag{13}$$

Proof. Assume that X is nonempty; otherwise, there is nothing to prove. Fix $I \in \mathcal{F}$. For $\varepsilon \in [0,1]$ consider the mapping

$$\varepsilon \mapsto \sum_{i \in I \cap X} \mu_I(i) \wedge \varepsilon - c \sum_{i \in I} \mu_I(i) \wedge \varepsilon$$

and observe that it is continuous. Moreover, for $\varepsilon = 1$ we see by (8) that

$$\sum_{i \in I \cap X} \mu_I(i) - c \sum_{i \in I} \mu_I(i) < c - c\mu_I(I) = c - c = 0,$$

and obviously $\varepsilon = 0$ is mapped to 0. This means that there exists the largest value $\varepsilon(I,X) \in [0,1)$ such that

$$\sum_{i \in I \cap X} \mu_I(i) \wedge \varepsilon(I,X) \geqslant c \sum_{i \in I} \mu_I(i) \wedge \varepsilon(I,X) \tag{14}$$

and for $\varepsilon > \varepsilon(I,X)$ the strict reverse inequality holds. Now, observe that for sufficiently small $\delta > 0$ we have

$$\sum_{i \in I \cap X} \mu_I(i) \wedge (\varepsilon(I,X) + \delta) = \sum_{i \in I \cap X} \mu_I(i) \wedge \varepsilon(I,X) + \delta \sum_{i \in I \cap X} \mathbb{1}_{\{\mu_I(i) > \varepsilon(I,X)\}} \tag{15}$$

and

$$\sum_{i \in I} \mu_I(i) \wedge (\varepsilon(I,X) + \delta) = \sum_{i \in I} \mu_I(i) \wedge \varepsilon(I,X) + \delta \sum_{i \in I} \mathbb{1}_{\{\mu_I(i) > \varepsilon(I,X)\}}. \tag{16}$$

By the definition of $\varepsilon(I,X)$,

$$\sum_{i \in I \cap X} \mu_I(i) \wedge (\varepsilon(I,X) + \delta) < c \sum_{i \in I} \mu_I(i) \wedge (\varepsilon(I,X) + \delta).$$

We combine the above inequality with (15) and (16) to get

$$\begin{aligned} \sum_{i \in I \cap X} \mu_I(i) \wedge \varepsilon(I,X) + \delta \sum_{i \in I \cap X} \mathbb{1}_{\{\mu_I(i) > \varepsilon(I,X)\}} \\ < c \sum_{i \in I} \mu_I(i) \wedge \varepsilon(I,X) + c\delta \sum_{i \in I} \mathbb{1}_{\{\mu_I(i) > \varepsilon(I,X)\}}, \end{aligned}$$

but because of (14) it follows that

$$\sum_{i \in I \cap X} \mathbb{1}_{\{\mu_I(i) > \varepsilon(I,X)\}} < c \sum_{i \in I} \mathbb{1}_{\{\mu_I(i) > \varepsilon(I,X)\}}.$$

Let

$$\tilde{I} := \{i \in I : \mu_I(i) > \varepsilon(I,X)\} \quad \text{and} \quad j(I,X) := |\tilde{I}|. \tag{17}$$

The result follows, since obviously

$$I_{j(I,X)} = \tilde{I}. \ \blacksquare \tag{18}$$

Since the quantities $\varepsilon(I,X)$ and $j(I,X)$ play a central role in our construction, we introduce them in a separate definition.

Definition 3.7. Let $X \subseteq [n]$ be c-bad (recall (8)) for some $c > 0$ and $I \in \mathcal{F}$. Let

$$F(\varepsilon) = \sum_{i \in I \cap X} \mu_I(i) \wedge \varepsilon - c \sum_{i \in I} \mu_I(i) \wedge \varepsilon.$$

We define $\varepsilon(I, X)$ as the largest ε for which $F(\varepsilon) \geqslant 0$. We also define

$$j(I, X) := |\{i \in I : \mu_I(i) > \varepsilon(I, X)\}|. \tag{19}$$

Definition 3.8 (Witness). Fix a c-bad set $X \subset [n]$. To each $I \in \mathcal{F}$ we associate the number $j(I, X)$ as in Definition 3.7. Now, fix $I \in \mathcal{F}$. Consider sets $\hat{I} \in \mathcal{F}$ such that

$$\hat{I}_{j(\hat{I}, X)} \setminus X \subset I \setminus X. \tag{20}$$

Among all $\hat{I}$ for which (20) holds we consider the ones for which $j(\hat{I}, X)$ is the smallest possible. Among the latter, we pick any $\hat{I}$ such that $\hat{I}_{j(\hat{I}, X)} \setminus X$ has the minimal number of elements and denote it by $I' = I'(I, X)$. We refer to I' as the (I, X)-*minimal set*, or simply the *minimal set*. We also define an (I, X)-*witness* by

$$W(I, X) := I'_{j(I', X)}.$$

If $j = j(I', X)$ and $t = |I'_j \setminus X|$ then it is easy to see that by Lemma 3.6 we have

$$j \geqslant t \geqslant (1 - c)j. \tag{21}$$

The cover of $\mathcal{F}$ we are interested in is given for a fixed c-bad X by

$$\mathcal{G}(X) = \{W(I, X) \setminus X : I \in \mathcal{F}\}. \tag{22}$$

Remark 3.9. Indeed, this is a cover of $\mathcal{F}$ since $W(I, X) \setminus X \subseteq I$ for every $I \in \mathcal{F}$. The sets $W(I, X) \setminus X$ are analogues of fragments in [10]. However, the notion of witness is more suitable for our proof.

Since for any X which is 1/2-bad, $\mathcal{G}(X)$ is a cover of $\mathcal{F}$, and $\mathcal{F}$ is not p-small, we have

$$\frac{1}{2} \leqslant \sum_{G \in \mathcal{G}(X)} p^{|G|}.$$

In addition, the number of sets $X \subset [n]$ with $|X| = m$ that are 1/2-bad is

$$\mathbf{P}\big(X \text{ is } 1/2\text{-bad} \,\big|\, |X| = m\big) \binom{n}{m}.$$

Therefore,

$$\frac{1}{2} \mathbf{P}\big(X \text{ is } 1/2\text{-bad} \,\big|\, |X| = m\big) \binom{n}{m} \leqslant \sum_{\substack{X \text{ is } 1/2\text{-bad} \\ |X| = m}} \sum_{G \in \mathcal{G}(X)} p^{|G|}. \tag{23}$$

In order to bound the above double sum we need to control from above the number of repetitions of the witnesses. This is deduced in Corollary 3.13 from the following two results.

LEMMA 3.10. *Take any c-bad sets X, Y and $I, J \in \mathcal{F}$. Denote $I' = I'(I, X)$ and $J' = J'(J, Y)$ (the minimal sets). Assume that*

(i) $j := j(I', X) = j(J', Y)$ (*recall* (19)),
(ii) $Z := I'_j \cup X = J'_j \cup Y$,
(iii) $t := |I'_j \setminus X| = |J'_j \setminus Y|$.

Then

$$\sum_{i \in I' \cap Y} \mu_{I'}(i) \wedge \varepsilon(I', X) \geqslant c \sum_{i \in I'} \mu_{I'}(i) \wedge \varepsilon(I', X),$$

where $\varepsilon(I', X)$ is defined as in (14).

REMARK 3.11. By the definition, $\varepsilon(I', Y)$ is the greatest number for which the above inequality is satisfied. This means that $\varepsilon(I', Y) \geqslant \varepsilon(I', X)$. This is a key consequence of Lemma 3.10.

Proof of Lemma 3.10. Let $(I'_j)^c := I' \setminus I'_j$, where we recall (see (18))

$$I'_j = \{i \in I' : \mu_{I'}(i) > \varepsilon(I', X)\}.$$

By the above and since $Y \subset Z$ we have

$$\begin{aligned}
&\sum_{i \in I' \cap Y} \mu_{I'}(i) \wedge \varepsilon(I', X) = \sum_{i \in (I'_j)^c \cap Y} \mu_{I'}(i) + \sum_{i \in I'_j \cap Y} \varepsilon(I', X) \\
&\quad = \sum_{i \in (I'_j)^c \cap Z} \mu_{I'}(i) - \sum_{i \in (I'_j)^c \cap (Z \setminus Y)} \mu_{I'}(i) + \varepsilon(I', X)|I'_j \cap Y| \\
&\quad = \sum_{i \in (I'_j)^c \cap Z} \mu_{I'}(i) - \sum_{i \in (I'_j)^c \cap (J'_j \setminus Y)} \mu_{I'}(i) + \varepsilon(I', X)\big(|I'_j \cap Z| - |I'_j \cap (Z \setminus Y)|\big) \\
&\quad = \sum_{i \in (I'_j)^c \cap Z} \mu_{I'}(i) - \sum_{i \in (I'_j)^c \cap ((J'_j \setminus I'_j) \setminus Y)} \mu_{I'}(i) + \varepsilon(I', X)\big(j - |I'_j \cap (J'_j \setminus Y)|\big),
\end{aligned}$$

where in the last two equalities we have used assumption (ii). By (iii), the last line is equal to

$$\begin{aligned}
\sum_{i \in (I'_j)^c \cap Z} \mu_{I'}(i) - \sum_{i \in (I'_j)^c \cap ((J'_j \setminus I'_j) \setminus Y)} \mu_{I'}(i) + \varepsilon(I', X)\big(j - t + |(J'_j \setminus I'_j) \setminus Y|\big) \\
\geqslant \sum_{i \in (I'_j)^c \cap Z} \mu_{I'}(i) + \varepsilon(I', X)(j - t).
\end{aligned}$$

The last inequality holds because if $i \in (I'_j)^c$ then by the definition of I'_j (see (18)) we have $\mu_{I'}(i) < \varepsilon(I', X)$. Now, by assumption (iii) once more,

$$\sum_{i \in (I'_j)^c \cap Z} \mu_{I'}(i) + \varepsilon(I', X)(j - t) = \sum_{i \in (I'_j)^c \cap Z} \mu_{I'}(i) + \varepsilon(I', X)|I'_j \cap X|$$
$$= \sum_{i \in I' \cap X} \mu_{I'}(i) \wedge \varepsilon(I', X) \geqslant c \sum_{i \in I'} \mu_{I'}(i) \wedge \varepsilon(I', X).$$

The last inequality follows again from the definition (14) of $\varepsilon(I', X)$. ∎

LEMMA 3.12. *Under the assumptions of Lemma* 3.10 *we have* $I'_j \setminus Y = J'_j \setminus Y$ *(i.e.* $W(I, X) \setminus Y = W(J, Y) \setminus Y$*).*

Proof. By the construction of the witness, together with assumption (ii) of Lemma 3.10, we have

$$J \setminus Y \supset J'_j \setminus Y = (J'_j \cup Y) \setminus Y = (I'_j \cup X) \setminus Y \supset I'_j \setminus Y. \tag{24}$$

Thus we must have

$$j_0 := j(I', Y) \geqslant j, \tag{25}$$

otherwise we would have

$$I'_{j_0} \setminus Y \subset I'_j \setminus Y \subset J \setminus Y,$$

which would contradict the optimality of $j = j(J', Y)$. Recall (18) and suppose that $j_0 = j(I', Y) > j$. The larger cardinality obviously indicates the lower threshold, therefore we have $\varepsilon(I', Y) < \varepsilon(I', X)$. This, however, contradicts Lemma 3.10 (see Remark 3.11); hence, we must have $j_0 = j$. This implies that $|I'_j \setminus Y| = |J'_j \setminus Y|$ by the construction of the witness. Since these two finite sets have the same cardinality and one is a subset of the other by (24), it follows that in fact $I'_j \setminus Y = J'_j \setminus Y$. ∎

COROLLARY 3.13. *Fix* m, t *and* $Z \subset [n]$ *such that* $|Z| = m + t$. *Fix also* $X \subset Z$ *which is c-bad and* $|X| = m$. *Then for any* $t \leqslant j \leqslant n$,

$$\begin{aligned} &|\{W(I, X) \setminus X : I \in \mathcal{F},\ |W(I, X)| = j, \\ &\qquad Z = W(I, X) \cup X,\ |W(I, X) \setminus X| = t\}| \leqslant \binom{j}{t}. \end{aligned} \tag{26}$$

Proof. Fix any X, I satisfying the conditions in (26). Take any other pair Y, J which also satisfies the same conditions. By Lemma 3.12, $W(J, Y) \setminus Y \subset W(I, X)$, so there are at most $\binom{j}{t}$ choices for $W(J, Y) \setminus Y$. ∎

We are in a position to bound the expression in (23) and prove the Key Lemma.

