Instytut Matematyczny Polskiej Akademii Nauk

Nierówności i prawa iterowanego logarytmu dla U-statystyk

Radosław Adamczak

Rozprawa doktorska napisana po kierunkiem
dr hab. Rafała Latały, prof. UW

Warszawa, styczeń 2007

Prace nad rozprawą wspomagane były przez granty
PO3A 019 30 i PO3A 012 29

Spis treści

Wstęp

Teoria U-statystyk, zapoczątkowana w latach czterdziestych XX w. pionierskimi pracami Hoeffdinga [27] i Halmosa [25], stanowi naturalne rozwinięcie, leżącej u podstaw rachunku prawdopodobieństwa, teorii sum niezależnych zmiennych losowych. Zgodnie z wprowadzoną przez Hoeffdinga definicją, U-statystykami rzędu d nazywamy zmienne losowe postaci

$$\sum_{\substack{1\leqslant i_1,\ldots,i_d\leqslant n\\ j\neq k\Rightarrow i_j\neq i_k}} h(X_{i_1},\ldots,X_{i_d}),$$

gdzie $X_1, X_2, \ldots$ to ciąg zmiennych i.i.d., zaś h jest pewną funkcją o d argumentach. W szczególności, U-statystyki rzędu 1 to sumy zmiennych i.i.d. Na przestrzeni lat definicja U-statystyk została w naturalny sposób rozszerzona, obejmując m. in. sumy jak powyższa, oparte niekoniecznie o ciąg i.i.d., czy sumy w których zamiast jednej funkcji h mamy całą rodzinę funkcji $h_{i_1,\ldots,i_d}$. Wprowadzono też tzw. U-statystyki uniezależnione. Dokładniejsze definicje zostaną zaprezentowane w następnych rozdziałach pracy, tutaj ograniczmy się tylko do stwierdzenia, że istotną rolę w badaniu U-statystyk odgrywa ich stosunkowo nieskomplikowana, algebraiczna definicja oraz związane z nią symetrie i regularność.

Badania nad U-statystykami motywowane są w dużym stopniu zastosowaniami, gdyż pojawiają się one jako kanoniczne estymatory nieobciążone, są składnikami wyższych rzędów w rozwinięciach asymptotycznych gładkich statystyk, a w przypadku zmiennych zależnych używane są do estymacji wymiaru korelacyjnego układów dynamicznych. U-statystyki mają także związki z teorią grafów losowych (np. z problemami zliczania liczby kopii danego podgrafu, występujących w grafie losowym). Istotnym aspektem teorii jest również wewnętrzna elegancja i charakterystyczna prostota sformułowania problemów, których rozwiązanie często okazuje się skomplikowane.

Teoria U-statystyk i związana z nią teoria U-procesów (odpowiednik teorii procesów empirycznych) w dużej mierze zajmuje się przeniesieniem na U-statystyki klasycznych wyników, dotyczących sum niezależnych zmiennych losowych. Dobrym przykładem jest tu już pierwsza praca Hoeffdinga [27], w której uzyskano wersję centralnego twierdzenia granicznego dla U-statystyk. Również Hoeffdingowi zawdzięczamy prawo wielkich liczb przy założeniu całkowalności funkcji h [28] oraz podstawowe nierówności eksponencjalne. Z kolei pierwsze prawo iterowanego logarytmu dla U-statystyk rzędu większego od 1 zostało udowodnione na początku lat siedemdziesiątych przez Serflinga [46]. Wspólną cechą powyższych wyników jest, że dają one jedynie warunki dostateczne na zachodzenie odpowiednich twierdzeń, przy czym warunki te są również warunkami koniecznymi jedynie w przypadku sum niezależnych zmiennych losowych (U-statystyk rzędu 1). Co więcej, twierdzenia te okazały się do pewnego stopnia nieprecyzyjne, dla przykładu centralne twierdzenie graniczne i prawo iterowanego logarytmu udowodnione zostały przy bardzo silnych normowaniach, które już w pewnych kanonicznych przypadkach dają trywialne, zerowe granice, podczas gdy zmiana normowania może prowadzić do nietrywialnych rozkładów granicznych czy niezerowej granicy w prawie iterowanego logarytmu. Bardziej precyzyjne odpowiedniki klasycznych twierdzeń rachunku prawdopodobieństwa pojawiły się na początku lat osiemdziesiątych, gdy Rubin i Vitale [44] wykazali, że przy słabszym normowaniu i dodatkowym założeniu, tzw. kanoniczności (pojęcie to zostanie zdefiniowane w dalszej części pracy), rozkłady U-statystyk dążą do rozkładów wielokrotnych całek stochastycznych względem odpowiednich procesów gaussowskich. Pojawiły sie też wersje prawa iterowanego logarytmu ze słabszym normowaniem [11]. W dalszym ciągu nie były jednak znane warunki konieczne i dostateczne dla żadnego z podstawowych twierdzeń granicznych.

Przełom w badaniach U-statystyk nastąpił na początku lat dziewięćdziesiątych, wraz z wprowadzeniem pojęcia *uniezależniania* (ang. *decoupling*). W szczególności wyniki de la Peñi i Montgomery'ego-Smitha pozwoliły na sprowadzenie badania klasycznych U-statystyk do tzw. U-statystyk uniezależnionych, które warunkowo można traktować jako sumy niezależnych zmiennych losowych. Pozwoliło to na wykazanie, że warunki dostateczne z pracy Rubina i Vitalego są warunkami koniecznymi w centralnym twierdzeniu granicznym dla U-statystyk (Giné, Zinn [23]). Okazało się także, że jest to jedyny przypadek w którym naturalne odpowiedniki warunków dla sum niezależnych zmiennych losowych są warunkami koniecznymi dla odpowiednich twierdzeń dla U-statystyk. W przypadku prawa wielkich liczb, odpowiednie warunki zostały znalezione przez Latałę i Zinna [36] i okazują się być dużo bardziej skomplikowane niż w przypadku klasycznym. Jeśli chodzi o prawo iterowanego logarytmu, warunki konieczne i dostateczne do niedawna znane były jedynie dla U-statystyk rzędu 2 (nie licząc oczywiście klasycznego przypadku U-statystyk rzędu 1). Zostały one znalezione przez Giné, Kwapienia, Latałę i Zinna [24] i podobnie jak w przypadku prawa wielkich liczb, okazują się być bardziej złożone niż ich odpowiedniki dla sum zmiennych i.i.d. Nadal, poza pewnymi szczególnymi przypadkami, nie jest znana postać granicy w prawie iterowanego logarytmu.

Niniejsza rozprawa poświęcona jest rozszerzeniu wyników znanych dla sum niezależnych zmiennych losowych i U-statystyk rzędu 2 na U-statystyki dowolnego rzędu.

W szczególności w obszarze zainteresowań pozostaje badanie wzrostu momentów oraz zachowania odchyleń U-statystyk rzeczywistych i o wartościach w przestrzeniach Banacha. Problemom tym poświęcona jest pierwsza część pracy, w której uogólniamy klasyczne wyniki dotyczące U-statystyk szczególnego rodzaju, jakimi są gaussowskie chaosy jednorodne, na chaosy generowane przez ogólniejszą klasę zmiennych, znajdujemy także ich wersje dla ogólnych U-statystyk o wartościach w przestrzeniach Banacha. Otrzymane oszacowania wyrażone sa poprzez wartości oczekiwane supremów procesów empirycznych, analogicznie jak w klasycznych nierównościach dla chaosów gaussowskich (Borell [7], Arcones-Giné [6]). Następnie przedstawiamy dokładne oszacowania momentów i odchyleń U-statystyk rzeczywistych i o wartościach w przestrzeni Hilberta, uogólniając wyniki Giné, Latały i Zinna, dotyczące rzeczywistych U-statystyk rzędu 2 [18].

Uzyskane nierówności eksponencjalne pozwalają w drugiej części pracy rozszerzyć podaną przez Giné, Kwapienia, Latałę i Zinna charakteryzację prawa iterowanego logarytmu na U-statystyki dowolnego rzędu. Wyniki prezentowane w tej części rozprawy pochodzą ze wspólnego artykułu z Rafałem Latałą. W niniejszej pracy udało się jednak uprościć część dowodów.

Podziękowania Pragnę wyrazić swoją wdzięczność Promotorowi, prof. Rafałowi Latale, który wprowadził mnie nie tylko w problematykę U-statystyk, ale także w teorię prawdopodobieństwa. Praca ta nigdy by nie powstała, gdyby nie jego wskazówki i opieka. Chciałbym także podziękować prof. Stanisławowi Kwapieniowi za liczne uwagi, związane z artykułami, w których przedstawione zostały główne wyniki pracy. Dużo zawdzięczam także rozmowom z innymi Doktorantami, zajmującymi się probabilistyką na Wydziale Matematyki, Informatyki i Mechaniki Uniwersytetu Warszawskiego oraz w Instytucie Matematycznym PAN.

Osobne podziękowania kieruję do Rodziców, którzy wspierali mnie przez cały okres pracy nad niniejsza rozprawą.

Ogólne założenia i notacja

Ze względu na zależność składników sum definiujących U-statystyki od wielu zmiennych, istotne jest wprowadzenie przejrzystej notacji, pozwalającej na częściowe uproszczenie dość skomplikowanych wzorów, pojawiających się przy badaniu tej klasy zmiennych losowych. Notacja, którą proponujemy jest oparta na książce [12], najpełniejszej jak dotąd pozycji na temat U-statystyk.

W pracy rozważamy U-statystyki stałego rzędu $d \in \mathbb{N}\backslash\{0\} =: \mathbb{N}^*$. Mogą być one oparte na jednej próbce, czyli ciągu $(X_i)_{i\in N^*}$ niezależnych zmiennych losowych lub na d próbkach, czyli macierzy $(X_i^{(k)})_{i\in\mathbb{N}^*,k=1,\ldots,d}$ niezależnych zmiennych losowych (niekoniecznie o tym samym rozkładzie). W obu przypadkach zakładamy, że zmienne przyjmują wartości w przestrzeni $(\Sigma, \mathcal{B})$, gdzie Σ jest przestrzenią polską, zaś $\mathcal{B}$ – σ-ciałem zbiorów borelowskich.

Przez $\mathrm{i}, \mathrm{j} \in (\mathbb{N}^*)^d$ będziemy oznaczać multiindeksy związane z zakresem sumowania w definicji U-statystyk. Przyjmujemy konwencję $\mathrm{i} = (i_1, \ldots, i_d), \mathrm{j} = (j_1, \ldots, j_d)$. Zdefiniujmy ponadto $|\mathrm{i}| = \max_{k\leqslant d} |i_k|$.

Wprowadźmy także następujące oznaczenia

$$I_d = \{1, 2, \ldots, d\},$$
$$I_n^d = \{\mathrm{i} \in (N^*)^d \colon |\mathrm{i}| \leqslant n, i_j \neq i_k \text{ dla } j \neq k\}.$$

W szczególności I_n^d **nie** oznacza potęgi kartezjańskiej zbioru I_n, aczkolwiek $I_n^1 = I_n$. Notacja ta jest spójna z wprowadzoną w [12].

Ponadto dla multiindeksu i zdefiniujmy

$$\mathbf{X}_\mathrm{i} = (X_{i_1}, \ldots, X_{i_d}),$$
$$\mathbf{X}_\mathrm{i}^{\mathrm{dec}} = (X_{i_1}^{(1)}, \ldots, X_{i_d}^{(d)}),$$

gdzie indeks „dec" pochodzi od słowa *decoupled*. W sytuacji gdy X_i będą miały te same rozkłady, będziemy pisać

$$X = (X_1, \ldots, X_d).$$

Będziemy odwoływać się także do podzbiorów zbioru indeksów. Dla $J \subseteq I_d$ oraz multiindeksu i oznaczmy

$$\mathbf{X}_{\mathrm{i}_J}^{\mathrm{dec}} = (X_{i_j}^{(j)})_{j\in J}.$$

Analogicznie

$$X_J = (X_j)_{j\in J}.$$

W powyższej notacji odpowiednie U-statystyki związane z ciągiem (X_i) oraz macierzą $(X_i^{(k)})$ można zdefiniować jako

$$\sum_{\mathrm{i}\in I_n^d} h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}) = \sum_{\mathrm{i}\in I_n^d} h_{\mathrm{i}}(X_{i_1},\dots,X_{i_d}) \quad \text{– } U\text{-statystyka zwyczajna,}$$

$$\sum_{|\mathrm{i}|\leqslant n} h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}) = \sum_{|\mathrm{i}|\leqslant n} h_{\mathrm{i}}(X_{i_1}^{(1)},\dots,X_{i_d}^{(d)}) \quad \text{– } U\text{-statystyka uniezależniona (decoupled),}$$

gdzie dla każdego i z zakresu sumowania $h_{\mathrm{i}}\colon \Sigma^d \to B$ jest funkcją borelowską o wartości w odpowiedniej przestrzeni Banacha.

Będziemy także rozpatrywać szczególny przypadek U-statystyk, jakim są chaosy, czyli wieloliniowe formy od niezależnych zmiennych losowych. Odpowiada to sytuacji $\Sigma = \mathbb{R}$ oraz $h_{\mathrm{i}}(x_1,\dots,x_d) = a_{\mathrm{i}}\prod_{k=1}^d x_k$, gdzie a_{i} są ustalonymi elementami przestrzeni B. W tym przypadku zmienimy notację i do oznaczenia zmiennych losowych definiujących chaos będziemy używać małych liter alfabetu greckiego, np. $(\eta_i)_{i\in\mathbb{N}^*}$, $(\eta_i^{(k)})_{i\in\mathbb{N}^*,k\in I_d}$. Ponadto będziemy oznaczać

$$\eta_{\mathrm{i}} = \prod_{k=1}^d \eta_{i_k},$$

$$\eta_{\mathrm{i}}^{\mathrm{dec}} = \prod_{k=1}^d \eta_{i_k}^{(k)}.$$

Oznaczenie to stoi w drobnym konflikcie z notacją $\mathbf{X}_{\mathrm{i}}, \mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}$, ponieważ jednak będzie się pojawiać w odpowiednim kontekście, nie powinno prowadzić do niejednoznaczności (oznaczenia te również są zgodne z notacją z [12]).