Proof of Lemma 3.4. We recall the notation $I'_j = W(I, X)$, where $j = j(I, X)$ (cf. (17)). Recall the construction of the cover (22). Then (21) together with Corollary 3.13 implies that

$$\begin{aligned}(27)\qquad &\sum_{\substack{X \text{ is } 1/2\text{-bad}\\ |X|=m}} \sum_{G\in\mathcal{G}(X)} p^{|G|}\\ &= \sum_{j\geqslant 1} \sum_{j/2\leqslant t\leqslant j} \sum_{|Z|=m+t} p^t |\{W(I,X)\setminus X : Z = I'_j \cup X,\ |I'_j\setminus X| = t\}|\\ &\leqslant \sum_{j\geqslant 1} \sum_{j/2\leqslant t\leqslant j} \sum_{|Z|=m+t} p^t \binom{j}{t} = \sum_{t=1}^{n}\sum_{j=t}^{2t}\binom{j}{t}\binom{n}{m+t}p^t.\end{aligned}$$

Clearly

$$\sum_{j=t}^{2t}\binom{j}{t} = \binom{2t+1}{t+1} \leqslant \frac{1}{2}2^{2t+1} = 4^t$$

and

$$\frac{\binom{n}{m+t}}{\binom{n}{m}} \leqslant \left(\frac{n}{m}\right)^t.$$

Those two inequalities applied to (27) give

$$\sum_{\substack{X \text{ is } 1/2\text{-bad}\\ |X|=m}} \sum_{G\in\mathcal{G}(X)} p^{|G|} \leqslant \binom{n}{m}\sum_{t=1}^{n}\left(4\frac{np}{m}\right)^t.$$

The above together with (23) implies the assertion. ∎

4. Positive infinitely divisible processes. The following notion is an analogue of p-smallness for an event in $\mathbb{R}^T$.

DEFINITION 4.1. We say that a random event A is δ-*small* if there exists a family $\mathcal{F}$ of pairs (g,u), where $g:\mathbb{R}^T\to\mathbb{R}_+$ is measurable (with respect to the cylindrical σ-algebra) and $u\geqslant 0$ such that

$$A \subset \bigcup_{(g,u)\in\mathcal{F}} \{|X|_g \geqslant u\} \quad\text{and}\quad \sum_{(g,u)\in\mathcal{F}} \mathbf{P}(|X|_g\geqslant u) \leqslant \delta.$$

In other words, Theorem 2.3 states that the event

$$\Big\{\sup_{t\in T}|X|_t \geqslant K\,\mathbf{E}\sup_{t\in T}|X|_t\Big\}$$

is 1/2-small.

Proof of Theorem 2.3. Note that if $S = \mathbf{E}\sup_{t\in T}|X|_t = \infty$ then the left-hand side of (2) is empty, so any small cover will work in this case. In particular, we can assume that

$$\sup_{t\in T}\int_{\mathbb{R}^T}|t(x)|\,d\nu(x)=\sup_{t\in T}\mathbf{E}\sum_{i\geqslant 1}|t(Y_i)|=\sup_{t\in T}\mathbf{E}|X|_t \leqslant \mathbf{E}\sup_{t\in T}|X|_t=S<\infty. \tag{28}$$

The proof is based on successive simplifications, which ultimately allow us to apply Theorem 2.1.

STEP 1. Consider some fixed large integer N and the functions $f,h:\mathbb{R}\to\mathbb{R}$ given by

$$f(s)=|s|\mathbb{1}_{\{|s|\geqslant 2^{-N}\}},\qquad h(s)=|s|\mathbb{1}_{\{|s|<2^{-N}\}}.$$

We split the process into two parts,

$$|X|_t=\sum_{i\geqslant 1}|t(Y_i)|\mathbb{1}_{\{|t(Y_i)|\geqslant 2^{-N}\}}+\sum_{i\geqslant 1}|t(Y_i)|\mathbb{1}_{\{|t(Y_i)|<2^{-N}\}}=|X|_{f(t)}+|X|_{h(t)}.$$

Since $|X|_t\geqslant |X|_{f(t)}$ we have (recall that K may differ at each occurrence)

$$\Big\{\sup_{t\in T}|X|_t\geqslant KS\Big\}\subset\Big\{\sup_{t\in T}|X|_{f(t)}\geqslant K\,\mathbf{E}\sup_{t\in T}|X|_{f(t)}\Big\}\cup\bigcup_{t\in T}\{|X|_{h(t)}\geqslant KS\}. \tag{29}$$

Markov's inequality implies that

$$\sum_{t\in T}\mathbf{P}(|X|_{h(t)}\geqslant KS)\leqslant\sum_{t\in T}\frac{\mathbf{E}|X|_{h(t)}}{KS}=\sum_{t\in T}\frac{\int_{\mathbb{R}^T}|t(x)|\mathbb{1}_{\{|t(x)|<2^{-N}\}}\,d\nu(x)}{KS}.$$

Since T is finite, by (28) we have $\sum_{t\in T}\mathbf{P}(|X|_{h(t)}\geqslant KS)<\frac{1}{16}$ for sufficiently large N. So by (29) and the above considerations it is enough to show that the event

$$\Big\{\sup_{t\in T}|X|_{f(t)}\geqslant K\,\mathbf{E}\sup_{t\in T}|X|_{f(t)}\Big\}$$

is 7/16-small. Thus, without loss of generality, we can assume that

$$\forall_{t\in T}\forall_{x\in\mathbb{R}^T}\ \ t(x)=0 \text{ or } |t(x)|\geqslant 2^{-N}. \tag{30}$$

In particular, this implies that

$$\nu(\{x\in\mathbb{R}^T:|t(x)|>0\})\leqslant 2^N\int_{\mathbb{R}^T}|t(x)|\,d\nu(x)\leqslant 2^N S. \tag{31}$$

STEP 2. Let $M:=16S|T|$. For each $t\in T$ denote $B(t):=\{x\in\mathbb{R}^T:|t(x)|\geqslant M\}$. Since for each $t\in T$,

$$\{\exists_i\ |t(Y_i)|\geqslant M\}=\{|X|_{\mathbb{1}_{B(t)}}\geqslant 1\},$$

we have

$$\Big\{\sup_{t\in T}|X|_t>KS\Big\}\subset\Big\{\sup_{t\in T}\sum_i|t(Y_i)|\mathbb{1}_{|t(Y_i)|\leqslant M}\geqslant KS\Big\}\cup\bigcup_{t\in T}\{|X|_{\mathbb{1}_{B(t)}}\geqslant 1\}.$$

Observe that

$$\sum_{t\in T}\mathbf{P}(|X|_{\mathbb{1}_{B(t)}}\geqslant 1)\leqslant\sum_{t\in T}\mathbf{E}|X|_{\mathbb{1}_{B(t)}}\leqslant\sum_{t\in T}\mathbf{E}\frac{|X|_t}{M}\leqslant\sum_{t\in T}\frac{S}{M}\leqslant\frac{1}{16}.$$

Thus, it is enough to show that the following event is 6/16-small:

$$\Big\{\sup_{t\in T}\sum_i|t(Y_i)|\mathbb{1}_{|t(Y_i)|\leqslant M}\geqslant KS\Big\}.$$

STEP 3. The third step is to replace the measurable functions $t\in T$ (treated as functions on $\mathbb{R}^T$) with the step functions. Define the following set of sequences, indexed by T:

$$\mathcal{S}:=\{(k(t))_{t\in T}:\forall_{t\in T}\ 0\leqslant k(t)\leqslant M4^N\text{ and }\exists_{t\in T}\ k(t)\neq 0\}.$$

Let

$$\mathcal{V}_N:=\{A^N_{(k(t))_{t\in T}}:(k(t))_{t\in T}\in\mathcal{S}\},$$

where

$$A^N_{(k(t))_{t\in T}}:=\bigcap_{t\in T}\Big\{x\in\mathbb{R}^T:\frac{k(t)-1}{4^N}<|t(x)|\leqslant\frac{k(t)}{4^N}\Big\}.$$

If $k(t)=0$ then (30) implies

$$\Big\{x\in\mathbb{R}^T:\frac{k(t)-1}{4^N}<|t(x)|\leqslant\frac{k(t)}{4^N}\Big\}=\{x\in\mathbb{R}^T:x(t)=0\}.$$

Since the sequence constantly equal to zero does not belong to $\mathcal{S}$, we have

$$\begin{aligned}\bigcup_{A\in\mathcal{V}_N}A&=\Big(\bigcap_{t\in T}\{x\in\mathbb{R}^T:|t(x)|\leqslant M\}\Big)\setminus\{x\in\mathbb{R}^T:\forall_{t\in T}\ t(x)=0\}\\&\subset\bigcup_{t\in T}\{x\in\mathbb{R}^T:0<|t(x)|\leqslant M\}.\end{aligned}\tag{32}$$

Because $\mathcal{V}_N$ consists of disjoint sets, by using (31) we obtain

$$\sum_{A\in\mathcal{V}_N}\nu(A)\leqslant\sum_{t\in T}\nu(\{x\in\mathbb{R}^T:0<|t(x)|\leqslant M\})\leqslant 2^N|T|S.\tag{33}$$

For each $t\in T$ and $A\in\mathcal{V}_N$ let $k(A,t)$ be the unique integer such that the number $t(A):=k(A,t)/4^N$ satisfies

$$\forall_{x\in A}\ t(A)-1/4^N<|t(x)|\leqslant t(A).\tag{34}$$

Thus, if we define

$$Y(A):=\sum_{i\geqslant 1}\mathbb{1}_{\{Y_i\in A\}},\tag{35}$$

by (32) and (34) we obtain

$$\sup_{t\in T}\sum_i|t(Y_i)|\mathbb{1}_{|t(Y_i)|\leqslant M}\leqslant\sup_{t\in T}\sum_{A\in\mathcal{V}_N}t(A)Y(A).$$

As a result,

$$\Big\{\sup_{t\in T}\sum_i |t(Y_i)|\mathbb{1}_{|t(Y_i)|\leqslant M}\geqslant KS\Big\}\subset\Big\{\sup_{t\in T}\sum_{A\in\mathcal{V}_N} t(A)Y(A)\geqslant KS\Big\},$$

and it is sufficient to show that the latter set is 6/16-small.