W szczególności będziemy rozpatrywać dwa rodzaje chaosów, tj.

$$\sum_{\mathrm{i}\in I_n^d} a_{\mathrm{i}}\eta_{\mathrm{i}} = \sum_{\mathrm{i}\in I_n^d} a_{\mathrm{i}} \prod_{k=1}^d \eta_{i_k} \quad \text{– chaos zwyczajny,}$$

$$\sum_{|\mathrm{i}|\leqslant n} a_{\mathrm{i}}\eta_{\mathrm{i}}^{\mathrm{dec}} = \sum_{|\mathrm{i}|\leqslant n} a_{\mathrm{i}} \prod_{k=1}^d \eta_{i_k}^{(k)} \quad \text{– chaos uniezależniony.}$$

Ponieważ w dowodach będziemy rozpatrywać także tzw. U-statystyki zrandomizowane, w rozdziale dot. praw iterowanego logarytmu użyjemy następującej notacji

$$\sum_{|\mathrm{i}|\leqslant n} \epsilon_{\mathrm{i}}^{\mathrm{dec}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}) = \sum_{|\mathrm{i}|\leqslant n} \Big[\prod_{k=1}^d \varepsilon_{i_k}^{(k)}\Big] h(X_{i_1}^{(1)},\dots,X_{i_d}^{(d)}),$$

gdzie $(\varepsilon_i^{(k)})_{i\in\mathbb{N}^*,k\in I_d}$ to macierz niezależnych zmiennych Rademachera (tzn. niezależnych, symetrycznych zmiennych losowych o wartościach ± 1), niezależnych od zmiennych $(X_i^{(k)})_{i\in\mathbb{N}^*,k\in I_d}$.

Wielokrotnie, licząc wartości oczekiwane warunkowo, będziemy całkować tylko względem pewnego zbioru zmiennych. Przyjmijmy zatem notację $\mathbb{E}_k$ na określenie całkowania jedynie względem zmiennych $(X_i^{(k)})_i$ bądź zmiennej X_k (w zależności od kontekstu, nie powinno to prowadzić do niejednoznaczności). Podobnie dla zbioru I, $\mathbb{E}_I$ będzie oznaczać całkowanie względem $(X_i^{(k)})_{k\in I,i}$ lub względem $(X_k)_{k\in I}$. Na przykład

$$\mathbb{E}_I \Big\| \sum_{|\mathrm{i}|\leqslant n} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}) \Big\|$$

oznacza całkę względem $(X_{i_k}^{(k)})_{k\in I, i\leqslant n}$, zaś

$$\mathbb{E}_I |h(X_1, \ldots, X_d)|$$

całkę względem $(X_k)_{k\in I}$.

Będziemy także stosowali standardowe oznaczenia dla przestrzeni L^p oraz p-tych norm ($\| \cdot \|_p$) zmiennych losowych.

Podstawowe fakty wykorzystywane w rozprawie

W niniejszym rozdziale prezentujemy kilka podstawowych twierdzeń z teorii U-statystyk, w szczególności twierdzenie o uniezależnianiu, pochodzące od de la Peñi i Montgomery'ego-Smitha. Przedstawimy także kilka klasycznych faktów, dotyczących związków między oszacowaniami momentów i odchyleń zmiennych losowych.

Twierdzenie o uniezależnianiu

Jak już wspomnieliśmy we Wstępie, jednym z najważniejszych twierdzeń, odpowiedzialnych za przełom w teorii U-statystyk, jest wynik dotyczący porównywania odchyleń i momentów U-statystyk zwyczajnych i uniezależnionych. Rezultat ten ma istotne znaczenie, gdyż U-statystyka uniezależniona może być warunkowo traktowana jako suma niezależnych zmiennych losowych, co pozwala stosować do niej cały klasyczny aparat matematyczny związany z takimi sumami. Umożliwia to w szczególności przeprowadzanie dowodów indukcyjnych, gdzie indukcja przebiega względem rzędu U-statystyki. Poniżej przedstawiamy wersje Twierdzenia o uniezależnianiu, pochodzącą z monografii de la Peñi i Giné [12].

Twierdzenie 1 (de la Peña). *Rozważmy liczby naturalne $n \geqslant d$ oraz $(X_i)_{i=1}^n$ – ciąg niezależnych zmiennych losowych o wartości w przestrzeni $(\Sigma, \mathcal{B})$. Rozważmy także $(X_i^{(k)})_{i \leqslant n}$, $k = 1, \ldots, d$ – niezależne kopie ciągu $(X_i)_{i=1}^n$. Niech ponadto $(B, \|\cdot\|)$ będzie ośrodkową przestrzenią Banacha oraz, dla każdego $\mathrm{i} = (i_1, \ldots, i_d) \in I_n^d$, niech $h_\mathrm{i}\colon \Sigma^d \to B$ będzie funkcją mierzalną, taką, że $\mathbb{E}\|h_\mathrm{i}(X_{i_1}, \ldots, X_{i_d})\| < \infty$. Niech $\Phi\colon [0,\infty) \to [0,\infty)$ będzie funkcją wypukłą, niemalejącą, spełniającą warunek $\mathbb{E}\Phi(\|h_\mathrm{i}(X_{i_1}, \ldots, X_{i_d})\|) < \infty$ dla wszystkich $\mathrm{i} \in I_n^d$. Wówczas,*

$$\mathbb{E}\Phi\Big(\Big\|\sum_{\mathrm{i}\in I_n^d} h_\mathrm{i}(X_{i_1}, \ldots, X_{i_d})\Big\|\Big) \leqslant \mathbb{E}\Phi\Big(C_d\Big\|\sum_{\mathrm{i}\in I_n^d} h_\mathrm{i}(X_{i_1}^{(1)}, \ldots, X_{i_d}^{(d)})\Big\|\Big),$$

gdzie $C_d = 2^d(d^d-1)((d-1)^{(d-1)}-1) \times \ldots \times 3$. Jeśli ponadto funkcje h_i są symetryczne, tzn. dla dowolnych $x_1, \ldots, x_d \in \Sigma$ oraz wszystkich permutacji σ zbioru

$I_d = \{1, \ldots, d\}$,

$$h_{i_1,\ldots,i_d}(x_1, \ldots, x_d) = h_{i_{\sigma_1},\ldots,i_{\sigma_d}}(x_{\sigma_1}, \ldots, x_{\sigma_d}),$$

to zachodzi nierówność odwrotna, tzn.

$$\mathbb{E}\Phi\Big(\frac{1}{2^{(2d-2)}(d-1)!}\Big\| \sum_{\mathrm{i} \in I_n^d} h_{\mathrm{i}}(X_{i_1}^{(1)}, \ldots, X_{i_d}^{(d)})\Big\|\Big), \leqslant \mathbb{E}\Phi\Big(\Big\| \sum_{\mathrm{i} \in I_n^d} h_{\mathrm{i}}(X_{i_1}, \ldots, X_{i_d})\Big\|\Big).$$

Prawdziwe jest także twierdzenie, mówiące o porównywaniu odchyleń U-statystyk zwyczajnych i uniezależnionych. Przedstawiamy je poniżej, również w wersji z [12]. Wynik ten jest formalnie silniejszy od poprzedniego twierdzenia, które wynika z niego poprzez całkowanie przez części, niemniej porównywanie momentów używane jest w jego dowodzie. W dalszej części pracy, do obu twierdzeń będziemy odwoływać się jako do *twierdzenia o uniezależnianiu.*

Twierdzenie 2 (de la Peña, Montgomery-Smith). *Rozważmy liczby naturalne $n \geqslant d$ oraz $(X_i)_{i=1}^n$ – ciąg niezależnych zmiennych losowych o wartości w przestrzeni $(\Sigma, \mathcal{B})$. Rozważmy także $(X_i^{(k)})_{i \leqslant n}$, $k = 1, \ldots, d$ – niezależne kopie ciągu $(X_i)_{i=1}^n$. Niech ponadto $(B, \|\cdot\|)$ będzie ośrodkową przestrzenią Banacha oraz, dla każdego $\mathrm{i} = (i_1, \ldots, i_d) \in I_n^d$, niech $h_{\mathrm{i}} \colon \Sigma^d \to B$ będzie funkcją mierzalną. Wówczas, istnieją stałe $C_d \in (0, \infty)$, zależne tylko od d, takie, że dla wszystkich $t > 0$,*

$$\mathbb{P}\Big(\Big\| \sum_{\mathrm{i} \in I_n^d} h_{\mathrm{i}}(X_{i_1}, \ldots, X_{i_d})\Big\| > t\Big) \leqslant C_d \mathbb{P}\Big(\Big\| \sum_{\mathrm{i} \in I_n^d} h_{\mathrm{i}}(X_{i_1}^{(1)}, \ldots, X_{i_d}^{(d)})\Big\| > t/C_d\Big).$$

Jeśli ponadto funkcje h_{i} są symetryczne, tzn. dla dowolnych $x_1, \ldots, x_d \in \Sigma$ oraz wszystkich permutacji σ zbioru $I_d = \{1, \ldots, d\}$,

$$h_{i_1,\ldots,i_d}(x_1, \ldots, x_d) = h_{i_{\sigma_1},\ldots,i_{\sigma_d}}(x_{\sigma_1}, \ldots, x_{\sigma_d}),$$

to zachodzi nierówność odwrotna, tzn. dla dowolnego $t > 0$,

$$\mathbb{P}\Big(\Big\| \sum_{\mathrm{i} \in I_n^d} h_{\mathrm{i}}(X_{i_1}^{(1)}, \ldots, X_{i_d}^{(d)})\Big\| > t\Big) \leqslant D_d \mathbb{P}\Big(\Big\| \sum_{\mathrm{i} \in I_n^d} h_{\mathrm{i}}(X_{i_1}, \ldots, X_{i_d})\Big\| > t/D_d\Big),$$

gdzie $D_d \in (0, \infty)$ zależą tylko od d.

Podstawowe związki między oszacowaniami momentów i odchyleń zmiennych losowych

Przedstawimy teraz klasyczne fakty dotyczące oszacowań odchyleń zmiennych losowych, z których będziemy intensywnie korzystać w dalszej części pracy. Zaczniemy od następującego prostego twierdzenia.

Twierdzenie 3. *Rozważmy nieujemną zmienną losową X. Dla dowolnego $p > 0$ zachodzi*

$$\mathbb{P}(X \geqslant e\|X\|_p) \leqslant e^{-p}.$$

W szczególności, jeżeli dla pewnego $p_0 \geqslant 1$ oraz liczb $\alpha_1, \ldots, \alpha_m, a_0, a_1, \ldots, a_m \geqslant 0$ zachodzi

$$\|X\|_p \leqslant a_0 + K \sum_{i=1}^{n} p^{\alpha_i} a_i$$

dla dowolnego $p \geqslant p_0$, to dla dowolnych $t \geqslant 0$,

$$\mathbb{P}(X \geqslant ea_0 + t) \leqslant L \exp\Big[-\frac{1}{L} \min_{i=1,\ldots,m} \Big(\frac{t}{a_i}\Big)^{1/\alpha_i}\Big],$$

gdzie L zależy tylko od K, p_0, m i $\min_{i \leqslant m} \alpha_i$.

Dowód. Pierwsza nierówność wynika wprost z nierówności Czebyszewa. Aby wykazać drugą nierówność, rozważmy

$$p = \frac{1}{L_1} \min_{i=1,\ldots,m} \Big(\frac{t}{a_i}\Big)^{1/\alpha_i}.$$

Zauważmy, że jeśli dla każdego $i = 1, \ldots, m$, $L_1^{\alpha_i} \geqslant mKe$, to

$$eK \sum_{i=1}^{m} p^{\alpha_i} a_i \leqslant eK \sum_{i=1}^{m} \frac{t}{L_1^{\alpha_i}} \leqslant t.$$

Zatem, jeśli $p \geqslant p_0$, to

$$\mathbb{P}(X \geqslant ea_0 + t) \leqslant \mathbb{P}(X \geqslant e\|X\|_p) \leqslant e^{-p}.$$

Jeśli z kolei $p < p_0$, to dla $L_2 = e^{p_0}$, mamy trywialną nierówność

$$\mathbb{P}(X \geqslant ea_0 + t) \leqslant 1 \leqslant L_2 e^{-p}.$$

Łącząc powyższe oszacowania, otrzymujemy tezę. □

Uwaga Prawdziwe jest także twierdzenie odwrotne, oszacowanie odchyleń postaci

$$\mathbb{P}(X \geqslant a_0 + t) \leqslant L \exp\Big[-\frac{1}{L} \min_{i=1,\ldots,m} \Big(\frac{t}{a_i}\Big)^{1/\alpha_i}\Big]$$

dla $t \geqslant 0$, implikuje nierówność $\|X\|_p \leqslant a_0 + K \sum_{i=1}^{m} p^{\alpha_i} a_i$, $p \geqslant 1$, gdzie K zależy tylko od L i $(\alpha_i)_i$. Dowód wynika w prosty sposób z wzoru na całkowanie przez części.

Będziemy także korzystać z następującej wersji nierówności Paleya-Zygmunda [42].