STEP 4. Let

$$\mathcal{A}_N := \{A : A\in\mathcal{V}_N,\ \nu(A)>0\},\qquad \mathcal{N}_N := \{A : A\in\mathcal{V}_N,\ \nu(A)=0\}. \tag{36}$$

Clearly,

$$\Big\{\sup_{t\in T}\sum_{A\in\mathcal{V}_N} t(A)Y(A)\geqslant KS\Big\}$$
$$\subset\Big\{\sup_{t\in T}\sum_{A\in\mathcal{A}_N} t(A)Y(A)\geqslant KS\Big\}\cup\bigcup_{A\in\mathcal{N}_N}\{|X|_{\mathbb{1}_A}\geqslant 1\}.$$

Since

$$\sum_{A\in\mathcal{N}_N}\mathbf{P}(|X|_{\mathbb{1}_A}\geqslant 1)=\sum_{A\in\mathcal{N}_N}\nu(A)=0, \tag{37}$$

it is sufficient to show that $\{\sup_{t\in T}\sum_{A\in\mathcal{A}_N} t(A)Y(A)\geqslant KS\}$ is 6/16-small. Observe that trivially from (33),

$$\sum_{A\in\mathcal{A}_n}\nu(A)\leqslant 2^N|T|S. \tag{38}$$

STEP 5. Denote $\bar{S} := \mathbf{E}\sup_{t\in T}\sum_{A\in\mathcal{A}_N} t(A)Y(A)$. Using (34) in the first inequality below and (38) in the second we get

$$S=\mathbf{E}\sup_{t\in T}\sum_{i\geqslant 1}|t(Y_i)|\geqslant\mathbf{E}\sup_{t\in T}\sum_{A\in\mathcal{A}_n}\Big(t(A)-\frac{1}{4^N}\Big)Y(A)=\bar{S}-\sum_{A\in\mathcal{A}_n}\frac{1}{4^N}\mathbf{E}Y(A)$$
$$=\bar{S}-\frac{1}{4^N}\sum_{A\in\mathcal{A}_N}\nu(A)\geqslant\bar{S}-\frac{2^N|T|S}{4^N}\geqslant\frac{\bar{S}}{2}$$

for N large enough. So,

$$\Big\{\sup_{t\in T}\sum_{A\in\mathcal{A}_n} t(A)Y(A)\geqslant KS\Big\}\subset\Big\{\sup_{t\in T}\sum_{A\in\mathcal{A}_n} t(A)Y(A)\geqslant K\bar{S}\Big\}$$

and it is enough to prove that the latter event is 6/16-small (note that the constant K is now different).

STEP 6. The next step is to divide each $A\in\mathcal{A}_N$ into subsets of equal measure and a "little" remainder. Let (by the definition of $\mathcal{A}_N$, and since it is finite, the numerator is strictly positive)

$$p := \frac{\min(1,S,\inf_{A\in\mathcal{A}_n}\nu(A))}{16\cdot 2^N S|T|}\in(0,1). \tag{39}$$

For each $A \in \mathcal{A}_N$ let $m(A) := \lfloor \nu(A)/p \rfloor$. Fix $A \in \mathcal{A}_N$. Since ν is atomless, there are ν-measurable disjoint sets $A_1, \ldots, A_{m(A)} \subset A$ such that $\nu(A_k) = p$. Let also $A_* := A \setminus \bigcup_{k=1}^{m(A)} A_k$. Obviously

$$\nu(A_*) \leqslant p \leqslant \frac{\nu(A)}{16 \cdot 2^N S|T|}. \tag{40}$$

For any $A \in \mathcal{A}_N$ the sets $(A_i)_{i \leqslant m(A)}$ are pairwise disjoint. Thus,

$$\forall_{A \in \mathcal{A}_n} \sum_{i=1}^{m(A)} \mathbb{1}_{Y(A_i)>0} \leqslant \sum_{i=1}^{m(A)} Y(A_i) \leqslant Y(A). \tag{41}$$

We define

$$\mathcal{B}_N := \{A_k : A \in \mathcal{A}_N,\, k \leqslant m(A)\},$$

and for each $B \in \mathcal{B}_N$,

$$t(B) := t(A) \quad \text{where } A \in \mathcal{A}_N \text{ is unique such that } B \subset A. \tag{42}$$

Since the constructed $t(A), t(B)$ are nonnegative, and using (41), we obtain

$$\bar{S} = \mathbf{E} \sup_{t \in T} \sum_{A \in \mathcal{A}_N} t(A) Y(A) \geqslant \mathbf{E} \sup_{t \in T} \sum_{B \in \mathcal{B}_N} t(B) \mathbb{1}_{Y(B)>0} =: S'.$$

By the above,

$$\Big\{ \sup_{t \in T} \sum_{A \in \mathcal{A}_N} t(A) Y(A) \geqslant K \bar{S} \Big\}$$
$$\subset \Big\{ \sup_{t \in T} \sum_{B \in \mathcal{B}_N} t(B) Y(B) \geqslant K S' \Big\} \cup \bigcup_{A \in \mathcal{A}_N} \{|X|_{\mathbb{1}_{A_*}} \geqslant 1\}.$$

From (40) and (38) we get

$$\sum_{A \in \mathcal{A}_N} \mathbf{P}(|X|_{\mathbb{1}_{A_*}} \geqslant 1) \leqslant \sum_{A \in \mathcal{A}_N} \mathbf{E}|X|_{\mathbb{1}_{A_*}} = \sum_{A \in \mathcal{A}_N} \nu(A_*) \leqslant \sum_{A \in \mathcal{A}_N} \frac{\nu(A)}{16 \cdot 2^N S|T|} \leqslant \frac{1}{16}.$$

So it is sufficient to show that the event

$$\Big\{ \sup_{t \in T} \sum_{B \in \mathcal{B}_N} t(B) Y(B) \geqslant K S' \Big\}$$

is 5/16-small.

Step 7. This step shows that the random variables $Y(B)$ can be replaced by independent Bernoulli variables. Let $n := |\mathcal{B}_N|$. Formulas (38) and (39) (and since for all $B \in \mathcal{B}_N$ we have $\nu(B) = p$) imply

$$np^2 = p \sum_{B \in \mathcal{B}_N} \nu(B) \leqslant p \sum_{A \in \mathcal{A}_N} \nu(A) \leqslant \frac{1}{16}. \tag{43}$$

Clearly,

$$\{Y(B) > 1\} = \{|X|_{\mathbb{1}_B} \geqslant 2\}.$$

So, we immediately obtain

$$\Big\{\sup_{t\in T}\sum_{B\in\mathcal{B}_N} t(B)Y(B)\geqslant KS'\Big\}$$
$$\subset\Big\{\sup_{t\in T}\sum_{B\in\mathcal{B}_N} t(B)\mathbb{1}_{\{Y(B)>0\}}\geqslant KS'\Big\}\cup\bigcup_{B\in\mathcal{B}_N}\{|X|_{\mathbb{1}_B}\geqslant 2\}.$$

Since $(Y_i)_{i\geqslant 1}$ is a Poisson point process, the random variables $(Y(B))_{B\in\mathcal{B}_N}$ are independent (we will need this property later) Poisson r.v.'s with parameter $\nu(B)=p$. Thus (recall (43)),

$$\sum_{B\in\mathcal{B}_N}\mathbf{P}(|X|_{\mathbb{1}_B}\geqslant 2)=\sum_{B\in\mathcal{B}_N}\mathbf{P}(Y(B)>1)=n(1-e^{-p}(1+p))\leqslant np^2\leqslant\frac{1}{16}.$$

We have reduced our task to proving that the following event is 1/4-small:

$$\Big\{\sup_{t\in T}\sum_{B\in\mathcal{B}_N} t(B)\mathbb{1}_{\{Y(B)>0\}}\geqslant KS'\Big\}.$$

Final Step. The random variables $(Y(B))_{B\in\mathcal{B}_N}$ are i.i.d. Poisson(p) r.v.'s (see the previous step). Thus, Theorem 2.1 ensures that the event

$$\Big\{\sup_{t\in T}\sum_{B\in\mathcal{B}_N} t(B)\mathbb{1}_{\{Y(B)>0\}}\geqslant KS'\Big\}$$

is $(1-e^{-p})$-small. Since $p\geqslant 1-e^{-p}\geqslant p/e$ (recall (39)), by Remark 3.5 we can argue that for sufficiently large K this event is $(2ep)$-small. This means that there exists a family $\mathcal{F}$ of subsets of $\mathcal{B}_N$ such that

$$\Big\{\sup_{t\in T}\sum_{B\in\mathcal{B}_N} t(B)\mathbb{1}_{\{Y(B)>0\}}\geqslant KS'\Big\}\subset\bigcup_{F\in\mathcal{F}}\bigcap_{B\in F}\{Y(B)>0\}$$

and

$$\sum_{F\in\mathcal{F}}(2ep)^{|F|}\leqslant\frac{1}{2}. \tag{44}$$

Now, for $F\in\mathcal{F}$ define

$$g_F(x):=\mathbb{1}_{\bigcup_{B\in F}}(x).$$

Note that

$$\bigcap_{B\in F}\{Y(B)>0\}\subset\Big\{\sum_{B\in F}Y(B)\geqslant|F|\Big\}=\{|X|_{g_F}\geqslant|F|\},$$

so

$$\Big\{\sup_{t\in T}\sum_{B\in\mathcal{B}_N} t(B)\mathbb{1}_{\{Y(B)>0\}}\geqslant KS'\Big\}\subset\bigcup_{F\in\mathcal{F}}\{|X|_{g_F}\geqslant|F|\}.$$

Since $\sum_{B\in F}Y(B)\sim$ Poiss$(p|F|)$, Chernoff's inequality implies (recall that

$p<1$ so that $\ln p<0$)

$$\mathbf{P}(|X|_{g_F}\geqslant |F|)=\mathbf{P}\Big(\sum_{B\in F}Y(B)\geqslant |F|\Big)\leqslant e^{|F|\ln p}\mathbf{E}e^{-\ln p\sum_{B\in F}Y(B)}$$
$$=e^{|F|\ln p}e^{p|F|(1/p-1)}\leqslant (ep)^{|F|}.$$

By recalling (44) we obtain

$$\sum_{F\in\mathcal{F}}\mathbf{P}(|X|_{g_F}\geqslant |F|)\leqslant \sum_{F\in\mathcal{F}}(ep)^{|F|}\leqslant \frac{1}{4}.$$

Thus, we have found a witness that the event

$$\Big\{\sup_{t\in T}\sum_{B\in\mathcal{B}_N}t(B)\mathbb{1}_{\{Y(B)>0\}}\geqslant KS'\Big\}$$

is 1/4-small. ∎

5. Positive empirical processes. The proof of Theorem 2.4 is similar to the one in the previous section. We want to partition E into a sequence of appropriately small sets with respect to the measure μ, argue that it is enough to investigate only one representative of each of the partition sets, and therefore reduce the problem to the selector case. The main difference is that the analogues of indicators from Step 7 are now not independent. To deal with this fact, we first prove a version of Theorem 2.1.

Let us use the notation $\binom{[n]}{m}$ for the class of all m-element subsets of $[n]$. Suppose that $Y\in\binom{[n]}{m}$ is chosen uniformly. We denote by $\mathbf{P}_Y$ and $\mathbf{E}_Y$ the probability and the expectation with respect to Y.

We aim to prove the following conditional version of Theorem 3.1.

THEOREM 5.1. *Let $\mathcal{F}$ be a family of subsets of $[n]$ which is not p-small. Suppose that for each $I\in\mathcal{F}$ we are given coefficients $(\mu_I(i))_{i\in I}$ with $\mu_I(i)\geqslant 0$ and $\sum_{i\in I}\mu_I(i)\geqslant 1$. Let Y be uniformly distributed on $\binom{[n]}{m}$ where $m\geqslant 9pn$. Then*

$$\mathbf{E}_Y\sup_{I\in\mathcal{F}}\sum_{i\in I\cap Y}\mu_I(i)\geqslant\frac{1}{10}.$$

Proof. Let X be given by (9). Then X conditionally on $\{|X|=m\}$ has the same distribution as Y (uniform). So by Lemma 3.4,

$$\mathbf{P}(Y\text{ is }1/2\text{-bad})\leqslant\sum_{t=1}^{n}\Big(4\frac{np}{m}\Big)^t\leqslant\sum_{t\geqslant 1}\Big(\frac{4}{9}\Big)^t\leqslant\frac{4}{5}.$$

Thus

$$\mathbf{E}_Y\sup_{I\in\mathcal{F}}\sum_{i\in I\cap Y}\mu_I(i)\geqslant\frac{1}{2}\mathbf{P}(Y\text{ is not }1/2\text{-bad})\geqslant\frac{1}{2}\Big(1-\frac{4}{5}\Big).\ ∎$$

Remark 5.2. Theorems 2.1 and 5.1 are in a sense presentations of the same result from two perspectives. Theorem 5.1 emphasizes the necessary relation between m and n for providing a p-small cover which we need in the application to empirical processes.

Theorem 5.3. *The family*

$$\mathcal{F} = \left\{ I \in \binom{[n]}{m} : \sup_{t\in T} \sum_{i\in I} t_i \geqslant 11\mathbf{E}_Y \sup_{t\in T} \sum_{i\in Y} t_i \right\}$$

is $\frac{m}{9n}$*-small.*

Proof. Assume for contradiction that $\mathcal{F}$ is not $\frac{m}{9n}$-small. Let

$$S := \mathbf{E}_Y \sup_{t\in T} \sum_{i\in Y} t_i.$$

For any $I \in \mathcal{F}$ we can find $t \in T$ such that for $\mu_I(i) := \frac{t_i}{11S}$, we have $\sum_{i\in I} \mu_I(i) \geqslant 1$, thus

$$\mathbf{E}_Y \sup_{I\in\mathcal{F}} \sum_{i\in I\cap Y} \mu_I(i) \leqslant \mathbf{E}_Y \sup_{t\in T} \sum_{i\in Y} \frac{t_i}{11S} = \frac{1}{11} < \frac{1}{10},$$

which contradicts Theorem 5.1. ∎

Lemma 5.4. *Let Y and Y' be uniformly distributed over $\binom{[n]}{m}$ and $\binom{[n]}{km}$ respectively, where k is a natural number such that $km \leqslant n$. Then*

$$\mathbf{E}_{Y'} \sup_{t\in T} \sum_{i\in Y'} t_i \leqslant k\mathbf{E}_Y \sup_{t\in T} \sum_{i\in Y} t_i.$$

Proof. Let π be a permutation of $[km]$, chosen uniformly at random. Let $y_1 < \cdots < y_{km}$ be the elements of Y'. Then

$$S' = \mathbf{E}_{\pi,Y'} \sup_{t\in T} \sum_{i=1}^{km} t_{y_{\pi(i)}} \leqslant \sum_{l=0}^{k-1} \mathbf{E}_{\pi,Y} \sup_{t\in T} \sum_{i=lm+1}^{(l+1)m} t_{y_{\pi(i)}} = kS. \ \blacksquare$$

Proposition 5.5. *Let $C \geqslant 1$, Y be uniformly distributed over $\binom{[n]}{m}$ and*

$$S = \mathbf{E}_Y \sup_{t\in T} \sum_{i\in Y} t_i.$$

Then the family

$$\mathcal{F} = \left\{ I \in \binom{[n]}{m} : \sup_{t\in T} \sum_{i\in I} t_i \geqslant 100CS \right\}$$

is Cm/n-small.

Proof. For any $I \in \binom{[n]}{m}$ we pick $t^I \in T$ such that

$$\sup_{t\in T} \sum_{i\in I} t_i = \sum_{i\in I} t_i^I.$$

Then

$$\mathcal{F} = \Big\{ I \in \binom{[n]}{m} : \sum_{i\in I} t_i^I \geqslant 100CS \Big\}.$$

Assume that $\mathcal{F}$ is not Cm/n-small. Let $k := \lceil 9C \rceil$. If $km > n$ then by Lemma 5.4 the family $\mathcal{F}$ is empty, which is a contradiction. Therefore we may assume that $km \leqslant n$. Take Y' uniformly distributed over $\binom{[n]}{km}$. By Theorem 5.1,

$$\mathbf{E}_{Y'} \sup_{I\in\mathcal{F}} \sum_{i\in Y'} t_i^I \geqslant 10CS. \tag{45}$$

On the other hand, by Lemma 5.4 we have

$$\mathbf{E}_{Y'} \sup_{I\in\mathcal{F}} \sum_{i\in Y'} t_i^I \leqslant \mathbf{E}_{Y'} \sup_{t\in T} \sum_{i\in Y'} t_i \leqslant kS. \tag{46}$$

Trivially (45) and (46) give a contradiction, so $\mathcal{F}$ is Cm/n-small. ■

Proof of Theorem 2.4. We proceed as in the case of infinitely divisible processes, gradually simplifying the functions $t \in T$. The proof is very similar but less technical. This is due to the fact that we are now working with the probability measure μ instead of the (potentially) infinite measure ν. For example, Step 1 in the proof of Theorem 2.3 is now superfluous. Only the last step requires a different argument. For this reason, we refer to the proof of Theorem 2.3 in a few places, instead of repeating the technical arguments.

Throughout the proof we use the notion of δ-smallness analogous to that from Definition 4.1, where we replace $|X|_g$ with $\frac{1}{d}\sum_{i=1}^d g(X_i)$. Let $S := \mathbf{E}\sup_{t\in T}\sum_{i=1}^d t(X_i)$ and let $M > 0$ be a large number to be specified later. In particular, we may assume that (since T is finite)

$$\sum_{t\in T} \mu(\{|t| \geqslant M\}) \leqslant \frac{1}{16d}. \tag{47}$$

Step 1. Fix N large enough. By repeating the argument from Step 3 of the proof of Theorem 2.3 we construct a family $\mathcal{V}_N$ of disjoint subsets of E, together with coefficients $(t(A))_{t\in T, A\in\mathcal{V}_N}$ such that

$$\bigcup_{A\in\mathcal{V}_N} A = \bigcap_{t\in T} \{x \in E : |t(x)| \leqslant M\} \tag{48}$$

and for any $A \in \mathcal{V}_N$ and $x \in A$,

$$t(A) - 4^{-N} \leqslant t(x) \leqslant t(A). \tag{49}$$

Notice the difference between (48) and (32). The reason is that μ is finite so we do not need to take care of the behavior of the measure around 0. Let

$X(A) := \sum_{i=1}^d \mathbb{1}_{\{X_i \in A\}}$. Then (48) and (49) imply that

$$\Big\{\sup_{t\in T}\sum_{i=1}^d t(X_i) > KS\Big\}$$
$$\subset \Big\{\sup_{t\in T}\sum_{A\in\mathcal{V}_N} t(A)X(A) > KS\Big\} \cup \bigcup_{t\in T}\Big\{\sum_{i=1}^d \mathbb{1}_{|t(X_i)|\geqslant M} \geqslant 1\Big\}.$$

Using Bernoulli's inequality and (47) yields

$$\begin{aligned}\sum_{t\in T}\mathbf{P}\Big(\Big\{\sum_{i=1}^d \mathbb{1}_{|t(X_i)|\geqslant M} \geqslant 1\Big\}\Big) &= \sum_{t\in T}(1-\mu(\{|t|\leqslant M\})^d)\\ &\leqslant d\sum_{t\in T}\mu(\{|t|>M\}) \leqslant \frac{1}{16}.\end{aligned}$$

So it is sufficient to show that the set $\{\sup_{t\in T}\sum_{A\in\mathcal{V}_N} t(A)X(A) > KS\}$ is 7/16-small.

Step 2. Let $\mathcal{A}_N := \{A\in\mathcal{V}_N : \mu(A) > 0\}$ and $\mathcal{N}_N := \{A\in\mathcal{V}_N : \mu(A) = 0\}$. Then

$$\Big\{\sup_{t\in T}\sum_{A\in\mathcal{V}_N} t(A)X(A) > KS\Big\}$$
$$\subset \Big\{\sup_{t\in T}\sum_{A\in\mathcal{A}_N} t(A)X(A) > KS\Big\} \cup \bigcup_{A\in\mathcal{N}_N}\Big\{\sum_{i=1}^d \mathbb{1}_A(X_i) \geqslant 1\Big\}.$$

By the definition of $\mathcal{N}_N$ we have $\sum_{A\in\mathcal{N}_N}\mathbf{P}(\sum_{i=1}^d \mathbb{1}_A(X_i)\geqslant 1) = 0$. Also, by arguing as in Step 5 of the proof of Theorem 2.3,

$$S \geqslant \frac{\mathbf{E}\sup_{t\in T}\sum_{A\in\mathcal{A}_N} t(A)X(A)}{2} =: \frac{\bar{S}}{2}.$$

Thus, it is enough to show that the following set is 7/16-small:

$$\Big\{\sup_{t\in T}\sum_{A\in\mathcal{A}_N} t(A)X(A) > K\bar{S}\Big\}.$$

Step 3. Let

$$p := \frac{\min(1, \inf_{A\in\mathcal{A}_N}\mu(A))}{16d^2} > 0. \tag{50}$$

We can divide each $A\in\mathcal{A}_N$ into disjoint sets $A_1,\ldots,A_{m(A)},A_*$, where $m(A) = \lfloor \mu(A)/p\rfloor$, $\mu(A_1) = \cdots = \mu(A_{m(A)}) = p$ and $\mu(A_*) < p$ (we recall that μ is atomless). Let (as before)

$$\mathcal{B}_N := \{A_k : A\in\mathcal{A}_N,\ k\leqslant m(A)\}.$$

From the definition of p (and since $\mathcal{A}_N$ consists of disjoint sets)

$$\sum_{A\in\mathcal{A}_N}\mathbf{P}\Big(\sum_{i=1}^{d}\mathbb{1}_{A_*}(X_i)\geqslant 1\Big)\leqslant d\sum_{A\in\mathcal{A}_N}\mu(A_*)\leqslant d\sum_{A\in\mathcal{A}_N}\frac{\mu(A)}{16d^2}\leqslant\frac{1}{16}. \tag{51}$$

Similar to Step 6 of the proof of Theorem 2.3, we show that the sets A^* are negligible, so it is sufficient to show that the event

$$\Big\{\sup_{t\in T}\sum_{B\in\mathcal{B}_N}t(B)X(B)\geqslant K\bar{S}\Big\} \tag{52}$$

is 5/16-small, where $(t(B))_{B\in\mathcal{B}_N}$ are defined as in (42).

Step 4. In this step we replace the variables $(X(B))_{B\in\mathcal{B}_N}$ by 0-1 variables (but not independent). By construction, the sets $(B)_{B\in\mathcal{B}_N}$ are disjoint and $\mu(B)=p$ for any $B\in\mathcal{B}_N$. So, if $n=|\mathcal{B}_N|$ we have

$$np=\sum_{B\in\mathcal{B}_N}\mu(B)\leqslant 1. \tag{53}$$

Now, observe that

$$\{X(B)>1\}=\Big\{\sum_{i=1}^{d}\mathbb{1}_B(X_i)\geqslant 2\Big\},$$

so (as before) for

$$\mathcal{E}:=\Big\{\sup_{t\in T}\sum_{B\in\mathcal{B}_N}t(B)X(B)\geqslant K\bar{S}\Big\}\cap\bigcap_{B\in\mathcal{B}_N}\{X(B)\in\{0,1\}\}$$

we have

$$\Big\{\sup_{t\in T}\sum_{B\in\mathcal{B}_N}t(B)X(B)\geqslant K\bar{S}\Big\}\subset\mathcal{E}\cup\bigcup_{B\in\mathcal{B}_N}\Big\{\sum_{i=1}^{d}\mathbb{1}_B(X_i)\geqslant 2\Big\}.$$

Since $(1-p)^d\geqslant 1-pd$, we obtain

$$\begin{aligned}\sum_{B\in\mathcal{B}_N}\mathbf{P}\Big(\sum_{i=1}^{d}\mathbb{1}_B(X_i)\geqslant 2\Big)&=n(1-(1-p)^d-dp(1-p)^{d-1})\\&\leqslant np^2d^2\leqslant 1/16,\end{aligned} \tag{54}$$

where the last inequality follows from (53) and (50). By the same argument as in Step 6 of the proof of Theorem 2.3,

$$\bar{S}=\mathbf{E}\sup_{t\in T}\sum_{A\in\mathcal{A}_N}t(A)X(A)\geqslant\mathbf{E}\sup_{t\in T}\sum_{B\in\mathcal{B}_N}t(B)X(B)=:S'.$$