Twierdzenie 4 (Paley,Zygmund, [12], Corollary 3.3.2). *Dla dowolnej nieujemnej zmiennej losowej Y, całkowalnej z kwadratem oraz dowolnej liczby $\lambda \in (0,1)$,*

$$\mathbb{P}(Y \geqslant \lambda \mathbb{E} Y) \geqslant (1-\lambda)^2 \frac{(\mathbb{E} Y)^2}{\mathbb{E} Y^2}.$$

Część I

Oszacowania momentów i odchyleń U-statystyk

Rozdział 1

Nierówności dla chaosów wielomianowych

W niniejszym rozdziale prezentujemy nierówności dla szczególnego typu U-statystyk, jakim są jednorodne chaosy wielomianowe, czyli formy wieloliniowe od wektorów losowych. Otrzymane wyniki uogólniają nierówności wyprowadzone prze Borella [7] i Arconesa-Giné [6] dla chaosów gaussowskich oraz uzupełniają nierówności Talagranda [47] oraz Bousqueta, Boucherona, Lugosi i Massarta [8] dla chaosów rademacherowych. Istotnym elementem jest prosta metoda dowodu, pozwalająca na jednolite potraktowanie przypadku gaussowskiego i rademacherowego, dotychczas rozpatrywanych przy pomocy różnych środków. Metoda ta, oparta na pewnym typie „tensoryzacji" pozwala także na odtworzenie nierówności Łochowskiego [40] dla chaosów generowanych przez zmienne o logarytmicznie wklęsłych „ogonach" oraz innych klas zmiennych losowych, posiadających odpowiednie własności koncentracyjne.

1.1 Koncentracja miary dla iloczynów tensorowych

Główne twierdzenie tej sekcji sformułujemy i udowodnimy używając iloczynów tensorowych. W natychmiastowy sposób będzie można je wyrazić w języku chaosów o wartościach w przestrzeniach Banacha. Zanim sformułujemy twierdzenie, przypomnijmy, że półnormą nazywamy jednorodną funkcję nieujemną na przestrzeni liniowo-topologicznej, spełniającą nierówność trójkąta (w przeciwieństwie do normy, półnorma może się zatem zerować na niezerowych wektorach).

Twierdzenie 5. *Niech $X = (X_1, \ldots, X_n)$ będzie wektorem losowym w $\mathbb{R}^n$ (niekoniecznie o współrzędnych niezależnych), dla którego istnieje stała K taka, że dla dowolnej półnormy $\varphi\colon \mathbb{R}^n \to \mathbb{R}_+$, zachodzi $\mathbb{E}\varphi(X) < \infty$ oraz*

$$\|\varphi(X) - \mathbb{E}\varphi(X)\|_p \leqslant K\sqrt{p} \sup_{|\alpha|=1} \varphi(\alpha)$$

dla wszystkich $p \geqslant 1$, gdzie $|\alpha|$ oznacza normę euklidesową wektora $\alpha \in \mathbb{R}^n$.

Wówczas, jeśli $X_1, \ldots, X_d$ są niezależnymi kopiami X oraz $\psi\colon \bigotimes_{i=1}^d \mathbb{R}^n \to \mathbb{R}_+$ jest półnormą, zachodzi nierówność

$$\|\psi(\bigotimes_{i=1}^d X_i) - \mathbb{E}\psi(\bigotimes_{i=1}^d X_i)\|_p \leqslant K_d \sum_{I \subseteq \{1,\ldots,d\}, I \neq \emptyset} p^{\#I/2}\mathbb{E} \sup_{|\alpha_k| \leqslant 1\colon k \in I} \psi\left(\bigotimes_{i=1}^d X_{i,I,(\alpha_k)_{k\in I}}\right),$$

gdzie

$$X_{i,I,(\alpha_k)_{k\in I}} = \begin{cases} X_i & \text{jeśli} \quad i \notin I, \\ \alpha_i & \text{jeśli} \quad i \in I, \end{cases}$$

zaś K_d jest stałą zależną tylko od K oraz d.

Dowód. Indukcja ze względu na d. Dla $d = 1$ teza twierdzenia pokrywa się z założeniem. Załóżmy zatem, że teza zachodzi dla dowolnej mniejszej niż d ($d \geqslant 2$) liczby zmiennych X_i. Używając warunkowo założenia indukcyjnego dla $d_1 = 1$ oraz półnormy

$$\varphi_1(x) = \psi(X_1 \otimes \ldots \otimes X_{d-1} \otimes x),$$

otrzymujemy

$$\mathbb{E}_{X_d}\left|\psi(\bigotimes_{i=1}^d X_i) - \mathbb{E}_{X_d}\psi(\bigotimes_{i=1}^d X_i)\right|^p \leqslant K^p p^{p/2} \sup_{|\alpha| \leqslant 1} \psi(\bigotimes_{i=1}^{d-1} X_i \otimes \alpha)^p.$$

Zauważmy, że funkcja $\varphi_2(x) = \sup_{|\alpha|\leqslant 1} \psi(x \otimes \alpha)$ jest półnormą na $\bigotimes_{i=1}^{d-1} \mathbb{R}^d$. Zatem, z założenia indukcyjnego, nierówności trójkąta w L^p oraz twierdzenia Fubiniego, dostajemy

$$\begin{aligned} &\mathbb{E}\left|\psi(\bigotimes_{i=1}^d X_i) - \mathbb{E}_{X_d}\psi(\bigotimes_{i=1}^d X_i)\right|^p \\ &\qquad \leqslant \tilde{K}_{d-1}^p \sum_{I \subseteq \{1,\ldots,d\}, d \in I} p^{p\#I/2}\Big(\mathbb{E} \sup_{|\alpha_k| \leqslant 1\colon k \in I} \psi(\bigotimes_{i=1}^d X_{i,I,(\alpha_k)_{k\in I}})\Big)^p, \end{aligned} \tag{1.1}$$

dla pewnego $\tilde{K}_{d-1}$ zależącego tylko od K_{d-1} oraz d.

Pozostaje oszacować $\mathbb{E}|\mathbb{E}_{X_d}\psi(\bigotimes_{i=1}^d X_i) - \mathbb{E}\psi(\bigotimes_{i=1}^d X_i)|^p$. Rozważmy w tym celu funkcję $\varphi_3\colon \bigotimes_{i=1}^{d-1} \mathbb{R}^n \to \mathbb{R}_+$, zdefiniowaną wzorem $\varphi_3(x) = \mathbb{E}\psi(x \otimes X_d)$. Łatwo sprawdzić, że φ_3 jest półnormą, a zatem z założenia indukcyjnego,

$$\mathbb{E}|\varphi_3(\bigotimes_{i=1}^{d-1} X_i) - \mathbb{E}\varphi_3(\bigotimes_{i=1}^{d-1} X_i)|^p \tag{1.2}$$
$$\leqslant \tilde{K}_{d-1}^p \sum_{I\subseteq\{1,\ldots,d-1\}, I\neq\emptyset} p^{p\#I/2}\Big(\mathbb{E}\sup_{|\alpha_k|\leqslant 1\colon k\in I} \varphi_3(\bigotimes_{i=1}^{d-1} X_{i,I,(\alpha_k)_{k\in I}})\Big)^p.$$

Wystarczy teraz zauważyć, że dla każdego $I \subseteq \{1,\ldots,d-1\}$,

$$\mathbb{E}\sup_{|\alpha_k|\leqslant 1\colon k\in I} \varphi_3\Big(\bigotimes_{i=1}^{d-1} X_{i,I,(\alpha_k)_{k\in I}}\Big) \leqslant \mathbb{E}\sup_{|\alpha_k|\leqslant 1\colon k\in I} \psi\Big(\bigotimes_{i=1}^{d} X_{i,I,(\alpha_k)_{k\in I}}\Big),$$

co wraz z (1.1) i (1.2) kończy dowód. $\square$

Natychmiastowym wnioskiem z powyższego twierdzenia i nierówności Czebyszewa są oszacowania odchyleń zmiennej losowej $\psi(\bigotimes_{i=1}^d X_i)$ od średniej.

Wniosek 1. *Przy założeniach Twierdzenia 5, istnieją stałe K_d, zależne tylko od K oraz d, takie, że dla $p \geqslant 1$,*

$$\mathbb{P}\Big(|\psi(\bigotimes_{i=1}^{d} X_i) - \mathbb{E}\psi(\bigotimes_{i=1}^{d} X_i)| \geqslant K_d \sum_{I\subseteq\{1,\ldots,d\}, I\neq\emptyset} p^{\#I/2}\mathbb{E}\sup_{|\alpha_k|\leqslant 1\colon k\in I} \psi(\bigotimes_{i=1}^{d} X_{i,I,(\alpha_k)_{k\in I}})\Big) \leqslant e^{-p}.$$

Uwaga Założenia Twierdzenia 5 są spełnione dla szerokiej klasy wektorów losowych, w szczególności dla wektorów posiadających własność koncentracji subgaussowskiej dla funkcji lipschitzowskich, wypukłych. Własność tę (ze stałą niezależną od wymiaru n) posiadają m.in. wszystkie wektory $X = (X_1,\ldots,X_n)$, gdzie zmienne X_i są niezależne i spełniają logarytmiczną nierówność Sobolewa ze wspólną stałą, np. standardowy wektor gaussowski w $\mathbb{R}^n$, $X \sim \mathcal{N}(0,\mathrm{Id})$ (zob. np. [38]). Własność koncentracji dla funkcji lipschitzowskich wypukłych posiadają także wektory o współrzędnych niezależnych i ograniczonych (ze stałą zależną tylko od normy L^∞ składników), a nawet dowolne wektory ograniczone. W tym ostatnim przypadku stała w nierównościach koncentracyjnych, a w konsekwencji stała K w założeniach Twierdzenia 5 zależy od macierzy współczynników Φ-mieszania między składowymi wektora X [45].

Warto także nadmienić, że Twierdzenie 5 można nieco zmodyfikować, mianowicie jeśli zastąpimy w założeniach $(\varphi(X) - \mathbb{E}\varphi(X))$ przez $(\varphi(X) - \mathbb{E}\varphi(X))_+$ oraz w tezie $\psi(\bigotimes_{i=1}^d X_i) - \mathbb{E}\psi(\bigotimes_{i=1}^d X_i)$ przez $(\psi(\bigotimes_{i=1}^d X_i) - \mathbb{E}\psi(\bigotimes_{i=1}^d X_i))_+$, twierdzenie pozostaje prawdziwe (dowód różni się tylko formalnymi zmianami). Szeroka klasa wektorów losowych o współrzędnych niezależnych, spełniających to zmodyfikowane założenie, została wprowadzona w pracy [1]. Poniżej przedstawiamy odpowiednie twierdzenie. Ponieważ zarówno metoda dowodu, jak i tematyka nie dotyczą bezpośrednio tematu niniejszej rozprawy, dowód twierdzenia zostanie pominięty.

TWIERDZENIE 6. *Niech $X_1, \ldots, X_n$ będą niezależnymi zmiennymi losowymi, dla których istnieją stałe $C, T \in \mathbb{R}_+$, $\alpha \in [0,1)$ takie, że dla $t > T$ zachodzą nierówności*

$$\mathbb{P}(X_i \geqslant t + C/t) \leqslant \alpha \mathbb{P}(X \geqslant t),$$
$$\mathbb{P}(X_i \leqslant -t - C/t) \leqslant \alpha \mathbb{P}(X \leqslant -t).$$

Wówczas istnieje stała K, zależna tylko od C, T, α taka, że dla dowolnej funkcji wypukłej 1-lipschitzowskiej $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ zachodzi $\mathbb{E}|\varphi(X_1, \ldots, X_n)| < \infty$ oraz

$$\mathbb{P}(\varphi(X_1, \ldots, X_n) \geqslant \mathbb{E}\varphi(X_1, \ldots, X_n) + t) \leqslant e^{-t^2/K}$$

dla wszystkich $t \geqslant 0$. W szczególności

$$\|(\varphi(X_1, \ldots, X_n) - \mathbb{E}\varphi(X_1, \ldots, X_n))_+\|_p \leqslant L\sqrt{p}$$

dla stałej L zależnej tylko od C, T, α.

1.2 Oszacowania momentów dla chaosów wielomianowych

Twierdzenie 5 ma interpretację w języku chaosów uniezależnionych. Zanim ją przedstawimy, wprowadźmy niezbędne definicje.

Rozważmy wektor losowy $\eta = (\eta_1, \ldots, \eta_n)$ oraz jego niezależne kopie $\eta^{(1)}, \ldots, \eta^{(d)}$. Niech $\mathcal{T}$ będzie przeliczalnym zbiorem d-wskaźnikowych macierzy $t\colon \{1, \ldots, n\}^d \to \mathbb{R}$, ograniczonym w metryce euklidesowej. Rozważmy zmienną losową

$$Z := \sup_{t \in \mathcal{T}} |\sum_{|\mathrm{i}| \leqslant n} t_{\mathrm{i}} \eta_{\mathrm{i}}^{\mathrm{dec}}| = \sup_{t \in \mathcal{T}} |\sum_{|\mathrm{i}| \leqslant n} t_{\mathrm{i}} \prod_{k=1}^{d} \eta_{i_k}^{(k)}|. \tag{1.3}$$

W tym miejscu po raz pierwszy stosujemy uproszczony zapis, wprowadzony w rozdziale OGÓLNE ZAŁOŻENIA I NOTACJA. Zapis ten będziemy konsekwentnie stosować w dalszej części rozprawy.

W następnej kolejności zdefiniujemy wielkości charakteryzujące „rozmiar” zbioru $\mathcal{T}$.

Definicja 1. *Dla $I \subseteq I_d$ niech*

$$\|\mathcal{T}\|_I = \mathbb{E}\sup_{t\in\mathcal{T}} \sup_{|\alpha^{(k)}|\leqslant 1, k\in I} \Big| \sum_{|\mathrm{i}|\leqslant n} t_{\mathrm{i}} \prod_{k\notin I} \eta_{i_k}^{(k)} \prod_{k\in I} \alpha_{i_k}^{(k)} \Big|,$$

gdzie wewnętrzne supremum dotyczy ciągów wektorów n-wymiarowych $(\alpha^{(k)})_{k\in I} \in (\mathbb{R}^n)^I$, indeksowanych przez zbiór I, zaś $|\alpha^{(k)}|$ oznacza normę euklidesową wektora $\alpha^{(k)} \in \mathbb{R}^n$.