So it is enough to show that the following event is 1/4-small:

$$\mathcal{E}':=\Big\{\sup_{t\in T}\sum_{B\in\mathcal{B}_N}t(B)X(B)\geqslant KS'\Big\}\cap\bigcap_{B\in\mathcal{B}_N}\{X(B)\in\{0,1\}\}. \tag{55}$$

Final Step. Recall that $\binom{\mathcal{B}_N}{d}$ is the family of d-element subsets of $\mathcal{B}_N$. Let $\mathcal{Y}$ be uniformly chosen from $\binom{\mathcal{B}_N}{d}$ and let $\mathcal{X} \in 2^{\mathcal{B}_N}$ be a random set defined as $\mathcal{X} := \{B \in \mathcal{B}_N : X(B) \geqslant 1\}$. We claim that

$$\frac{8}{7}\mathbf{E}\sup_{t\in T}\sum_{B\in\mathcal{B}_N} t(B)X(B) = \frac{8}{7}\mathbf{E}\sup_{t\in T}\sum_{B\in\mathcal{X}} t(B)X(B) \geqslant \mathbf{E}\sup_{t\in T}\sum_{B\in\mathcal{Y}} t(B). \tag{56}$$

By (51), (54) and the union bound,

$$\begin{aligned}\mathbf{P}&\left(|\mathcal{X}| \neq d \text{ or } \exists_{B\in\mathcal{B}_N}\ X(B) \geqslant 2\right)\\ &\leqslant \sum_{A\in\mathcal{A}_n}\mathbf{P}\Big(\sum_i \mathbb{1}_{A_*}(X_i) \geqslant 1\Big) + \sum_{B\in\mathcal{B}_N}\mathbf{P}\Big(\sum_i \mathbb{1}_B(X_i) \geqslant 2\Big) \leqslant \frac{1}{8}.\end{aligned}$$

Thus,

$$\begin{aligned}\mathbf{E}&\sup_{t\in T}\sum_{B\in\mathcal{X}} t(B)X(B)\\ &\geqslant \frac{7}{8}\mathbf{E}\Big(\sup_{t\in T}\sum_{B\in\mathcal{X}} t(B)X(B) \mid |\mathcal{X}| = d,\ \forall_{B\in\mathcal{B}_N}\ X(B) \in \{0,1\}\Big).\end{aligned}$$

Now, if we take any $\mathcal{I} = \{B_1,\ldots,B_d\} \in \binom{\mathcal{B}_N}{d}$ and any other $\mathcal{J} \in \binom{\mathcal{B}_N}{d}$, we see that

$$\mathbf{P}(\mathcal{X} = \mathcal{I}) = \sum_{\pi\in\mathrm{Perm}([d])}\mathbf{P}(X_{\pi(1)} \in B_1,\ldots,X_{\pi(d)} \in B_d) = d!p^d = \mathbf{P}(\mathcal{X} = \mathcal{J}).$$

Thus $\mathcal{X}$ conditioned on $\{|\mathcal{X}| = d,\ X(B) \in \{0,1\} : B \in \mathcal{B}_N\}$ has the same distribution as $\mathcal{Y}$, which concludes the argument for (56).

Now, by Proposition 5.5 there exists a $4ed/n$-small family $\mathcal{G}$ consisting of the sets $G = \{B_1,\ldots,B_{|G|} : B_i \in \mathcal{B}_N\}$ such that

$$\begin{aligned}&\left\{\mathcal{I} \in \binom{\mathcal{B}_N}{d} : \sup_{t\in T}\sum_{B\in\mathcal{I}} t(B) \geqslant 400e\,\mathbf{E}\sup_{t\in T}\sum_{B\in\mathcal{B}_N} t(B)X(B)\right\}\\ &\qquad\subset \left\{\mathcal{I} \in \binom{\mathcal{B}_N}{d} : \sup_{t\in T}\sum_{B\in\mathcal{I}} t(B) \geqslant 400e\,\mathbf{E}\sup_{t\in T}\sum_{B\in\mathcal{B}_N} t(B)X(B)\right\}\\ &\qquad\subset \bigcup_{G\in\mathcal{G}}\bigcap_{B\in G}\{\mathcal{I} : B \in \mathcal{I}\}\end{aligned}$$

and $\sum_{G\in\mathcal{G}}(4ed/n)^{|G|} \leqslant 1/2$. Then clearly (remembering the definition (55) of $\mathcal{E}'$)

$$\mathcal{E}' \subset \bigcup_{G\in\mathcal{G}}\{X(B) = 1,\ B \in G\} \cap \{X(B) \in \{0,1\},\ B \in \mathcal{B}_N\}.$$

We define $g_G(x) := \sum_{B\in G} \mathbb{1}_B(x)$ and see that

$$\{X(B) = 1,\, B \in G\} \cap \{X(B) \in \{0,1\},\, B \in \mathcal{B}_N\} \subset \Big\{\sum_{i=1}^{d} g_G(X_i) \geqslant |G|\Big\}.$$

Thus, it is enough to show that

$$\sum_{G\in\mathcal{G}} \mathbf{P}\Big(\sum_{i=1}^{d} g_G(X_i) \geqslant |G|\Big) \leqslant 1/4.$$

The variables $(g_G(X_i))_{i\leqslant d}$ are independent Bernoulli random variables with success probability $|G|p$. Hence,

$$\mathbf{P}\Big(\sum_{i\leqslant d} g_G(X_i) \geqslant |G|\Big) \leqslant \binom{d}{|G|}(p|G|)^{|G|} \leqslant (edp)^{|G|},$$

so the result follows for sufficiently small p. ∎

Appendix. The aim of this section is to show how the function F of Definition 3.7 can be used in a more general setting. For the sake of clarity, and since the proof steps are conceptually not far from those presented in Section 3, we will not go into every detail of the argument.

Consider independent random variables $X_1,\dots,X_d$ valued in $[n]$ with the same probability distribution, that is, $\mathbf{P}(X_i = j) = p(j) > 0$, $j \in [n]$. Let $T \subset \mathbb{R}_+^d$ and let $f : [n] \to \mathbb{R}_+$. We assume that T is permutationally invariant, i.e., for any permutation $\sigma : [d] \to [d]$ and any $t \in T$ we have $(t_{\sigma(i)})_{i=1}^d \in T$. We are interested in

$$S(T) := \mathbf{E}\sup_{t\in T} \sum_{i=1}^{d} t_i f(X_i) = \mathbf{E}\sum_{i=1}^{d} t_i^X f(X_i),$$

where for $x \in [n]^d$, $t^x \in T$ is defined as the point at which $\sup_{t\in T}\sum_{i=1}^d t_i f(x_i)$ is attained and t^X is its random counterpart. Since T is permutationally invariant, we see that for any permutation σ and $X_\sigma = (X_{\sigma(i)})_{i=1}^d$, we can choose $t^{X_\sigma} = t^X(\sigma)$, where $t_i^X(\sigma) = t_{\sigma(i)}^X$ for $i \in [d]$.

For $x \in [n]^d$ and $1 \leqslant j \leqslant n$ let $s_d^x(j) := |\{i \in [d] : x_i = j\}|$ and let $s_d^x := (s_d^x(j))_{j=1}^n$ be the multiset generated by x. In general, we refer to an element of $[d]^n = \{0,1,\dots,d\}^n$ as a *multiset*. Let $\mathcal{M}_d := \{s = (s(j))_{j=1}^n \in \{0,1,\dots,d\}^n : \sum_{j=1}^n s(j) = d\}$, which we refer to as the collection of multisets *of length d*. Observe that the multiset s_d^X generated by $X = (X_1,\dots,X_d)$ is the random element of $\mathcal{M}_d$ distributed according to

$$\mathbf{P}(s_d^X = s) = d!\prod_{j=1}^{n} \frac{p(j)^{s(j)}}{s(j)!}. \tag{57}$$

We will need the following coefficients:

$$\mu_j^{s_d^x} := \frac{1}{d! s_d^x(j)} \sum_\sigma \sum_{i=1}^d t^x_{\sigma(i)} f(j) \mathbb{1}_{x_{\sigma(i)}=j}, \qquad j = 1, \ldots, n.$$

The following simple observation is a starting point for our result.

LEMMA 5.6. *We have*

$$S(T) = \mathbf{E} \sum_{j=1}^n \mu_j^{s_d^X} s_d^X(j).$$

Proof. Indeed,

$$\begin{aligned} S(T) &= \mathbf{E} \sum_{i=1}^d t^X f(X_i) = \mathbf{E} \sum_{j=1}^n \frac{1}{s_d^X(j)} \sum_{i=1}^d t_i^X f(j) \mathbb{1}_{X_i=j} s_d^X(j) \\ &= \mathbf{E} \sum_{j=1}^n \frac{1}{d! s_d^X(j)} \sum_\sigma \sum_{i=1}^d t^X_{\sigma(i)} f(j) \mathbb{1}_{X_{\sigma(i)}=j} s_d^X(j) = \mathbf{E} \sum_{j=1}^n \mu_j^{s_d^X} s_d^X(j). \ \blacksquare \end{aligned}$$

DEFINITION 5.7. Let $\mathcal{H}, \mathcal{G} \subset [n]^d$ be families of multisets. We say that $\mathcal{G}$ is a cover of $\mathcal{H}$ if for every multiset $s \in \mathcal{H}$ there exists $w \in \mathcal{G}$ such that for $1 \leqslant j \leqslant n$, $w(j) \leqslant s(j)$.

DEFINITION 5.8. We say that a family $\mathcal{H}$ of multisets is *δ-small* if there exists a cover $\mathcal{G}$ of $\mathcal{H}$ which satisfies

$$\sum_{w \in \mathcal{G}} A(w) \leqslant \delta, \quad \text{where} \quad A(w) := \frac{\prod_{j=1}^n (edp(j))^{w(j)}}{\prod_{j=1}^n w(j)!}.$$

THEOREM 5.9. *There exists a constant $K > 0$ such that the family*

$$\tilde{\mathcal{F}} = \Big\{ s_d^x \in \mathcal{M}_d : \sum_{j=1}^n \mu_j^{s_d^x} s_d^x(j) \geqslant K S(T) \Big\}$$

is $1/2$-small.

The following result shows that $A(w)$ is a natural quantity, which leads to the characterization of the small cover of $\tilde{\mathcal{F}}$ in terms of simple witnessing events, which is in line with the previous results (Theorems 2.3 and 2.4). See also the discussion in [10, p. 1310].

THEOREM 5.10. *Let $\mathcal{G}$ be the cover of $\tilde{\mathcal{F}}$ from Theorem 5.9. For fixed $w \in \mathcal{G}$, consider the event*

$$E = \{ s_d^X \in \mathcal{M}_d : w(j) \leqslant s_d^X(j), 1 \leqslant j \leqslant n \}.$$

Then there is a function g_w and a number u_w with $E \subset \{\sum_{i=1}^d g_w(X_i) \geqslant u_w\}$, *where* $X=(X_1,\ldots,X_d)$ *generates* $s_d^X\in E$*. Moreover,*

$$\sum_{w\in\mathcal{G}}\mathbf{P}\Big(\sum_{i=1}^d g_w(X_i)\geqslant u_w\Big)\leqslant\frac{1}{2}.$$

Proof. Suppose that $w\in\mathcal{G}$. Define

$$g_w(j)=\log\left(1+\frac{w(j)}{dp(j)}\right)\quad\text{and}\quad u_w=\sum_{j=1}^n w(j)\log\left(1+\frac{w(j)}{dp(j)}\right).$$

Then X generating $s_d^X\in E$ satisfies

$$\begin{aligned}\sum_{i=1}^d g_w(X_i)&=\sum_{j=1}^n s_d^X(j)g_w(j)=\sum_{j=1}^n s_d^X(j)\log\left(1+\frac{w(j)}{dp(j)}\right)\\&\geqslant\sum_{j=1}^n w(j)\log\left(1+\frac{w(j)}{dp(j)}\right)=u_w,\end{aligned}$$

so $E\subset\{\sum_{i=1}^d g_w(X_i)\geqslant u_w\}$. Now, we show that

$$\mathbf{P}\Big(\sum_{i=1}^d g_w(X_i)\geqslant u_w\Big)\leqslant A(w).$$

Suppose that $\sum_{j=1}^n w(j)=t$. By the definition of the cover we know that $t\leqslant d$. By Markov's inequality, we obtain

$$\mathbf{P}\Big(\sum_{i=1}^d g_w(X_i)\geqslant u_w\Big)=\mathbf{P}\Big(\exp\Big(\sum_{i=1}^d g_w(X_i)\Big)\geqslant e^{u_w}\Big)\leqslant[\mathbf{E}\exp(g_w(X_1))]^d e^{-u_w}.$$

Now,

$$\begin{aligned}e^{-u_w}&=e^{-\sum_{j=1}^n w(j)\log(1+\frac{w(j)}{dp(j)})}=\prod_{j=1}^n e^{-w(j)\log(1+\frac{w(j)}{dp(j)})}\\&=\prod_{j=1}^n\left(\frac{dp(j)}{dp(j)+w(j)}\right)^{w(j)}\leqslant\prod_{j=1}^n\left(\frac{dp(j)}{w(j)}\right)^{w(j)}\end{aligned}$$

and

$$\begin{aligned}\mathbf{E}\exp(g_w(X_1))&=\mathbf{E}\exp\left(\log\left(1+\frac{w(X_1)}{dp(j)}\right)\right)=\mathbf{E}\left(1+\frac{w(X_1)}{dp(j)}\right)\\&=1+\sum_{j=1}^n\frac{w(j)}{dp(j)}p(j)=1+\frac{t}{d}.\end{aligned}$$

Putting the above together gives

$$\mathbf{P}\Big(\sum_{i=1}^{d} g_w(X_i) \geqslant u_w\Big) \leqslant \Big(1+\frac{t}{d}\Big)^d \prod_{j=1}^{n}\Big(\frac{dp(j)}{w(j)}\Big)^{w(j)} \leqslant e^t \prod_{j=1}^{n}\Big(\frac{dp(j)}{w(j)}\Big)^{w(j)}$$
$$= \prod_{j=1}^{n}\Big(\frac{edp(j)}{w(j)}\Big)^{w(j)} \leqslant A(w),$$

since $w(j)^{w(j)} \geqslant w(j)!$. So, by Theorem 5.9,

$$\sum_{w\in\mathcal{G}} \mathbf{P}\Big(\sum_{i=1}^{d} g_w(X_i) \geqslant u_w\Big) \leqslant 1/2. \blacksquare$$

To simplify notation, we write S_d for s_d^X. Consider S_{Cd} for some $C > 0$, which we define as the multiset of length Cd generated by $Y^1, \ldots, Y^C$, independent copies of X. Clearly, S_{Cd} is distributed on $\mathcal{M}_{Cd}$ according to

$$\mathbf{P}(S_{Cd} = s) = (Cd)! \prod_{j=1}^{n} \frac{p(j)^{s(j)}}{s(j)!}.$$

We aim to prove the following.

THEOREM 5.11. *Suppose that $\mathcal{F} \subset \mathcal{M}_d$ is not $1/2$-small and that for every $s \in \mathcal{F}$ we are given coefficients $\mu_j^s \geqslant 0$ satisfying $\sum_{j=1}^n \mu_j^s s(j) \geqslant 1$. Then there exists $C > 0$ such that*

$$\mathbf{E}\sup_{s\in\mathcal{F}} \sum_{j=1}^{n} \mu_j^s S_{Cd}(j) \geqslant \frac{1}{4},$$

where S_{Cd} is distributed as above.

Proof of Theorem 5.9. Suppose for contradiction that $\tilde{\mathcal{F}}$ is not $1/2$-small. Let

$$\tilde{S}(T) = \mathbf{E} \sup_{s_d^x\in\tilde{\mathcal{F}}} \sum_{j=1}^{n} \mu_j^{s_d^x} S_{Cd}(j).$$

Consider the coefficients $\tilde{\mu}_j^x := \frac{\mu_j^{s_d^x}}{KS(T)}$. Then, for $s_d^x \in \tilde{\mathcal{F}}$, we have $\sum_{j=1}^n \tilde{\mu}_j^x s_d^x(j) \geqslant 1$. Now, by the argument similar to that for Lemma 5.4, and by Theorem 5.11, we have

$$CS(T) \geqslant \tilde{S}(T) = KS(T)\,\mathbf{E} \sup_{s_d^x\in\tilde{\mathcal{F}}} \sum_{j=1}^{n} \tilde{\mu}_j^x S_{Cd}(j) \geqslant \frac{1}{4} KS(T),$$

which is a contradiction for $K > 4C$. $\blacksquare$

From now on, we fix the family $\mathcal{F}$ from Theorem 5.11. We need the following definition.

Definition 5.12. Fix $C \in \mathbb{N}$ and $c \in (0,1)$. We say $s \in \mathcal{M}_{Cd}$ is *bad* if

$$\sup_{r\in\mathcal{F}} \sum_{j=1}^{n} \mu^r(j)s(j) < c.$$

The proof of Theorem 5.11 is based on the following result which is an analogue of the pivotal Lemma 3.4.

Lemma 5.13. *Let $c = 1/2$. There exists $C > 0$ such that*

$$\mathbf{P}(S_{Cd} \textit{ is bad}) \leqslant \frac{1}{2}.$$

Proof of Theorem 5.11. Conditioning on S_{Cd} being not bad gives

$$\mathbf{E}\sup_{s\in\mathcal{F}} \sum_{j=1}^{n} \mu_j^s S_{Cd}(j) \geqslant \frac{1}{2}\mathbf{P}(S_{Cd} \text{ is not bad}) \geqslant \frac{1}{4},$$

where in the last inequality we use Lemma 5.13. ∎

The rest of this section is devoted to the proof of Lemma 5.13. Observe that without loss of generality we can normalize μ_j^s for every $s \in \mathcal{F}$ so that $\sum_{j=1}^{n} \mu_j^s s(j) = 1$. In particular, $0 \leqslant \mu_j^s \leqslant 1$.

Lemma 5.14. *Fix $r \in \mathcal{F}$ and $s \in \mathcal{M}_{Cd}$ which is bad. Then there exists $\varepsilon(r,s) \in [0,1]$ such that*

$$\sum_{j=1}^{n} \mathbb{1}_{\{\mu_j^r > \varepsilon(r,s)\}} s(j) \leqslant c \sum_{j=1}^{n} \mathbb{1}_{\{\mu_j^r > \varepsilon(r,s)\}} r(j). \tag{58}$$

Proof. Define

$$F_{r,s}(\varepsilon) := \sum_{j=1}^{n} (\mu_j^r \wedge \varepsilon) s(j) - c \sum_{j=1}^{n} (\mu_j^r \wedge \varepsilon) r(j). \tag{59}$$

Obviously, $F_{r,s}(0) = 0$ and since s is bad we have

$$F_{r,s}(1) = \sum_{j=1}^{n} \mu_j^r s(j) - c \sum_{j=1}^{n} \mu_j^r r(j) < c - c = 0.$$

Define $\varepsilon(r,s) := \sup\{\varepsilon \in [0,1) : F_{r,s}(\varepsilon) \geqslant 0\}$. Following the proof of Lemma 3.6, we check that $\varepsilon(r,s)$ satisfies (58). ∎

Definition 5.15. For fixed $r \in \mathcal{F}$ and $s \in \mathcal{M}_{Cd}$ bad, we define $\varepsilon(r,s)$ as the largest number such that $F_{r,s}(\varepsilon(r,s)) \geqslant 0$ (recall (59)), i.e. the largest number for which

$$\sum_{j=1}^{n} \mathbb{1}_{\{\mu_j^r > \varepsilon(r,s)\}} s(j) \leqslant c \sum_{j=1}^{n} \mathbb{1}_{\{\mu_j^r > \varepsilon(r,s)\}} r(j). \tag{60}$$

Definition 5.16 (Witness). Fix $r \in \mathcal{F}$ and $s \in \mathcal{M}_{Cd}$ bad. Let $A(r,s) := \{j \in [n] : \mu_j^r > \varepsilon(r,s)\}$ and

$$j(r,s) := \sum_{j \in A(r,s)} r(j).$$

We say that $\hat{r} \in \mathcal{F}$ is *admissible* for (r,s) if

$$\forall_{j \in A(\hat{r},s)}\, \hat{r}(j) - s(j) \leqslant r(j). \tag{61}$$

Let

$$t(\hat{r},s) := \sum_{j \in A(\hat{r},s)} (\hat{r}(j) - s(j))_+ = \sum_{j \in A(\hat{r},s)} \max(\hat{r}(j) - s(j), 0).$$

We choose from all $\hat{r} \in \mathcal{F}$ admissible for (r,s) one with the smallest $j(\hat{r},s)$ and then with the smallest $t(\hat{r},s)$. We denote any chosen element by r_* and call it a *witness* for (r,s) (with a small abuse of notation).

Remark 5.17. In the same way as for (21), we deduce from Lemma 5.14 that

$$j(r_*,s) \geqslant t(r_*,s) \geqslant (1-c)j(r_*,s).$$

Before we present the cover of $\mathcal{F}$ we need some notation. For two multisets $x, y \in [n]^d$ we write $x+y$ for the multiset $(x(j)+y(j))_j$, $x-y$ for $(\max\{x(j)-y(j), 0\})_j$, and for $A \subset [n]$ we write $x\mathbb{1}_A$ for the multiset with entries $x(j)$ if $j \in A$ and 0 otherwise. We also write $x \leqslant y$ if $x(j) \leqslant y(j)$ for every j.

We construct the cover of $\mathcal{F}$ with multisets $(r_* - s)\mathbb{1}_{A(r_*,s)}$, where we first fix bad $s \in \mathcal{M}_{Cd}$ and then we proceed with the construction of r_* as in Definition 5.16. These are analogues of the sets $W(I,X) \setminus X$ of Section 3. Define

$$\mathcal{G}(s) := \{(r_* - s)\mathbb{1}_{A(r_*,s)} : r \in \mathcal{F}\}. \tag{62}$$

We see that $\mathcal{G}(s)$ is a cover of $\mathcal{F}$, because if $j \in A(s, r_*)$, then

$$(r_*(j) - s(j))\mathbb{1}_{A(r_*,s)} \leqslant r(j)$$

by (61) and otherwise

$$(r_*(j) - s(j))\mathbb{1}_{A(r_*,s)} = 0 \leqslant r(j).$$

Lemma 5.18. *Let $s, s' \in \mathcal{M}_{Cd}$ be bad and let $r, r' \in \mathcal{F}$. Suppose that r_* is a witness of (r,s) and that r'_* is a witness of (r',s'). Assume that*

(a) $\tilde{j} := j(r,s) = j(r',s')$,
(b) $z := s + (r_* - s)\mathbb{1}_{A(r_*,s)} = s' + (r'_* - s')\mathbb{1}_{A(r'_*,s')}$,
(c) $t := t(r,s) = t(r',s')$.

Then $\varepsilon(r_*,s') \geqslant \varepsilon(r_*,s)$.