Następujące twierdzenie jest równoważną wersją Twierdzenia 5, wyrażoną przy pomocy chaosów (równoważność jest natychmiastowym wnioskiem z definicji iloczynu tensorowego i twierdzenia Hahna-Banacha).

Twierdzenie 7. *Niech $\eta = (\eta_1, \ldots, \eta_n)$ będzie wektorem losowym takim, że dla dowolnej półnormy φ na $\mathbb{R}^n$,*

$$\|\varphi(\eta) - \mathbb{E}\varphi(\eta)\|_p \leqslant K\sqrt{p} \sup_{|\alpha|=1} \varphi(\alpha)$$

dla wszystkich $p \geqslant 1$. Wówczas, dla dowolnego ograniczonego, przeliczalnego zbioru d–wskaźnikowych macierzy, zmienna losowa Z, zdefiniowana wzorem (1.3), spełnia

$$\|Z - \mathbb{E}Z\|_p \leqslant K_d \sum_{I\subseteq\{1,\ldots,d\}, I\neq\emptyset} p^{\#I/2} \|\mathcal{T}\|_I, \tag{1.4}$$

gdzie K_d zależy tylko od K oraz d. W szczególności dla $p \geqslant 1$ zachodzi nierówność

$$\mathbb{P}\Big(|Z - \mathbb{E}Z| \geqslant K_d e \sum_{I\subseteq\{1,\ldots,d\}, I\neq\emptyset} p^{\#I/2} \|\mathcal{T}\|_I\Big) \leqslant e^{-p}. \tag{1.5}$$

Uwaga Z reguły rozpatruje się *chaosy zwyczajne*, czyli zmienne losowe postaci

$$Z = \sup_{t\in\mathcal{T}} \Big|\sum_{\mathrm{i}\in I_n^d} t_{\mathrm{i}} \eta_{\mathrm{i}}\Big| = \sup_{t\in\mathcal{T}} \Big|\sum_{\mathrm{i}\in I_n^d} t_{\mathrm{i}} \prod_{k=1}^{d} \eta_{i_k}\Big|.$$

Dość łatwo można sprawdzić, że ta klasa chaosów jest ogólniejsza od chaosów uniezależnionych, tzn. dla każdego chaosu uniezależnionego można, odpowiednio zmieniając macierze (t_{i}) znaleźć reprezentację w postaci chaosu zwyczajnego.

Dla chaosów zwyczajnych oraz η będącego wektorem niezależnych zmiennych Rademachera, nierówność (1.5) została udowodniona w przypadku $d = 2$ przez Talagranda [47]. Została ona potem rozszerzona na dowolne d przez Boucherona, Bousqueta, Lugosi i Massarta [8], z tą różnicą, że rozpatrują oni tylko górny „ogon", tj. $\mathbb{P}(Z \geqslant \mathbb{E}Z + K_d \sum_{I\subseteq\{1,\ldots,d\}, I\neq\emptyset} p^{\#I/2}\|\mathcal{T}\|_I)$. W przypadku chaosów gaussowskich, Borell [7] oraz Arcones i Giné [6] udowodnili podobne oszacowania dla

$\mathbb{P}(Z \geqslant K_d[\mathbb{E}Z + \sum_{I\subseteq\{1,\ldots,d\}, I\neq\emptyset} p^{\#I/2}\|\mathcal{T}\|_I])$, ich rezultaty są szczególnym przypadkiem oszacowań Latały i Łochowskiego (zob. Twierdzenie 8 poniżej). Nierówności, jak te rozpatrywane przez Borella i Arconesa, Giné mogą być natychmiast uzyskane z Twierdzenia 7 przy pomocy twierdzenia o uniezależnianiu (ponieważ zgodnie z tym twierdzeniem, zarówno „ogony", jak i średnie chaosu zwyczajnego i uniezależnionego są porównywalne). Nie można jednak bezpośrednio uzyskać w ten sposób dla chaosów zwyczajnych koncentracji wokół średniej (używając twierdzenia o uniezależnianiu tracimy stałą 1 przy $\mathbb{E}Z$). Z drugiej strony w Twierdzeniu 7 nie zakładamy, że wektor η ma współrzędne niezależne. Ponadto dowód przebiega analogicznie dla zmiennych Rademachera i gaussowskich, zaś metody dowodu wspomnianych nierówności znacznie różnią się od siebie.

Dowód Twierdzeń 5 i 7 może być przeniesiony na chaosy generowane przez zmienne losowe o innych własnościach koncentracji. Zilustrujemy to podając alternatywny dowód wspomnianej już nierówności Łochowskiego [40], dotyczącej chaosów generowanych przez niezależne zmienne o logarytmicznie wklęsłych „ogonach".

Definicja 2. *Niech $\mathcal{N} = (\eta_i^{(k)})_{k\leqslant d, i\leqslant n}$ będzie macierzą niezależnych, symetrycznych zmiennych losowych o logarytmicznie wklęsłych „ogonach", tj. takich, że funkcje*

$$\mathcal{N}_i^{(k)}(t) = -\log\mathbb{P}(|X_i^{(k)}| \geqslant t),\ t\geqslant 0,$$

są wypukłe. Załóżmy dodatkowo (dla znormalizowania), że

$$\inf\{t\colon \mathcal{N}_i^{(k)}(t)\geqslant 1\} = 1,$$

i zdefiniujmy funkcje $\tilde{\mathcal{N}}_i^{(k)}$, wzorem

$$\tilde{\mathcal{N}}_i^{(k)}(t) = \begin{cases} t^2 & \text{dla} \quad |t|\leqslant 1 \\ \mathcal{N}_i^{(k)}(|t|) & \text{dla} \quad |t| > 1. \end{cases}$$

Rozpatrzmy znów zmienną Z, zdefiniowaną wzorem (1.3) i wprowadźmy odpowiedniki wielkości $\|\mathcal{T}\|_I$ (dokładniej, będą to odpowiedniki wielkości $p^{\#I/2}\|\mathcal{T}\|_I$).

Definicja 3. *Dla $I\subseteq\{1,\ldots,d\}$ oraz $p\geqslant 1$ niech*

$$\|\mathcal{T}\|_{\mathcal{N},I,p} = \mathbb{E}\sup_{t\in\mathcal{T}} \sup_{\alpha^{(k)}\in\mathcal{A}_{k,p}\colon k\in I} \Big|\sum_{|\mathrm{i}|\leqslant n} t_{\mathrm{i}} \prod_{k\in I}\alpha_{i_k}^{(k)} \prod_{k\notin I}\eta_{i_k}^{(k)}\Big|,$$

gdzie

$$\mathcal{A}_{k,p} = \{\alpha\in\mathbb{R}^n\colon \sum_{i=1}^n \tilde{\mathcal{N}}_i^{(k)}(\alpha_i)\leqslant p\}.$$

Następujące twierdzenie zostało udowodnione dla $d = 1$ (przypadek kombinacji liniowych o współczynnikach wektorowych) przez Latałę [34] oraz dla dowolnego d przez Łochowskiego [40].

Twierdzenie 8 (Latała, Łochowski). *Istnieją stałe K_d (zależne tylko od d) takie, że dla wszystkich $p \geqslant 1$,*

$$\frac{1}{K_d} \sum_{I \subseteq \{1,\ldots,d\}} \|\mathcal{T}\|_{\mathcal{N},I,p} \leqslant \|Z\|_p \leqslant K_d \sum_{I \subseteq \{1,\ldots,d\}} \|\mathcal{T}\|_{\mathcal{N},I,p}.$$

Po sformułowaniu powyższego wyniku w języku półnorm, natychmiast widać, że dowód Twierdzenia 5 można bardzo łatwo zaadaptować, by uzyskać następujący

Fakt 1. *Teza Twierdzenia 8 dla $d = 1$ formalnie implikuje tezę tego twierdzenia dla dowolnego d.*

Dowód. Oszacowania z góry dowodzi się analogicznie jak w Twierdzeniu 5. Jedyną różnicę stanowi fakt, że nie pojawiają się czynniki $p^{\#Ip/2}$, gdyż zależność od p ukryta jest w zbiorach $\mathcal{A}_{k,p}$, zastępujących tu kule euklidesowe z Twierdzenia 5.

Dowód oszacowania z dołu jest jeszcze prostszy. Ponownie używamy indukcji względem d.

Oczywiście $\|Z\|_p \geqslant \|Z\|_1 = \|\mathcal{T}\|_{\mathcal{N},\emptyset,p}$. Ponadto, dla dowolnego niepustego zbioru $I \subseteq \{1,\ldots,d\}$ (niech r będzie dowolnym elementem I), z założenia indukcyjnego

$$\|Z\|_p \geqslant \frac{1}{K_1}\Big\| \sup_{t\in\mathcal{T}} \sup_{\alpha^{(r)}\in\mathcal{A}_{r,p}} |\sum_{|\mathbf{i}|\leqslant n}^{n} t_{\mathbf{i}} \prod_{k\neq r} \eta_{i_k}^{(k)} \alpha_{i_r}^{(r)}|\Big\|_p \geqslant \frac{1}{K_1 K_{d-1}} \|\mathcal{T}\|_{\mathcal{N},I,p}.$$

□

Z powyższego twierdzenia dostajemy

Wniosek 2. *Istnieją stałe K_d, c_d takie, że dla wszystkich $t \geqslant 1$,*

$$\begin{aligned}
\mathbb{P}\Big(Z \geqslant K_d \sum_{I\subseteq\{1,\ldots,d\}} \|\mathcal{T}\|_{\mathcal{N},I,t}\Big) &\leqslant e^{-t},\\
\mathbb{P}\Big(Z \geqslant \frac{1}{K_d} \sum_{I\subseteq\{1,\ldots,d\}} \|\mathcal{T}\|_{\mathcal{N},I,t}\Big) &\geqslant e^{-t} \wedge c_d.
\end{aligned}$$

Dowód. Wynikanie oszacowań odchyleń z oszacowań momentów jest wynikiem dość klasycznym i dowód podajemy jedynie dla kompletności. Oszacowanie górne wynika bezpośrednio z Twierdzenia 8 i nierówności Czebyszewa. Przytoczony poniżej dowód drugiej nierówności pochodzi od Kwapienia i Gluskina [31] (przypadek sum niezależnych zmiennych losowych). Z nierówności Paleya-Zygmunda otrzymujemy, że dla każdej liczby $\lambda \in (0,1)$ i dowolnego $p \geqslant 1$

$$\mathbb{P}(Z \geqslant \lambda \|Z\|_p) \geqslant (1-\lambda^p)^2 \frac{(\mathbb{E}Z^p)^2}{\mathbb{E}Z^{2p}}. \tag{1.6}$$

Zauważmy teraz, że z wypukłości funkcji $\mathcal{N}_i^{(k)}$ oraz warunku $\mathcal{N}_i^{(k)}(0) = 0$ wynika, że $\mathcal{N}_i^{(k)}(tx) \geqslant t\mathcal{N}_i^{(k)}(x)$ dla $x \geqslant 0$ i $t \geqslant 1$. Łatwo sprawdzić, że nierówność ta zachodzi także dla funkcji $\tilde{\mathcal{N}}_i^{(k)}$. Wynika stąd, że dla $t \geqslant 1$ i $p \geqslant 0$, $\|\mathcal{T}\|_{\mathcal{N},I,tp} \leqslant t^{\#I}\|\mathcal{T}\|_{\mathcal{N},I,p}$. W szczególności, Twierdzenie 8 daje

$$\|Z\|_{2p} \leqslant \tilde{K}_d \|Z\|_p,$$

dla wszystkich $p \geqslant 1$. Łącząc to oszacowanie z (1.6) dla $\lambda = 1/4$ i Twierdzeniem 8, otrzymujemy dla $p \geqslant 1$,

$$\mathbb{P}\left(Z \geqslant \frac{1}{4K_d} \sum_{I \subseteq \{1,\ldots,d\}} \|\mathcal{T}\|_{\mathcal{N},I,p}\right) \geqslant e^{-2p\log \tilde{K}_d}/2 \geqslant e^{-2p(1+\log \tilde{K}_d)}.$$

Podstawiając teraz $p = t/[1 \vee (2 + 2\log \tilde{K}_d)]$, oraz korzystając z faktu $\|\mathcal{T}\|_{\mathcal{N},I,p} \geqslant \bar{K}_d^{-1}\|\mathcal{T}\|_{\mathcal{N},I,t}$, dostajemy

$$\mathbb{P}\left(Z \geqslant \frac{1}{4K_d\bar{K}_d} \sum_{I \subseteq \{1,\ldots,d\}} \|\mathcal{T}\|_{\mathcal{N},I,t}\right) \geqslant e^{-t}$$

dla $t \geqslant 1 \vee (2 + 2\log \tilde{K}_d)$, skąd już łatwo wynika druga nierówność tezy. □

Posługując się twierdzeniem o uniezależnianiu, możemy wywnioskować, odpowiedniki Twierdzenia 8 i Wniosku 2 dla chaosów zwyczajnych.

WNIOSEK 3. *Niech $\mathcal{T}$ będzie zbiorem d-wskaźnikowych symetrycznych macierzy $(t_{\mathbf{i}})_{|\mathbf{i}|\leqslant n}$, takich, że $t_{\mathbf{i}} = 0$ dla $\mathbf{i} \notin I_n^d$. Rozważmy chaos*

$$Z = \sup_{t\in\mathcal{T}} |\sum_{i\in I_n^d} t_{\mathbf{i}}\eta_{\mathbf{i}}|,$$

gdzie $(\eta_i)_{i\leqslant n}$ to ciąg niezależnych, symetrycznych zmiennych losowych o logarytmicznie wklęsłych ogonach. Niech dla $1 \leqslant k \leqslant d$, $(n_i^{(k)})_{i\leqslant n}$ będą niezależnymi kopiami ciągu $(n_i)_{i\leqslant n}$,

$$\mathcal{N}_i^{(k)}(t) = \mathcal{N}_i(t) = -\log \mathbb{P}(|X_i| \geqslant t),\ t \geqslant 0,$$

zaś funkcje $\tilde{\mathcal{N}}_i^{(k)}$ oraz wielkości $\|\mathcal{T}\|_{\mathcal{N},I,p}$ będą zdefiniowane jak w Definicjach 2 i 3.

Wówczas, istnieją stałe K_d, c_d,zależne tylko od d, takie, że dla $p \geqslant 1$,

$$\frac{1}{K_d} \sum_{I \subseteq \{1,\ldots,d\}} \|\mathcal{T}\|_{\mathcal{N},I,p} \leqslant \|Z\|_p \leqslant K_d \sum_{I \subseteq \{1,\ldots,d\}} \|\mathcal{T}\|_{\mathcal{N},I,p}$$

oraz dla $t \geqslant 1$,

$$\begin{aligned} \mathbb{P}\left(Z \geqslant K_d \sum_{I \subseteq \{1,\ldots,d\}} \|\mathcal{T}\|_{\mathcal{N},I,t}\right) &\leqslant e^{-t}, \\ \mathbb{P}\left(Z \geqslant \frac{1}{K_d} \sum_{I \subseteq \{1,\ldots,d\}} \|\mathcal{T}\|_{\mathcal{N},I,t}\right) &\geqslant e^{-t} \wedge c_d. \end{aligned}$$

Uwaga Oczywiście w analogiczny sposób można uzyskiwać nierówności dla chaosów z ogólnych nierówności dla kombinacji liniowych typu

$$\Big\| \sup_{t \in \mathcal{T}} |\sum_{i=1}^n t_i \eta_i| \Big\|_p \leqslant K\Big(\Big\| \sup_{t \in \mathcal{T}} |\sum_{i=1}^n t_i \eta_i| \Big\|_1 + \sup_{t \in \mathcal{T}} \sup_{\beta \in \mathcal{B}_{p,n}} |\sum_{i=1}^n \beta_i t_i| \Big),$$

gdzie $\mathcal{B}_{p,n}$ jest pewnym podzbiorem przestrzeni $\mathbb{R}^n$. Nierówności jak powyższa, ze stałą K niezależną od n, dla niezależnych zmiennych η_i można uzyskać na przykład ze zmodyfikowanych nierówności logarytmicznych Sobolewa, które gwarantują koncentrację miary dla funkcji lipschitzowskich. Przykłady tego typu nierówności i miar je spełniających można znaleźć w pracach [9, 16, 17, 37]

Niedogodnością przedstawionych nierówności dla chaosów jest fakt, że oszacowania momentów i nierówności koncentracyjne są wyrażone poprzez wartości oczekiwane supremów procesów empirycznych, które w ogólnym przypadku mogą być trudne do wyznaczenia. W przypadku supremum chaosów rzeczywistych, bądź równoważnie chaosów o wartościach w przestrzeniach Banacha, $\|\mathcal{T}\|_{I,p}$ zależy w istotny sposób od geometrii zbioru $\mathcal{T}$ i nawet w przypadku $d = 1$, dokładne oszacowania znane są tylko w przypadku gaussowskim i wyrażone są w języku miar majoryzujących. Tym niemniej, w przypadku rzeczywistym (tj. dla $\mathcal{T}$ jednoelementowego) pożądanym wynikiem byłoby znalezienie oszacowań momentów chaosów wielomianowych przez wartości „deterministyczne". Wyniki tego typu uzyskane zostały przez Latałę w przypadku dowolnych zmiennych o logarytmicznie wklęsłych „ogonach " i $d = 2$ [34] oraz w przypadku chaosów gaussowskich dowolnego rzędu [35]. Poniżej przedstawiamy ten ostatni wynik, gdyż będzie on punktem odniesienia dla zaprezentowanych w dalszej części rozprawy nierówności dla rzeczywistych U-statystyk (podobnie jak Twierdzenie 7 dla oszacowań momentów U-statystyk o wartościach w przestrzeniach Banacha).

Poniższa definicja jest nieco szersza niż jest to wymagane dla przedstawienia oszacowań momentów chaosów gaussowskich. Będzie ona jednak wykorzystana w pełnej ogólności w dalszej części pracy.

DEFINICJA 4. *Dla dowolnego skończonego, niepustego zbioru I, niech $\mathcal{P}_I$ oznacza rodzinę złożoną ze wszystkich rozbić zbioru I na niepuste, parami rozłączne podzbiory. Niech ponadto, dla $\mathcal{J} \in \mathcal{P}_I$, $\deg \mathcal{J}$ oznacza liczbę elementów rozbicia $\mathcal{J}$. Kładziemy też $\mathcal{P}_I = \{\emptyset\}$ oraz $\deg \emptyset = 0$.*

DEFINICJA 5. *Niech $(a_{\mathrm{i}})_{|\mathrm{i}| \leqslant n}$ będzie d-wskaźnikową macierzą liczb rzeczywistych. Dla $\mathcal{J} = \{J_1, \ldots, J_k\} \in \mathcal{P}_{I_d}$ zdefiniujmy*

$$\|(a_{\mathrm{i}})_{\mathrm{i}}\|_{\mathcal{J}} = \sup\Big\{ \sum_{\mathrm{i}} a_{\mathrm{i}} x^{(1)}_{\mathrm{i}_{J_1}} \cdots x^{(k)}_{\mathrm{i}_{J_k}} \colon \sum_{\mathrm{i}_{J_1}} (x^{(1)}_{\mathrm{i}_{J_1}})^2 \leqslant 1, \ldots, \sum_{\mathrm{i}_{J_k}} (x^{(k)}_{\mathrm{i}_{J_k}})^2 \leqslant 1 \Big\}.$$

TWIERDZENIE 9 (Latała). *Istnieją stałe K_d takie, że dla dowolnej d-wskaźnikowej macierzy liczb rzeczywistych $(a_{\mathrm{i}})_{|\mathrm{i}| \leqslant n}$, oraz $g^{(k)}_i$, $i = 1, \ldots, n$, $k = 1, \ldots, d$ – niezależnych zmiennych losowych o rozkładzie $\mathcal{N}(0,1)$,*

$$\frac{1}{K_d} \sum_{\mathcal{J} \in \mathcal{P}_{I_d}} p^{\deg \mathcal{J}/2} \|(a_{\mathrm{i}})_{\mathrm{i}}\|_{\mathcal{J}} \leqslant \Big\| \sum_{|\mathrm{i}| \leqslant n} a_{\mathrm{i}} g^{\mathrm{dec}}_{\mathrm{i}} \Big\|_p \leqslant K_d \sum_{\mathcal{J} \in \mathcal{P}_{I_d}} p^{\deg \mathcal{J}/2} \|(a_{\mathrm{i}})_{\mathrm{i}}\|_{\mathcal{J}}.$$

1.3 Zastosowania. Wykładnicza całkowalność chaosu rademacherowego

Oszacowania typu Twierdzenia 8 i Wniosku 2, pomimo wspomnianej powyżej słabości, mogą okazać się przydatne. Zilustrujemy to, podając dowód wykładniczej całkowalności chaosu rademacherowego w przestrzeniach Banacha. Rezultat ten został udowodniony dla szeregów rademacherowych w przestrzeniach Banacha przez Kwapienia [30] oraz dla uogólnionych sum rademacherowych i chaosów rzędu 2 przez Ledoux i Talagranda [39]. Przedstawiony tu dowód, choć częściowo korzysta z idei opisanych w [39], jest znacznie prostszy i pozwala w przejrzysty sposób udowodnić rezultat dla chaosów dowolnego rzędu.

Następująca definicja chaosu rademacherowego w przestrzeni Banacha, pochodzi z [39] i jest nieco ogólniejsza niż standardowa.

DEFINICJA 6. *Niech B będzie przestrzenią Banacha, dla której istnieje przeliczalny zbiór $D = \{\psi_j\}$ funkcjonałów taki, że dla dowolnego $x \in B$*

$$\|x\|_B = \sup_j |\psi_j(x)|.$$

Zmienną losową X o wartościach w B nazywamy uogólnionym chaosem rademacherowym *rzędu d, jeśli istnieje d-wskaźnikowa macierz $(x_{\mathrm{i}})_{|\mathrm{i}| < \infty}$ (spełniająca $x_{\mathrm{i}} = 0$, gdy $i_k = i_l$ dla pewnych $k \neq l$) taka, że dla każdego j, ciąg zmiennych rzeczywistych $\sum_{|\mathrm{i}| \leqslant n} \psi_j(x_{\mathrm{i}}) \epsilon_{\mathrm{i}}$ jest zbieżny p.n. przy $n \to \infty$ oraz rozkłady skończenie wymiarowe procesu $(\psi_j(X))_{j \in \mathbb{N}}$ są takie same jak procesu $(\sum_{|\mathrm{i}| < \infty} \psi_j(x_{\mathrm{i}}) \epsilon_{\mathrm{i}})_{j \in \mathbb{N}}$.*

Udowodnimy następujące twierdzenie.

Twierdzenie 10 ([39], rozdz. 4 dla $d = 1, 2$). *Jeśli X jest chaosem rademacherowym rzędu d, to dla dowolnego $\lambda \in \mathbb{R}$ zachodzi*

$$\mathbb{E}\exp(\lambda \|X\|^{2/d}) < \infty.$$

Zanim przejdziemy do dowodu, zauważmy, że możemy założyć, że w Definicji 6, $B = \ell_\infty$, ψ_j jest rzutowaniem na j-tą współrzędną, zaś macierz x_{i} jest symetryczna, tzn. $x_{i_1,\ldots,i_d} = x_{i_{\pi(1)},\ldots,i_{\pi(d)}}$ dla każdej permutacji $\pi\colon I_d \to I_d$. Zachodzi następujący lemat.

Lemat 1. *Jeśli X jest chaosem rademacherowym w ℓ_∞, x_{i} odpowiadającą mu macierzą symetryczną, zaś ψ_j – rzutowaniem na j-tą współrzędną, to dla dowolnego $I \subseteq I_d$ oraz ustalonej wartości i_{I^c}, zależność*

$$\psi_j(Y) = \sum_{|\mathrm{i}_I| < \infty} \psi_j(x_{\mathrm{i}}) \prod_{k \in I} \varepsilon_{i_k}$$

definiuje chaos rademacherowy rzędu $\#I$ o wartościach w ℓ_∞.

Dowód. Z klasycznej własności hiperkontrakcji dla chaosu rademacherowego (modulo stałe wynika ona w szczególności z Twierdzenia 8) mamy

$$\|\sum_{|\mathrm{i}| \leqslant n} y_{\mathrm{i}} \epsilon_{\mathrm{i}}\|_4 \leqslant 3^{d/2} \|\sum_{|\mathrm{i}| \leqslant n} y_{\mathrm{i}} \epsilon_{\mathrm{i}}\|_2,$$

gdzie $(y_{\mathrm{i}})_{|\mathrm{i}| \leqslant n}$ to dowolna d-wskaźnikowa macierz o współczynnikach z ośrodkowej przestrzeni Banacha. Rozważmy $X_n^m = \sum_{|\mathrm{i}| < n} (\psi_1(x_{\mathrm{i}}), \ldots, \psi_m(x_{\mathrm{i}}))\epsilon_{\mathrm{i}}$, zmienną o wartościach w ℓ_∞^m. Z nierówności Paleya-Zygmunda

$$\mathbb{P}\Big(\|X_n^m\| \geqslant \frac{1}{2}\|X_n^m\|_2\Big) \geqslant \frac{9}{16} \frac{(\mathbb{E}\|X_n^m\|^2)^2}{\mathbb{E}\|X_n^m\|^4} \geqslant c_d,$$

dla pewnej stałej $c_d > 0$. Niech t będzie dowolną liczbą taką, że $\mathbb{P}(\|X_n^m\| \geqslant t) \leqslant c_d/2$. Z powyższej nierówności wynika, że

$$\|X_n^m\|_2 \leqslant 2t.$$

Zauważmy, że jeśli M spełnia $\mathbb{P}(\|X\| \geqslant M) \leqslant c_d/4$, to dla każdego m, $\mathbb{P}(\|X_\infty^m\| \geqslant M) \leqslant c_d/4$. Zatem dla dużych n (być może zależnych od m), $\mathbb{P}(\|X_n^m\| \geqslant M) \leqslant c_d/2$, czyli $\|X_n^m\|_2 \leqslant 2M$. Rozpatrzmy teraz $\tilde{X}_n^m$ – uniezależniony odpowiednik chaosu X_n^m, dany wzorem

$$\tilde{X}_n^m = \sum_{|\mathrm{i}| \leqslant n} (\psi_1(x_{\mathrm{i}}), \ldots, \psi_m(x_{\mathrm{i}}))\epsilon_{\mathrm{i}}^{\mathrm{dec}}.$$

Z twierdzenia o uniezależnianiu mamy

$$\mathbb{E}\|\tilde{X}_n^m\|^2 \leqslant C_d \mathbb{E}\|X_n^m\|^2 \leqslant 4C_d M^2.$$

Ponadto, zdefiniujmy przy ustalonym i_{I^c},

$$\tilde{Y}_n^m = \sum_{|\mathrm{i}_I| \leqslant n} (\psi_1(x_\mathrm{i}), \ldots, \psi_m(x_\mathrm{i})) \prod_{k \in I} \varepsilon_{i_k}^{(k)}$$

oraz

$$Y_n^m = \sum_{|\mathrm{i}_I| \leqslant n} (\psi_1(x_\mathrm{i}), \ldots, \psi_m(x_\mathrm{i})) \prod_{k \in I} \varepsilon_{i_k}.$$

Z zasady kontrakcji ([39], Theorem 4.4) wynika, że $\mathbb{E}\|\tilde{Y}_n^m\|^2 \leqslant \mathbb{E}\|\tilde{X}_n^m\|^2 \leqslant 4C_d M^2$. Zatem, znów z twierdzenia o uniezależnianiu, $\mathbb{E}\|Y_n^m\|^2 \leqslant 4C_d M^2$. Ale $(Y_n^m)_n$ jest martyngałem, stąd zbiega p.n. i w $L^2(\ell_\infty^m)$ do pewnej zmiennej Y_∞^m. To już pozwala nam zdefiniować $(\psi_j(Y))_{j \in \mathbb{N}} = (\sum_{|\mathrm{i}_I| < \infty} \psi_j(x_\mathrm{i}) \prod_{k \in I} \varepsilon_{i_k})_{j \in \mathbb{N}}$ z tezy lematu jako proces. Pozostaje wykazać, że z prawdopodobieństwem 1, $(\psi_j(Y))_{j \in \mathbb{N}} \in \ell_\infty$. Ale, ze zbieżności Y_n^m w $L^2(\ell_\infty^m)$ wynika, że $\mathbb{E}\sup_{j \leqslant m} |\psi_j(Y)|^2 \leqslant 4C_d M^2$, więc z twierdzenia Lebesgue'a o zbieżności monotonicznej $\mathbb{E}\sup_j |\psi_j(Y)|^2 < \infty$. □

Uwaga Założenie, że macierz $(x_\mathrm{i})_\mathrm{i}$ jest symetryczna (wykorzystane w dowodzie przy powołaniu się na twierdzenie o uniezależnianiu) jest w powyższym lemacie istotne. Rzeczywiście, już w przypadku jednowymiarowym, dla $d = 2$ oraz $x_{ij} = 1$ gdy $i > j$, $x_{ij} = -1$ gdy $i < j$, mamy $\sum_{i,j} x_{ij}\varepsilon_i\varepsilon_j = 0$, ale szereg $\sum_i x_{i1}\varepsilon_i$ nie jest zbieżny p.n.