Proof. We have to prove that $F_{r_*,s'}(\varepsilon(r_*,s)) \geqslant 0$ (recall the definition (59)). Since $\varepsilon(r_*,s')$ is the largest number for which $F_{r_*,s'}$ is nonnegative, we conclude that $\varepsilon(r_*,s') \geqslant \varepsilon(r_*,s)$. By the definition of $A(r_*,s)$ and by adding and subtracting $s(j)$ we obtain

$$\begin{aligned}
\sum_{j=1}^{n}(\mu_j^{r_*} \wedge \varepsilon(r_*,s))s'(j) &= \varepsilon(r_*,s) \sum_{j\in A(r_*,s)} s'(j) + \sum_{j\in A^c(r_*,s)} \mu_j^{r_*} s'(j) \\
&= \varepsilon(r_*,s) \sum_{j\in A(r_*,s)} (s'(j)-s(j)) + \varepsilon(r_*,s) \sum_{j\in A(r_*,s)} s(j) \\
&\quad + \sum_{j\in A^c(r_*,s)} \mu_j^{r_*}(s'(j)-s(j)) + \sum_{j\in A^c(r_*,s)} \mu_j^{r_*} s(j) \\
&= \sum_{j=1}^{n}(\mu_j^{r_*} \wedge \varepsilon(r_*,s))s(j) + \varepsilon(r_*,s) \sum_{j\in A(r_*,s)} (s'(j)-s(j)) \\
&\quad + \sum_{j\in A^c(r_*,s)} \mu_j^{r_*}(s'(j)-s(j)) \\
&= \mathrm{I} + \mathrm{II} + \mathrm{III}.
\end{aligned}$$

Notice that $\mathrm{I} \geqslant c\sum_{j=1}^{n}(\varepsilon(r_*,s) \wedge \mu_j^{r_*})r_*(j)$, so it is enough to show that $\mathrm{II}+\mathrm{III} \geqslant 0$. We have

$$\begin{aligned}
\mathrm{II} &= \varepsilon(r_*,s) \sum_{j\in A(r_*,s)\cap A(r_*',s')} (s'(j)-s(j)) \\
&\quad + \varepsilon(r_*,s) \sum_{j\in A(r_*,s)\cap A^c(r_*',s')} (s'(j)-s(j)) \\
&= A + B.
\end{aligned}$$

Now, assumption (b) can be written as

$$z = s\mathbb{1}_{A^c(r_*,s)} + (r_* \vee s)\mathbb{1}_{A(r_*,s)} = s'\mathbb{1}_{A^c(r_*',s')} + (r_*' \vee s')\mathbb{1}_{A(r_*',s')}.$$

Observe that for $j \in A(r_*,s) \cap A(r_*',s')$, by (b) we have

$$r_*'(j) \vee s'(j) = r_*(j) \vee s(j),$$

so we can write

$$A = \varepsilon(r_*,s) \sum_{j\in A(r_*,s)\cap A(r_*',s')} (s'(j) - (r_*'(j) \vee s'(j)) + (r_*(j) \vee s(j)) - s(j)).$$

For $j \in A(r_*,s) \cap A^c(r_*',s')$, by (b) we have $s'(j) = r_*(j) \vee s(j)$, so

$$B = \varepsilon(r_*,s) \sum_{j\in A(r_*,s)\cap A^c(r_*',s')} ((r_*(j) \vee s(j)) - s(j)).$$

Therefore,

$$\begin{aligned}\mathrm{II} &= A+B\\ &= \varepsilon(r_*,s)\sum_{j\in A(r_*,s)}((r_*(j)\vee s(j))-s(j))\\ &\quad+\varepsilon(r_*,s)\sum_{j\in A(r_*,s)\cap A(r'_*,s')}(s'(j)-(r'_*(j)\vee s'(j)))\\ &= \varepsilon(r_*,s)\sum_{j\in A(r_*,s)}(r_*(j)-s(j))_+-\varepsilon(r_*,s)\sum_{j\in A(r_*,s)\cap A(r'_*,s')}(r'_*(j)-s'(j))_+.\end{aligned}$$

Similarly, by noticing that $s(j)=s'(j)$ whenever $j\in A^c(r_*,s)\cap A^c(r'_*,s')$, by (b) we get

$$\begin{aligned}\mathrm{III} &= \sum_{j\in A^c(r_*,s)}\mu_j^{r_*}(s'(j)-s(j)) = \sum_{j\in A^c(r_*,s)\cap A^c(r'_*,s')}\mu_j^{r_*}(s'(j)-s(j))\\ &\quad+\sum_{j\in A^c(r_*,s)\cap A(r'_*,s')}\mu_j^{r_*}(s'(j)-s(j))\\ &= \sum_{j\in A^c(r_*,s)\cap A(r'_*,s')}\mu_j^{r_*}(s'(j)-s(j)).\end{aligned}$$

On the set $A^c(r_*,s)\cap A(r'_*,s')$, by (b) we have $s(j)=r'_*(j)\vee s'(j)$, so

$$\begin{aligned}\mathrm{III} &= \sum_{j\in A^c(r_*,s)\cap A(r'_*,s')}\mu_j^{r_*}(s'(j)-s(j))\\ &= \sum_{j\in A^c(r_*,s)\cap A(r'_*,s')}\mu_j^{r_*}(s'(j)-(r'_*(j)\vee s'(j)))\\ &= -\sum_{j\in A^c(r_*,s)\cap A(r'_*,s')}\mu_j^{r_*}(r'_*(j)-s'(j))_+.\end{aligned}$$

Finally,

$$\begin{aligned}\mathrm{II}+\mathrm{III} &= \varepsilon(r_*,s)\sum_{j\in A(r_*,s)}(r_*(j)-s(j))_+\\ &\quad-\varepsilon(r_*,s)\sum_{j\in A(r_*,s)\cap A(r'_*,s')}(r'_*(j)-s'(j))_+\\ &\quad-\sum_{j\in A^c(r_*,s)\cap A(r'_*,s')}\mu_j^{r_*}(r'_*(j)-s'(j))_+.\end{aligned}$$

For $j\in A^c(r_*,s)$, $\mu_j^{r_*}\leqslant\varepsilon(r_*,s)$. Hence, by (c),

$$\begin{aligned}\mathrm{II}+\mathrm{III} &\geqslant \varepsilon(r_*,s)\sum_{j\in A(r_*,s)}(r_*(j)-s(j))_+-\varepsilon(r_*,s)\sum_{j\in A(r'_*,s')}(r'_*(j)-s'(j))_+\\ &= \varepsilon(r_*,s)t-\varepsilon(r_*,s)t=0. \ \blacksquare\end{aligned}$$

Lemma 5.19. *Under the assumptions of Lemma* 5.18,

$$(r_* - s')\mathbb{1}_{A(r_*,s')} = (r'_* - s')\mathbb{1}_{A(r'_*,s')}. \tag{63}$$

Proof. In order to prove (63) we need to prove that $\varepsilon(r_*, s') = \varepsilon(r_*, s)$. Due to Lemma 5.18 it is enough to show that $\varepsilon(r_*, s') \leqslant \varepsilon(r_*, s)$. Assume for contradiction that $\varepsilon(r_*, s') > \varepsilon(r_*, s)$. By the same argument as in the proof of Lemma 3.12, we deduce that $A(r_*, s') \subset A(r_*, s)$ and, in turn, that $j(r_*, s') < j(r_*, s)$. We have two cases to consider.

Case 1: $j \in A(r_*, s') \cap A^c(r'_*, s')$. Since $A(r_*, s') \subset A(r_*, s)$ and by (b), we have

$$r_*(j) \vee s(j) = s'(j).$$

In particular, $s'(j) \geqslant r_*(j)$, so

$$(r_*(j) - s'(j))_+ = 0 \leqslant r'_*(j).$$

Case 2: $j \in A(r_*, s') \cap A(r'_*, s')$. Again, since $A(r_*, s') \subset A(r_*, s)$ and by (b), we have

$$r_*(j) \vee s(j) = r'_*(j) \vee s'(j).$$

We have two subcases. If $s'(j) \geqslant r'_*(j)$, then $r_*(j) \vee s(j) = s'(j)$, so $r_*(j) \leqslant r_*(j) \vee s(j) = s'(j)$, and hence, as before,

$$(r_*(j) - s'(j))_+ = 0 \leqslant r'_*(j).$$

If $s'(j) < r'_*(j)$, then $r'_*(j) \geqslant r_*(j)$, which implies that

$$(r_*(j) - s'(j))_+ \leqslant (r'_*(j) - s'(j))_+ \leqslant r'_*(j).$$

Consequently, for $j \in A(r_*, s')$,

$$(r_*(j) - s'(j))_+ \leqslant r'_*(j),$$

which means that r_* is admissible for (r', s'). Moreover, $j(r_*, s') < j(r_*, s)$. But, by (a), $j(r_*, s) = j(r'_*, s')$, which contradicts the assumption that r'_* is a witness for (r', s'), i.e. that $j(r'_*, s')$ is the smallest possible, in particular, that $j(r'_*, s') \leqslant j(r_*, s')$. Hence, $\varepsilon(r_*, s') \leqslant \varepsilon(r_*, s)$, which together with Lemma 5.18 gives $\varepsilon(r_*, s') = \varepsilon(r_*, s)$.

The main consequence of the fact proved above is that $A(r_*, s') = A(r_*, s)$. To prove (63) we again consider two cases.

Case 1: $j \in A(r_*, s) \cap A^c(r'_*, s')$. Then, by (b), we have $s'(j) = s(j) \vee r_*(j)$ and, in turn,

$$(r_*(j) - s'(j))_+ \leqslant (r_*(j) \vee s(j) - s'(j))_+ = 0.$$

Case 2: $j \in A(r_*, s) \cap A(r'_*, s')$. Then, by (b), $r_*(j) \vee s(j) = r'_*(j) \vee s'(j)$ and

$$\begin{aligned}(r_*(j) - s'(j))_+ \leqslant (r_*(j) \vee s(j) - s'(j))_+ &= (r'_*(j) \vee s'(j) - s'(j))_+ \\ &= (r'_*(j) - s'(j))_+.\end{aligned}$$

So, since $A(r_*, s') = A(r_*, s)$, we can deduce that

$$(r_* - s')\mathbb{1}_{A(r_*,s')} \leqslant (r'_* - s')\mathbb{1}_{A(r_*,s')\cap A(r'_*,s')} \leqslant (r'_* - s')\mathbb{1}_{A(r'_*,s')}.$$

But r'_* provides minimal $t(r'_*, s')$, so we conclude that

$$(r_* - s')\mathbb{1}_{A(r_*,s)} = (r_* - s')\mathbb{1}_{A(r_*,s')} = (r'_* - s')\mathbb{1}_{A(r'_*,s')}. \quad \blacksquare$$

COROLLARY 5.20. *Fix a positive integer* t, $z \in \mathcal{M}_{Cd+t}$ *and bad* $s \in \mathcal{M}_{Cd}$ *such that* $s \leqslant z$. *Then, for any* $\tilde{j} \geqslant t$,

$$(64)\qquad |\{(r_* - s)\mathbb{1}_{A(r_*,s)} : r \in \mathcal{F},\ j(r_*, s) = \tilde{j},\ z = s + (r_* - s)\mathbb{1}_{A(r_*,s)},\ t(r_*, s) = t\}| \leqslant \binom{\tilde{j}}{t}.$$