LEMAT 2. *Istnieją stałe L_d (zależne tylko od d) oraz dla każdego $\alpha > 0$ istnieje stała $\varepsilon_d(\alpha)$ taka, że dla dowolnego chaosu rademacherowego X (rzędu d) i dowolnego M spełniającego $\mathbb{P}(\|X\| \geqslant M) < \varepsilon_d(\alpha)$, dla wszystkich $t \geqslant 1$ zachodzi nierówność*

$$\mathbb{P}(\|X\| > L_d M t^{d/2}) < e^{-\alpha t}.$$

DOWÓD. Zmienne Rademachera mają logarytmicznie wklęsłe „ogony", zatem do chaosów rademacherowych, zdefiniowanych przez skończone sumy, można stosować Wniosek 3. Łatwo sprawdzić, że w tym przypadku

$$\mathcal{A}_{k,p} = \mathcal{A}_p = \{(\alpha_i)\colon \sum_i \alpha_i^2 \leqslant p, |\alpha_i| \leqslant 1\}. \tag{1.7}$$

Standardowe metody, pozwalają natychmiast rozszerzyć nierówności Wniosku 3 na przypadek uogólnionych chaosów, wprowadzonych w Definicji 6 przy czym $\mathcal{T} = \{(\psi_j(x_\mathrm{i}))_\mathrm{i}\colon j \in \mathbb{N}\}$ (skończoność $\|\mathcal{T}\|_{\mathcal{N},I,t}$ wynika np. z istnienia drugiego momentu uogólnionego chaosu, wykazanego w dowodzie Lematu 1).

Zdefiniujmy

$$\phi(t) = \sum_{I \subseteq \{1,\ldots,d\}} \|\mathcal{T}\|_{\mathcal{N},I,t}.$$

Z (1.7) wynika, że

$$\phi(xt) \leqslant t^{d/2}\phi(x) \tag{1.8}$$

dla dowolnych $x \geqslant 0, t \geqslant 1$ (własności tej użyliśmy już w dowodzie Wniosku 2) Ponieważ ϕ jest niemalejąca, wynika stąd, że jest ciągła.

Jeśli $\phi(x) \leqslant K_d M$ dla wszystkich x (gdzie K_d to stała z Wniosku 3), to dla $t \geqslant 1$,

$$\mathbb{P}(\|X\| \geqslant MK_d^2 t^{d/2}) \leqslant \inf_x \mathbb{P}(\|X\| \geqslant K_d\phi(x)) \leqslant \inf_x e^{-x} = 0.$$

W przeciwnym razie, $K_d M = \phi(x)$ dla pewnego x. Zatem, jeśli $\varepsilon \geqslant \mathbb{P}(\|X\| \geqslant M)$, to z Wniosku 3,

$$\varepsilon \geqslant \mathbb{P}(\|X\| \geqslant \phi(x)/K_d) \geqslant c_d \wedge e^{-x},$$

co dla ε dostatecznie małych, daje $x \geqslant -\log\varepsilon$. Ponadto, z (1.8), $K_d M t^{d/2} \geqslant \phi(tx)$ dla $t \geqslant 1$, a zatem dla $\varepsilon < c_d \wedge e^{-\alpha}$,

$$\mathbb{P}(\|X\| \geqslant K_d^2 M t^{d/2}) \leqslant \mathbb{P}(\|X\| \geqslant K_d\phi(tx)) \leqslant e^{-tx} \leqslant e^{t\log\varepsilon} \leqslant e^{-\alpha t}.$$

□

Dowód Twierdzenia 10. Indukcja ze względu na d. Niech

$$S_n = \sum_{i_1,\ldots,i_d \geqslant n} x_{\mathrm{i}}\epsilon_{\mathrm{i}},$$

gdzie powyższa równość rozumiana jest zgodnie z Definicją 6. Zbieżność w sensie Definicji 6 wynika z Lematu 1, ponieważ

$$\sum_{n \leqslant i_1,\ldots,i_d \leqslant m} x_{\mathrm{i}}\epsilon_{\mathrm{i}} = \sum_{|\mathrm{i}| \leqslant m} x_{\mathrm{i}}\epsilon_{\mathrm{i}} - \sum_{I \subsetneq I_d} \sum_{|\mathrm{i}_{I^c}| < n} \prod_{k \in I^c} \varepsilon_{i_k} \Bigg[\sum_{n \leqslant i_k \leqslant m,\, k \in I} x_{\mathrm{i}} \prod_{k \in I} \varepsilon_{i_k} \Bigg], \tag{1.9}$$

a wyrażenia w nawiasach zgodnie z lematem, definiują w granicy (względem m) uogólnione chaosy o wartościach w ℓ^∞.

Ciąg $\|S_n\|$ jest odwrotnym podmartyngałem względem $\sigma(\varepsilon_i \colon i \geqslant n)$, a więc zbiega p.n. do pewnej zmiennej losowej, która, z prawa zero-jedynkowego, jest p.n. stała. Zatem istnieje $M > 0$ takie, że dla dowolnego $\varepsilon > 0$, dla odpowiednio dużych n, zachodzi $\mathbb{P}(\|S_n\| \geqslant M) < \varepsilon$. Z Lematu 2, dla dowolnego $\beta > 0$, istnieje n takie, że dla $t \geqslant 1$,

$$\mathbb{P}(\|S_n\| \geqslant L_d M t^{d/2}) < e^{-\beta t}.$$

Równoważnie

$$\mathbb{P}(\alpha\|S_n\|^{2/d} \geqslant t) \leqslant \exp\left(-\frac{\beta t}{\alpha(L_d M)^{2/d}}\right)$$

dla $t \geqslant (L_d M)^{2/d}\alpha$, skąd oczywiście wynika $\mathbb{E}e^{\alpha\|S_n\|^{2/d}} < \infty$ dla $\alpha < \beta/(L_d M)^{2/d}$. Z (1.9) wynika, że $\|S_1 - S_n\|$ jest ograniczone przez sumę chaosów niższego rzędu (dla $d \geqslant 2$) lub przez zmienną ograniczoną ($d = 1$), a zatem (dla $d \geqslant 2$ razem z założeniem indukcyjnym) nierówność Höldera daje

$$\mathbb{E}e^{\alpha\|X\|^{2/d}} < \infty$$

dla $\alpha < \beta/(L_d M)^{d/2}$. To kończy dowód, gdyż β może być wybrane dowolnie. $\square$

Rozdział 2

Oszacowania momentów dla U-statystyk o wartościach w przestrzeniach Banacha

Celem niniejszego rozdziału jest znalezienie odpowiedników Twierdzenia 7 dla kanonicznych U-statystyk o wartościach w ośrodkowej przestrzeni Banacha $(B, \|\cdot\|)$.

Rozważmy zatem niezależne zmienne losowe $(X_i^{(k)})_{i \in I_n, k \in I_d}$ o wartości w przestrzeni $(\Sigma, \mathcal{B})$. Zgodnie z założeniami z rozdziału Ogólne założenia i notacja Σ jest przestrzenią polską, zaś $\mathcal{B}$ – σ-ciałem zbiorów borelowskich. Rozważmy także d-wskaźnikową macierz $(h_{\mathrm{i}})_{|\mathrm{i}| \leqslant n}$ funkcji borelowskich $h_{\mathrm{i}} \colon \Sigma^d \to B$, *kanonicznych* (*całkowicie zdegenerowanych*) ze względu na $(X_i^{(k)})_{i,k}$, tzn. takich że

$$\mathbb{E}_k h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}) = 0,$$

dla dowolnego $|\mathrm{i}| \leqslant n$ oraz $k \in I_d$.

Głównym obiektem rozważań będzie uniezależniona U-statystyka

$$Z = \sum_{|\mathrm{i}| \leqslant n} h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}).$$

Wprowadźmy teraz wielkości, za pomocą których oszacujemy momenty zmiennej Z. Nasze oszacowania będą dotyczyć p-tych momentów, dla $p \geqslant 2$, przy czym występować w nich będą w szczególności p-te momenty poszczególnych składników U-statystyki (podobnie jak w znanych nierównościach dla sum niezależnych zmiennych losowych). Dlatego też możemy założyć, że te momenty istnieją, w przeciwnym razie nasze oszacowania będą trywialne (i oczywiście żadne nietrywialne oszacowania nie mogą być uzyskane). W szczególności dla każdego i, $\mathbb{E}\|h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\|^2 < \infty$, co sprawia, że całki w poniższej definicji są dobrze określone.

DEFINICJA 7. *Dla dowolnych $J \subseteq I \subseteq I_d$, $I \neq \emptyset$ oraz ustalonej wartości* i_{I^c}, *zdefiniujmy*

$$\|\!|(h_{\mathrm{i}})_{\mathrm{i}_I}|\!\|_{I,J} = \mathbb{E}_{I\setminus J} \sup \Big\{ \mathbb{E}_J \sum_{|\mathrm{i}_I| \leqslant n} \langle \phi, h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\rangle \prod_{j\in J} f_{i_j}^{(j)}(X_{i_j}^{(j)})\colon \phi \in B^*, \|\phi\| \leqslant 1,$$
$$f_i^{(j)}\colon \Sigma \to \mathbb{R}, j \in J, i \in I_n, \mathbb{E}\sum_i |f_i^{(j)}(X_i^{(j)})|^2 \leqslant 1, j \in J \Big\}. \tag{2.1}$$

Zdefiniujmy ponadto $\|\!|(h_{\mathrm{i}})_{\mathrm{i}_\emptyset}|\!\|_{\emptyset,\emptyset} = \|h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\|$.

Uwaga Powyższa definicja zależy oczywiście od $(X_i^{(k)})_{i,k}$, aby skrócić notację nie będziemy tego jednak explicite zaznaczać. Zauważmy także, że dla $I = I_d$, $\|\!|(h_i)_{\mathrm{i}}|\!\|_{I,J}$ jest „deterministyczną" normą, podczas gdy dla $I \neq I_d$, jest zmienną losową, zależną od $(X_{i_k}^{(k)})_{k\in I^c}$. Warto także zauważyć, że dla $J = \emptyset$, mamy $\|\!|(h_{\mathrm{i}})_{\mathrm{i}}|\!\|_{I,J} = \mathbb{E}_I \| \sum_{\mathrm{i}_I} h_i(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\|$.

W przypadku, gdy $\Sigma = \mathbb{R}$ oraz $h_{\mathrm{i}}(x_1, \ldots, x_d) = a_{\mathrm{i}} \prod_{k=1}^d x_k$, dla pewnych $a_{\mathrm{i}} \in B$ (tzn., gdy U-statystyka jest uniezależnionym chaosem o wartościach w B) oraz $\mathbb{E}(X_i^{(j)})^2 = 1$ dla dowolnych $i \in I_n, j \in I_d$, mamy

$$\|(h_{\mathrm{i}})_{\mathrm{i}_I}\|_{I_d,J} = \|\{(\langle \phi, a_{\mathrm{i}}\rangle)_{\mathrm{i}}\colon \phi \in \Phi\}\|_J,$$

gdzie Φ jest dowolnym przeliczalnym gęstym podzbiorem kuli jednostkowej w przestrzeni sprzężonej do $\mathrm{span}\{a_{\mathrm{i}}\colon |\mathrm{i}| \leqslant n\}$, zaś prawa strona została zdefiniowana w Definicji 1.

Poniżej przedstawimy kilka oszacowań momentów zmiennej Z, w ogólnej przestrzeni Banacha oraz przy pewnych dodatkowych założeniach odnośnie własności geometrycznych przestrzeni.

2.1 Nierówności dla U-statystyk w ogólnych przestrzeniach Banacha

Głównym rezultatem niniejszej sekcji jest

TWIERDZENIE 11. *Istnieją stałe K_d takie, że dla $p \geqslant 2$ zachodzi*

$$\mathbb{E}\|Z\|^p \leqslant K_d^p \Bigg[\sum_{I\subseteq I_d} \sum_{J \subseteq I} p^{p(\#J/2 + \#I^c)} \sum_{\mathrm{i}_{I^c}} \mathbb{E}_{I^c} \|\!|(h_{\mathrm{i}})_{\mathrm{i}_I}|\!\|_{I,J}^p \Bigg]. \tag{2.2}$$

Ponadto, jeśli dla wszystkich $|\mathrm{i}| \leqslant n$, $\|h_{\mathrm{i}}\|_\infty \leqslant \kappa$, to

$$\mathbb{E}\|Z\|^p \leqslant K_d^p \Big(\Big[\sum_{J\subseteq I_d} p^{\#Jp/2} \|\!|(h_{\mathrm{i}})_{\mathrm{i}}|\!\|_{I_d,J}^p \Big] + n^{p(d-1)} p^{p(d+1)/2} \kappa^p \Big). \tag{2.3}$$

Dowód powyższego twierdzenia opierać się będzie na nierówności Talagranda dla supremów procesów empirycznych [47], a dokładniej, na oszacowaniach momentów takich supremów, wywnioskowanych z nierówności Talagranda przez Giné, Latałę i Zinna [18], które można otrzymać także z ogólnych oszacowań momentów funkcji niezależnych zmiennych losowych, podanych przez Boucherona, Bousqueta, Lugosi i Massarta [8].

Lemat 3 ([18, Proposition 3.1],[8, Theorem 12]). *Niech $X_1, \ldots, X_n$ będą niezależnymi zmiennymi losowymi o wartościach w $(\Sigma, \mathcal{B})$ zaś $\mathcal{T}$ przeliczalną klasą rzeczywistych funkcji borelowskich, określonych na Σ, taką że dla każdej funkcji $f \in \mathcal{T}$ oraz $i \in I_n$, $\mathbb{E}f(X_i) = 0$ i $\mathbb{E}f(X_i)^2 < \infty$. Rozważmy zmienną losową $S := \sup_{f\in\mathcal{T}} |\sum_i f(X_i)|$. Wówczas, dla $p \geqslant 1$,*

$$\mathbb{E}S^p \leqslant K^p \left[(\mathbb{E}S)^p + p^{p/2}\sigma^p + p^p \mathbb{E} \max_i \sup_{f\in\mathcal{T}} |f(X_i)|^p \right],$$

gdzie

$$\sigma^2 = \sup_{f\in\mathcal{T}} \sum_i \mathbb{E}f(X_i)^2,$$

zaś K jest stałą uniwersalną.

W dowodzie Twierdzenia 11 będziemy potrzebować następującego prostego wniosku z Lematu 3.

Lemat 4. *Niech B będzie przestrzenią Banacha, dla której istnieje przeliczalny zbiór $D = \{\psi_j\}$ funkcjonałów taki, że dla dowolnego $x \in B$*

$$\|x\|_B = \sup_j |\psi_j(x)|.$$

Niech ponadto $X_1, \ldots, X_n, Y$ będą niezależnymi zmiennymi losowymi o wartościach w $(\Sigma, \mathcal{B})$. Rozważmy $E = L^1_Y(B)$ - przestrzeń wszystkich całkowalnych funkcji o wartościach w B, postaci $f(Y)$, takich że $\psi_j \circ f$ jest mierzalne dla wszystkich j. Rozważmy funkcje $h_i\colon \Sigma^2 \to B$ $(i = 1, \ldots, n)$ takie, że $\psi_j \circ h_i$ jest mierzalne dla wszystkich j, $\mathbb{E}_X h_i(X_i, Y) = 0$, Y-p.n. oraz $h_i(X_i, Y) \in E$, X-p.n. Niech $S = \sum_i h_i(X_i, Y) \in E$. Wówczas dla wszystkich $p \geqslant 2$,

$$\begin{aligned} \mathbb{E}\|S\|_E^p &\leqslant K^p \left[(\mathbb{E}\|S\|_E)^p + p^{p/2}\sigma^p + p^p \mathbb{E}_X \max_i \|h_i(X_i, Y)\|_E^p \right] \\ &\leqslant K^p \left[(\mathbb{E}\|S\|_E)^p + p^{p/2}\sigma^p + p^p \mathbb{E}_X \sum_i \|h_i(X_i, Y)\|_E^p \right], \end{aligned}$$

gdzie

$$\begin{aligned} \sigma &= \sup_{f=(f_i(X_i))\colon \sum \mathbb{E}f_i^2(X_i) \leqslant 1} \mathbb{E}_Y \sup_j |\sum_i \mathbb{E}_X \psi_j(h_i(X_i, Y) f_i(X_i))| \\ &\leqslant \mathbb{E}_Y \sup_j (\sum_i \mathbb{E}_X \psi_j(h_i(X_i, Y))^2)^{1/2}. \end{aligned}$$

Dowód. Skonstruujemy najpierw przeliczalny zbiór wektorów postaci $\phi(Y) = (\phi_1(Y), \phi_2(Y), \ldots)$ taki, że $\sum_j |\phi_j(Y)| = 1$ p.n. oraz dla wszystkich $g(Y) \in E$,

$$\|g(Y)\|_E = \sup_{\phi} |\sum_j \mathbb{E}_Y \phi_j(Y)\psi_j(g(Y))|. \tag{2.4}$$

Zauważmy, że dla każdego wektora losowego $k(Y) = (k_1(Y), \ldots, k_n(Y)) \in L^1_Y(\ell^n_\infty)$ istnieje wektor $\phi(Y) = (\phi_1(Y), \ldots, \phi_n(Y))$ taki, że

$$\sum_{j=1}^{n} |\phi_j(Y)| = 1 \text{ p.n.} \tag{2.5}$$

oraz

$$\mathbb{E} \max_{j \leqslant n} |k_j(Y)| = |\sum_{j=1}^{n} \mathbb{E}\phi_j(Y)k_j(Y)|,$$

gdyż dla każdej wartości Y możemy zdefiniować $\phi_l(Y) = \operatorname{sgn} k_l(Y)$ dla $l = \min\{i \leqslant n\colon |k_i(Y)| = \max_{j\leqslant n} |k_j(Y)|\}$ oraz $\phi_l(Y) = 0$ dla pozostałych wartości l.

Ponieważ wszystkie takie ciągi $\phi(Y)$, jako funkcjonały na $L^1_Y(\ell^n_\infty)$, mają normę 1 oraz $L^1_Y(\ell^n_\infty)$ jest ośrodkowa (z założenia, że Σ jest przestrzenią polską), istnieje przeliczalny zbiór $\mathcal{T}_n$ wektorów $\phi(Y)$, spełniających (2.5) i takich, że

$$\mathbb{E} \max_{j \leqslant n} |k_j(Y)| = \sup_{\phi \in \mathcal{T}_n} |\sum_{j=1}^{n} \mathbb{E}\phi_j(Y)k_j(Y)|$$

dla wszystkich $k(Y) \in L^1_Y(\ell^n_\infty)$. Ponadto, dla $g(Y) \in E$,

$$\|g(Y)\|_E = \sup_n \mathbb{E} \max_{j\leqslant n} |\psi_j(g(Y))| = \sup_n \sup_{\phi \in \mathcal{T}_n} |\sum_{j \leqslant n} \mathbb{E}\phi_j(Y)\psi_j(g(Y))|,$$

więc, aby otrzymać (2.4), wystarczy wziąć zbiór składających się ze wszystkich wektorów z $\phi(Y) \in \bigcup \mathcal{T}_n$, uzupełnionych zerami do ciągów nieskończonej długości.

Zachodzi zatem równość

$$\|S\|_E = \sup_{\phi} |\sum_i \sum_j \mathbb{E}_Y \psi_j(h_i(X_i, Y))\phi_j(Y)| = \sup_{\phi} |\sum_i g^i_{\phi}(X_i)|,$$

co pozwala oszacować $\|S\|_p$ przy pomocy Lematu 3 (formalnie dotyczy on sytuacji, gdy do wszystkich X_i jest przykładana ta sama funkcja, ale łatwo zauważyć, że prosty zabieg formalny pozwala zastosować go także w naszym przypadku. Zamiast zmiennych X_i można mianowicie rozpatrywać zmienne $\tilde{X}_i = (X_i, i)$ oraz funkcję $g_\phi(x, i) = g^i_\phi(x)$, co pozwala zastosować Lemat 3, gdyż jego założenia nie wymagają identycznych rozkładów zmiennych, definiujących proces empiryczny). Mamy

$$\begin{aligned}\sigma &= \sup_\phi (\sum_i \mathbb{E}_X g^i_\phi(X_i)^2)^{1/2} \\ &= \sup_{f=(f_i):\ \sum_i \mathbb{E} f_i(X_i)^2 \leqslant 1} \sup_\phi |\sum_j \mathbb{E}_Y \sum_i \mathbb{E}_X \psi_j(h_i(X_i,Y) f_i(X_i))\phi_j(Y)| \\ &= \sup_f \mathbb{E}_Y \sup_j |\sum_i \mathbb{E}_X \psi_j(h_i(X_i,Y) f_i(X_i))|.\end{aligned}$$

□

Będziemy także potrzebowali następującego prostego faktu, wiążącego normy $\|\!|(h_{\mathrm{i}})_{\mathrm{i}_I}|\!\|_{I,J}$ z warunkowymi drugimi momentami odpowiednich U-statystyk.

Lemat 5. *Dla dowolnej przestrzeni Banacha B, każdego $J \subseteq I \subseteq I_d$ oraz dowolnej ustalonej wartości i_{I^c} zachodzi*

$$\|\!|(h_{\mathrm{i}})_{\mathrm{i}_I}|\!\|_{I,J} \leqslant \sqrt{\mathbb{E}_I \| \sum_{\mathrm{i}_I} h_{\mathrm{i}}(\mathbf{X}^{\mathrm{dec}}_{\mathrm{i}})\|^2}.$$

Dowód. Z nierówności Cauchy'ego-Schwarza i kanoniczności h_{i},

$$\begin{aligned}\|\!|(h_{\mathrm{i}})_{\mathrm{i}_I}|\!\|^2_{I,J} &\leqslant \mathbb{E}_{I\setminus J} \sup_{\phi \in B^*, \|\phi\| \leqslant 1} \mathbb{E}_J \sum_{\mathrm{i}_J} \langle \phi, \sum_{\mathrm{i}_{I\setminus J}} h_{\mathrm{i}}(\mathbf{X}^{\mathrm{dec}}_{\mathrm{i}})\rangle^2 \\ &= \mathbb{E}_{I\setminus J} \sup_{\phi \in B^*, \|\phi\| \leqslant 1} \mathbb{E}_J \langle \phi, \sum_{\mathrm{i}_I} h_{\mathrm{i}}(\mathbf{X}^{\mathrm{dec}}_{\mathrm{i}})\rangle^2 \leqslant \mathbb{E}_I \| \sum_{\mathrm{i}_I} h_{\mathrm{i}}(\mathbf{X}^{\mathrm{dec}}_{\mathrm{i}})\|^2.\end{aligned}$$

□

Dowód Twierdzenia 11. Indukcja ze względu na d. Dla $d = 1$, obie nierówności są prostym wnioskiem z Lematu 3, gdyż

$$\begin{aligned}\|\!|(h_i)_i|\!\|_{\{1\},\emptyset} &= \mathbb{E}\|Z\|, \\ \|\!|(h_i)_i|\!\|_{\{1\},\{1\}} &= \sup_{\|\phi\|\leqslant 1} \sqrt{\sum_i \mathbb{E}\langle \phi, h_i(X_i)\rangle^2} \\ \sum_i \mathbb{E}\|\!|(h_i)_i|\!\|^p_{\emptyset,\emptyset} &= \sum_i \mathbb{E}\|h_i(X_i)\|^p \geqslant \mathbb{E}\max_i \|h_i(X_i)\|^p\end{aligned}$$

oraz jeśli $\|h_i\|_\infty \leqslant \kappa$, to oczywiście $\mathbb{E}\max_i \|h_i\|^p \leqslant \kappa^p$.

Rozważmy najpierw nierówność (2.2). Załóżmy, że jest ona prawdziwa dla wszystkich liczb mniejszych niż d. Oznaczmy $\tilde{I}^c = I^c \setminus \{d\}$ dla $I \subseteq I_d$. Założenie indukcyjne dla $d_1 = d - 1$, zastosowane warunkowo na $X^{(d)}$ razem z Twierdzeniem Fubiniego, daje

$$\mathbb{E}\|Z\|^p \leqslant K^p_{d-1} \sum_{I \subseteq \{1,\ldots,d-1\}} \sum_{J \subseteq I} \left[p^{p(\#J/2 + \#\tilde{I}^c)} \sum_{\mathrm{i}_{\tilde{I}^c}} \mathbb{E}_{\tilde{I}^c} \mathbb{E}_d \|\!|(\sum_{i_d} h_{\mathrm{i}})_{\mathrm{i}_I}|\!\|^p_{I,J} \right]. \tag{2.6}$$

Zauważmy, że z ośrodkowości B oraz $\oplus_{i=1}^{n} L^2(X_i^{(j)})$, w Definicji 7 możemy zawęzić supremum do przeliczalnego zbioru funkcjonałów i funkcji. Zatem $\mathbb{E}_d \|\!|(\sum_{i_d} h_{\mathrm{i}})_{\mathrm{i}_I}\|\!|_{I,J}^p$ może być oszacowane przy użyciu Lematu 4 (zastosowanego warunkowo względem $(X_{i_k}^{(k)})_{k\in\tilde{I}^c}$ dla $I \neq I_{d-1}$, z $Y = (X_i^{(j)})_{i\in I_n, j\in I\setminus J}$). Mamy

$$p^{p(\#J/2+\#\tilde{I}^c)} \sum_{\mathrm{i}_{\tilde{I}^c}} \mathbb{E}_{\tilde{I}^c}(\mathbb{E}_d \|\!|(\sum_{i_d} h_{\mathrm{i}})_{\mathrm{i}_I}\|\!|_{I,J})^p = p^{p(\#J/2+\#(I\cup\{d\})^c)} \\ \times \sum_{\mathrm{i}_{(I\cup\{d\})^c}} \mathbb{E}_{(I\cup\{d\})^c} \|\!|(h_{\mathrm{i}})_{\mathrm{i}_{I\cup\{d\}}}\|\!|_{I\cup\{d\},J}^p.$$