Proof. Consider any pair (r, s) satisfying (64) where $s \in \mathcal{M}_{Cd}$ is bad and any other pair (r', s') satisfying (64) where $s' \in \mathcal{M}_{Cd}$ is bad. By Lemma 5.19 we have

$$(r'_* - s')\mathbb{1}_{A(r'_*,s')} = (r_* - s')\mathbb{1}_{A(r_*,s)} \leqslant r_*\mathbb{1}_{A(r_*,s)},$$

so there are at most $\binom{\tilde{j}}{t}$ such multisets. $\blacksquare$

Proof of Lemma 5.13. We are ready for the final computation. The strategy is the same as in the proof of Lemma 3.4. We want to replace bad $s \in \mathcal{M}_{Cd}$ by suitably chosen $z \in \mathcal{M}_{Cd+t}$. Fix $\tilde{j}, t, z$ from Lemma 5.18. Consider $w(r', s') := (r'_* - s')\mathbb{1}_{A(r'_*,s')}$ and $w(r, s) = (r_* - s)\mathbb{1}_{A(r_*,s)}$ for which the parameters $\tilde{j}, t, z$ are the same. Fix a single $r(\tilde{j}, t, z) := r_* 1_{A(r_*,s)} \leqslant z$. By Corollary 5.20 we know that $w(r', s') \leqslant r(\tilde{j}, t, z)$. Note that $s = z - w(r_*, s)$. We use the notation $|w| := \sum_{j=1}^{n} w(j)$ for any multiset w. Recall the cover $\mathcal{G}(s)$ (see (62)) of $\mathcal{F}$. Since we assume that $\mathcal{F}$ is not small, for any bad $s \in \mathcal{M}_{Cd}$ we have

$$\sum_{w(r,s)\in\mathcal{G}(s)} A(w(r, s)) \geqslant 1/2$$

(recall Definition 5.8). Therefore, we can write

$$\frac{1}{2}\mathbf{P}(S_{Cd} \text{ is bad}) = \frac{1}{2}\sum_{\substack{s\in\mathcal{M}_{Cd}\\ s \text{ is bad}}} \mathbf{P}(S_{Cd} = s) \leqslant \sum_{\substack{s\in\mathcal{M}_{Cd}\\ s \text{ is bad}}} \mathbf{P}(S_{Cd} = s) \sum_{w(r,s)\in\mathcal{G}(s)} A(w(r, s)).$$

The crucial step, which is a result of the above discussion and Remark 5.17, is that we have

$$\sum_{\substack{s\in\mathcal{M}_{Cd}\\ s\text{ is bad}}}\mathbf{P}(S_{Cd}=s)\sum_{w(r,s)\in\mathcal{G}(s)}A(w(r,s))$$
$$\leqslant\sum_{\tilde{j}\geqslant 1}\sum_{(1-c)\tilde{j}\leqslant t\leqslant\tilde{j}}\sum_{z\in\mathcal{M}_{Cd+t}}\sum_{\substack{w\leqslant r(\tilde{j},t,z)\\ |w|=t}}\mathbf{P}(S_{Cd}=z-w)A(w).$$

Now, by (57), Definition 5.8 and since $w\leqslant z$, we get

$$\begin{aligned}\mathbf{P}(S_{Cd}=z-w)A(w)&=\frac{(Cd)!\prod_{j=1}^{n}p(j)^{z(j)-w(j)}}{\prod_{j=1}^{n}(z(j)-w(j))!}\,\frac{(ed)^t\prod_{j=1}^{n}p(j)^{w(j)}}{\prod_{j=1}^{n}w(j)!}\\&=(Cd)!(ed)^t\prod_{j=1}^{n}\frac{p(j)^{z(j)}}{(z(j)-w(j))!w(j)!},\end{aligned}$$

where we have used $\sum_{j=1}^{n}w(j)=t$. Now, since

$$\mathbf{P}(S_{Cd+t}=z)=(Cd+t)!\prod_{j=1}^{n}\frac{p(j)^{z(j)}}{z(j)!},$$

we have

$$\prod_{j=1}^{n}\frac{p(j)^{z(j)}}{(z(j)-w(j))!w(j)!}=\frac{1}{(Cd+t)!}\mathbf{P}(S_{Cd+t}=z)\prod_{j=1}^{n}\binom{z(j)}{w(j)}.$$

Because of (60), we can bound

$$\begin{aligned}\sum_{j\in A(r_*,s)}z(j)&=\sum_{j\in A(r_*,s)}s(j)\vee r_*(j)=\sum_{j\in A(r_*,s)}\big(s(j)+(r_*(j)-s(j))_+\big)\\&\leqslant c\tilde{j}+t\leqslant\tilde{j}+t\end{aligned}$$

since $c=\frac{1}{2}$. Using the fact that $w(j)=0$ for $j\notin A(r_*,s)$, we get

$$\prod_{j=1}^{n}\binom{z(j)}{w(j)}=\prod_{j\in A(r_*,s)}\binom{z(j)}{w(j)}\leqslant\prod_{j\in A(r_*,s)}2^{z(j)}\leqslant 2^{\tilde{j}+t}.$$

Consequently,

$$\sum_{\tilde{j}\geqslant 1}\sum_{(1-c)\tilde{j}\leqslant t\leqslant\tilde{j}}\sum_{z\in\mathcal{M}_{Cd+t}}\sum_{\substack{w\leqslant r(\tilde{j},t,z)\\ |w|=t}}\mathbf{P}(S_{Cd}=z-w)A(w)$$
$$\leqslant\sum_{\tilde{j}\geqslant 1}\sum_{(1-c)\tilde{j}\leqslant t\leqslant\tilde{j}}\sum_{z\in\mathcal{M}_{Cd+t}}\sum_{\substack{w\leqslant r(\tilde{j},t,z)\\ |w|=t}}2^{\tilde{j}+t}(ed)^t\frac{(Cd)!}{(Cd+t)!}\mathbf{P}(S_{Cd+t}=z).$$

Now, by Corollary 5.20,

$$\sum_{\substack{w\leqslant r(\tilde{j},t,z)\\ |w|=t}} 1 \leqslant \binom{\tilde{j}}{t},$$

which together with the fact that $\frac{(Cd)!}{(Cd+t)!} \leqslant (Cd)^{-t}$ gives

$$\begin{aligned}
&\sum_{\tilde{j}\geqslant 1}\ \sum_{(1-c)\tilde{j}\leqslant t\leqslant \tilde{j}}\ \sum_{z\in\mathcal{M}_{Cd+t}}\ \sum_{\substack{w\leqslant r(\tilde{j},t,z)\\ |w|=t}} 2^{\tilde{j}+t}(ed)^t \frac{(Cd)!}{(Cd+t)!}\mathbf{P}(S_{Cd+t}=z)\\
&\qquad \leqslant \sum_{\tilde{j}\geqslant 1}\ \sum_{(1-c)\tilde{j}\leqslant t\leqslant \tilde{j}}\ \sum_{z\in\mathcal{M}_{Cd+t}} \binom{\tilde{j}}{t} 2^{\tilde{j}+t}(ed)^t (Cd)^{-t}\mathbf{P}(S_{Cd+t}=z)\\
&\qquad = \sum_{\tilde{j}\geqslant 1}\ \sum_{(1-c)\tilde{j}\leqslant t\leqslant \tilde{j}} \binom{\tilde{j}}{t} 2^{\tilde{j}+t}(e/C)^t,
\end{aligned}$$

where the last equality follows since we are summing the probability mass function of S_{Cd+t}. Now, by substituting $\bar{t} := \tilde{j} - t$, we get

$$\sum_{\tilde{j}\geqslant 1}\ \sum_{(1-c)\tilde{j}\leqslant t\leqslant \tilde{j}} \binom{\tilde{j}}{t} 2^{\tilde{j}+t}(e/C)^t = \sum_{\tilde{j}\geqslant 1} 2^{\tilde{j}} 2^{\tilde{j}} (e/C)^{\tilde{j}} \sum_{0\leqslant \bar{t}\leqslant c\tilde{j}} \binom{\tilde{j}}{\bar{t}} 2^{-\bar{t}}(e/C)^{-\bar{t}}.$$

We bound $\binom{\tilde{j}}{\bar{t}} \leqslant 2^{\tilde{j}}$, $2^{-\bar{t}} \leqslant 1$ and

$$\sum_{0\leqslant \bar{t}\leqslant c\tilde{j}} (e/C)^{-\bar{t}} \leqslant \frac{(C/e)^{c\tilde{j}+1}}{C/e-1}.$$

Hence,

$$\sum_{\tilde{j}\geqslant 1} 2^{\tilde{j}} 2^{\tilde{j}} (e/C)^{\tilde{j}} \sum_{0\leqslant \bar{t}\leqslant c\tilde{j}} \binom{\tilde{j}}{\bar{t}} 2^{-\bar{t}}(e/C)^{-\bar{t}} \leqslant \frac{C/e}{C/e-1} \sum_{\tilde{j}\geqslant 1} 8^{\tilde{j}} (e/C)^{(1-c)\tilde{j}}.$$

Finally, for $C = [41^2 e]$ and since $c = 1/2$ we get

$$\frac{1}{2}\mathbf{P}(S_{Cd} \text{ is bad}) \leqslant \frac{41^2}{41^2-2} \cdot \frac{8}{\sqrt{41^2-1}-8} \leqslant \frac{1}{4}.$$

Hence, $\mathbf{P}(S_{Cd} \text{ is bad}) \leqslant \frac{1}{2}$. ∎

Acknowledgements. The authors would like to thank the organizers of the “Workshop in Convexity and High-Dimensional Probability” held May 23–27, 2022 at Georgia Institute of Technology, where some progress on this paper was made.

We are grateful to the anonymous referees for their careful reading of the initial version of this paper and pointing out all the needed clarifications.

Funding. The authors were supported by National Science Center, Poland, grants 2019/35/B/ST1/04292 (W. Bednorz) and 2021/40/C/ST1/00330 (R. Meller).

References

[1] W. Bednorz and R. Martynek, *The suprema of infinitely divisible processes*, Ann. Probab. 50 (2022), 397–417.

[2] W. Bednorz, R. Martynek and R. Meller, *Small cover approach to the suprema of positive canonical processes*, arXiv:2401.14045v2 (2026).

[3] Q. Dubroff, J. Kahn and J. Park, *Note on a conjecture of Talagrand: expectation thresholds vs. fractional expectation thresholds*, arXiv:2412.00917v1 (2024).

[4] J. Kahn, B. Narayanan and J. Park, *The threshold for the square of a Hamilton cycle*, Proc. Amer. Math. Soc. 149 (2021), 3201–3208.

[5] R. Latała, *On the spectral norm of Rademacher matrices*, Trans. Amer. Math. Soc. 379 (2026), 7019–7039.

[6] R. Latała and W. Świątkowski, *Norms of randomized circulant matrices*, Electron. J. Probab. 27 (2022), art. 80, 23 pp.

[7] R. Latała, R. van Handel and P. Youssef, *The dimension-free structure of nonhomogeneous random matrices*, Invent. Math. 214 (2018), 1031–1080.

[8] N. Lord, *Binomial averages when the mean is an integer*, Math. Gaz. 9 (2010), 331–332.

[9] R. Meller, *Spectral norm of matrices with independent entries up to polyloglog*, arXiv: 2512.23673v2 (2026).

[10] J. Park and H. T. Pham, *On a conjecture of Talagrand on selector processes and a consequence on positive empirical process*, Ann. of Math. 199 (2024), 1293–1321.

[11] J. Park and H. T. Pham, *A proof of the Kahn–Kalai Conjecture*, J. Amer. Math. Soc. 37 (2024), 235–243.

[12] M. Talagrand, *Are many small sets explicitly small?* in: Proceedings of the 2010 ACM International Symposium on Theory of Computing, ACM, 2010, 13–35.

[13] M. Talagrand, *Upper and Lower Bounds for Stochastic Processes*, 2nd ed., Ergeb. Math. Grenzgeb. 60, Springer, Cham, 2021.

Witold Bednorz, Rafał Martynek
Institute of Mathematics
University of Warsaw
02-097 Warszawa, Poland
E-mail: w.bednorz@mimuw.edu.pl
r.martynek@mimuw.edu.pl

Rafał Meller
Institute of Mathematics
University of Warsaw
02-097 Warszawa, Poland
and
Institute of Mathematics
Polish Academy of Sciences
00-656 Warszawa, Poland
E-mail: r.meller@mimuw.edu.pl