Dodatkowo, σ z lematu jest ograniczone przez $\|\!|(h_{\mathrm{i}})_{\mathrm{i}_{I\cup\{d\}}}\|\!|_{I\cup\{d\},J\cup\{d\}}$ oraz

$$p^{p(\#J/2+\#\tilde{I}^c)} \sum_{\mathrm{i}_{\tilde{I}^c}} \mathbb{E}_{\tilde{I}^c} p^{p/2} \|\!|(h_{\mathrm{i}})_{\mathrm{i}_{I\cup\{d\}}}\|\!|_{I\cup\{d\},J\cup\{d\}}^p = p^{p(\#(J\cup\{d\})/2+\#(I\cup\{d\})^c)} \\ \times \sum_{\mathrm{i}_{(I\cup\{d\})^c}} \mathbb{E}_{(I\cup\{d\})^c} \|\!|(h_{\mathrm{i}})_{\mathrm{i}_{I\cup\{d\}}}\|\!|_{I\cup\{d\},J\cup\{d\}}^p.$$

Aby zakończyć dowód nierówności (2.2) pozostaje zauważyć, że

$$p^{p(\#J/2+\#\tilde{I}^c)} \sum_{\mathrm{i}_{\tilde{I}^c}} p^p \sum_{i_d} \mathbb{E}_{\tilde{I}^c} \mathbb{E}_d \|\!|(h_{\mathrm{i}})_{\mathrm{i}_I}\|\!|_{I,J}^p \quad = \quad p^{p(\#J/2+\#I^c)} \sum_{\mathrm{i}_{I^c}} \mathbb{E}_{I^c} \|\!|(h_{\mathrm{i}})_{\mathrm{i}_I}\|\!|_{I,J}^p.$$

Krok indukcyjny w przypadku nierówności (2.3) jest analogiczny. Z założenia indukcyjnego

$$\mathbb{E}\|\sum_{\mathrm{i}} h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\|^p \leqslant K_{d-1}^p \Big(\Big[\sum_{J\subseteq I_{d-1}} p^{\#Jp/2} \mathbb{E}_d \|(\sum_{i_d} h_{\mathrm{i}})_{\mathrm{i}_{I_{d-1}}}\|_{I_{d-1},J}^p \Big] \\ + p^{dp/2} n^{p(d-2)} (n\kappa)^p \Big). \tag{2.7}$$

Podobnie jak w przypadku pierwszej nierówności, z Lematu 4 otrzymujemy

$$\mathbb{E}_d \|\!|(\sum_{i_d} h_{\mathrm{i}})_{\mathrm{i}_{I_{d-1}}}\|\!|_{I_{d-1},J}^p \leqslant K^p \Big[\|\!|(h_{\mathrm{i}})_{\mathrm{i}}\|\!|_{I_d,J}^p \\ + p^{p/2} \|\!|(h_{\mathrm{i}})_{\mathrm{i}}\|\!|_{I_d,J\cup\{d\}}^p \\ + p^p \mathbb{E}_d \max_{i_d} \|\!|(h_{\mathrm{i}})_{\mathrm{i}_{I_{d-1}}}\|\!|_{I_{d-1},J}^p \Big].$$

Zauważmy, że dla każdego i_d, z Lematu 5

$$\|\!|(h_{\mathrm{i}})_{\mathrm{i}_{I_{d-1}}}\|\!|_{I_{d-1},J} \leqslant \sqrt{\mathbb{E}_{I_{d-1}} \Big\| \sum_{\mathrm{i}_{I_{d-1}}} h_{\mathrm{i}} \Big\|^2} \leqslant n^{(d-1)}\kappa.$$

Otrzymujemy zatem

$$\mathbb{E}_d |\!|\!|(\sum_{i_d} h_{\mathrm{i}})_{\mathrm{i}_{I_{d-1}}} |\!|\!|^p_{I_{d-1},J} \leqslant K^p \Big(|\!|\!|(h_{\mathrm{i}})|\!|\!|^p_{I_d,J} + p^{p/2} |\!|\!|(h_{\mathrm{i}})|\!|\!|_{I_d,J\cup\{d\}} + p^p n^{p(d-1)} \kappa^p \Big).$$

Wstawiając powyższą nierówność do (2.7) otrzymujemy tezę (zauważmy, że $p\#J/2+p \leqslant p(d-1)/2+p = p(d+1)/2$).

□

Uwaga W kontekście uwagi o związkach między normami $|\!|\!|\cdot|\!|\!|_{I,J}$ oraz $\|\mathcal{T}\|$, poczynionej po Definicji 7, Twierdzenie 11 może być uważane za odpowiednik nierówności dla chaosów, przedstawionych w Sekcji 1. Ponadto, nierówność (2.3) jest wzmocnieniem wyniku uzyskanego niedawno przez Giné i Masona [19], w którym nie występują normy $|\!|\!|\cdot|\!|\!|_{I,J}$, a jedynie wariancja ze współczynnikiem $p^{pd/2}$, odpowiadającym w (2.3) $J = I_d$.

2.2 U-statystyki w przestrzeniach typu 2

W następnej kolejności zaprezentujemy wzmocnienie przedstawionych dotąd nierówności, które możemy uzyskać po nałożeniu na przestrzeń B pewnych warunków geometrycznych. Będziemy rozważać przestrzenie typu 2, które pojawiają się dość naturalnie w rachunku prawdopodobieństwa, np. przy badaniu praw iterowanego logarytmu (dla przykładu, w przestrzeniach tych upraszczają się warunki konieczne i dostateczne na prawo iterowanego logarytmu dla sum niezależnych zmiennych losowych [39], z kolei dla U-statystyk, całkowalność z kwadratem jądra U-statystyki jest warunkiem dostatecznym dla prawa iterowanego logarytmu [12]).

Zacznijmy od przypomnienia podstawowych definicji.

DEFINICJA 8. *Przestrzeń Banacha B nazywamy przestrzenią typu 2, ze stałą K, jeśli dla dowolnego skończonego ciągu $(x_i)_{i\leqslant n}$, elementów B, zachodzi*

$$\Big\| \sum_{i=1}^n x_i \varepsilon_i \Big\|_2 \leqslant K \Big(\sum_{i=1}^n \|x_i\|^2 \Big)^{1/2},$$

gdzie $\varepsilon_1, \ldots, \varepsilon_n$ to niezależne zmienne Rademachera.

Przestrzeń B nazywamy przestrzenią kotypu 2, ze stałą K, jeśli dla dowolnego skończonego ciągu $(x_i)_{i\leqslant n}$, elementów B, zachodzi

$$\Big(\sum_{i=1}^n \|x_i\|^2 \Big)^{1/2} \geqslant \frac{1}{K} \Big\| \sum_{i=1}^n x_i \varepsilon_i \Big\|_2$$

Stosując klasyczne nierówności symetryzacyjne ([39], Lemma 6.3), otrzymujemy następujący

Fakt 2. *Jeżeli B jest przestrzenią typu 2 ze stałą K, zaś $X_1, \ldots, X_n$ niezależnymi zmiennymi losowymi o średniej* 0, *to*

$$\mathbb{E}\Big\|\sum_{i=1}^{n} X_i\Big\|^2 \leqslant 4K^2 \sum_{i=1}^{n} \mathbb{E}\|X_i\|^2.$$

Podobnie, dla kanonicznych U-statystyk rzędu d zachodzi nierówność

$$\mathbb{E}\Big\|\sum_{|\mathrm{i}|\leqslant n} h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big\|^2 \leqslant (4K^2)^d \sum_{|\mathrm{i}|\leqslant n} \mathbb{E}\|h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\|^2.$$

Łącząc powyższą nierówność z Lematem 5, otrzymujemy

Lemat 6. *Jeśli B jest typu* 2, *to*

$$\|\!|(h_{\mathrm{i}})_{\mathrm{i}_I}|\!\|_{I,J} \leqslant K_d(B)\sqrt{\sum_{\mathrm{i}_I} \mathbb{E}_I\|h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\|^2},$$

gdzie $K_d(B)$, zależy tylko od d i stałej w definicji typu 2 dla przestrzeni B.

Twierdzenie 12. *Niech B będzie przestrzenią Banacha typu 2. Istnieją stałe $K_d(B)$, zależne tylko od d i stałej w definicji typu 2 dla B takie, że dla $p \geqslant 2$,*

$$\begin{aligned}\mathbb{E}|Z|^p \quad \leqslant \quad & K_d(B)^p \Bigg[\max_{I\subseteq I_d} \mathbb{E}_{I^c} \max_{\mathrm{i}_{I^c}}(\sum_{\mathrm{i}_I} \mathbb{E}_I\|h_{\mathrm{i}}\|^2)^{p/2} \qquad (2.8)\\ & + \sum_{I\subseteq I_d}\sum_{J\subseteq I} p^{p(\#J/2+\#I^c)}\mathbb{E}_{I^c}\max_{\mathrm{i}_{I^c}} \|\!|(h_{\mathrm{i}})_{\mathrm{i}_I}|\!\|_{I,J}^p\Bigg].\end{aligned}$$

Naszym celem jest zatem zastąpienie zewnętrznych sum w nierówności z Twierdzenia 11 przez maksima. Wykorzystamy następujące fakty.

Lemat 7 ([18], nierówność (2.6)). *Niech $\xi_1, \ldots, \xi_N$ będą niezależnymi, nieujemnymi zmiennymi losowymi. Wówczas, dla $p > 1$ oraz $\alpha > 0$, zachodzi*

$$p^{\alpha p}\sum_i \mathbb{E}\xi_i^p \leqslant 2(1+p^\alpha)\max\left[p^{\alpha p}\,\mathbb{E}\max_i \xi_i^p, (\sum_i \mathbb{E}\xi_i)^p\right].$$

W poniższych lematach, dla uproszczenia notacji, będziemy pisać g_{i} na oznaczenie $g_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})$.

Lemat 8 ([18], Corollary 2.2.). *Rozważmy nieujemne funkcje $g_{\mathrm{i}}\colon \Sigma^d \to \mathbb{R}_+$. Dla dowolnego $p \geqslant 1$,*

$$\max_{I\subseteq I_d}\mathbb{E}_I\max_{\mathrm{i}_I}(\sum_{\mathrm{i}_{I^c}}\mathbb{E}_{I^c}g_{\mathrm{i}})^p \leqslant \mathbb{E}(\sum_{\mathrm{i}} g_{\mathrm{i}})^p \leqslant K_d^p\sum_{I\subseteq I_d} p^{\#Ip}\mathbb{E}_I\max_{\mathrm{i}_I}(\sum_{\mathrm{i}_{I^c}}\mathbb{E}_{I^c}g_{\mathrm{i}})^p.$$

Lemat 9. *Dla $\alpha > 0$ i dowolnych nieujemnych funkcji $g_{\mathrm{i}}\colon \Sigma^d \to \mathbb{R}_+$ oraz $p > 1$ zachodzi*

$$p^{\alpha p}\sum_{\mathrm{i}\in I_n^d}\mathbb{E}g_{\mathrm{i}}^p \leqslant K_d^p p^{\alpha d}\left[p^{\alpha p}\mathbb{E}\max_{\mathrm{i}} g_{\mathrm{i}}^p + \sum_{I\subsetneq I_d} p^{\#Ip}\mathbb{E}_I \max_{\mathrm{i}_I}(\sum_{\mathrm{i}_{I^c}}\mathbb{E}_{I^c}g_{\mathrm{i}})^p\right].$$

Dowód Lematu 9. Indukcja ze względu na d. Dla $d = 1$, nierówność, z $K_1 = 4$, wynika z Lematu 7. Załóżmy, że teza zachodzi dla wszystkich liczb mniejszych od $d \geqslant 2$. Stosując założenie indukcyjne do $\mathbb{E}_{\{1,\ldots,d-1\}}$ i używając tej samej notacji, co w dowodzie Twierdzenia 11, otrzymujemy

$$\begin{aligned} p^{\alpha p}\sum_{\mathrm{i}}\mathbb{E}g_{\mathrm{i}}^p \quad \leqslant \quad & K_{d-1}^p p^{\alpha(d-1)}\mathbb{E}_d\sum_{i_d}\Big[p^{\alpha p}\mathbb{E}_{\{d\}^c}\max_{\mathrm{i}_{\{d\}^c}} g_{\mathrm{i}}^p \\ & + \sum_{I\subsetneq\{1,\ldots,d-1\}} p^{\#Ip}\mathbb{E}_I\max_{\mathrm{i}_I}(\sum_{\mathrm{i}_{\tilde{I}^c}}\mathbb{E}_{\tilde{I}^c}g_{\mathrm{i}})^p\Big]. \end{aligned}$$

Lemat 7 oraz Lemat 8 dają

$$\begin{aligned} K_{d-1}^p p^{\alpha(d-1)}\mathbb{E}_d\sum_{\mathrm{i}_d} p^{\alpha p}\mathbb{E}_{\{d\}^c}\max_{\mathrm{i}_{\{d\}^c}} g_{\mathrm{i}}^p &= K_{d-1}^p p^{\alpha(d-1)}\mathbb{E}_{\{d\}^c}\mathbb{E}_d\sum_{i_d} p^{\alpha p}\max_{\mathrm{i}_{\{d\}^c}} g_{\mathrm{i}}^p \\ \leqslant K_d^p p^{\alpha d} p^{\alpha p}\mathbb{E}\max_{\mathrm{i}} g_{\mathrm{i}}^p &+ K_d^p p^{\alpha d}\mathbb{E}_{\{d\}^c}(\sum_{\mathrm{i}}\mathbb{E}_d g_{\mathrm{i}})^p \\ \leqslant K_d^p p^{\alpha d} p^{\alpha p}\mathbb{E}\max_{\mathrm{i}} g_{\mathrm{i}}^p &+ K_d^p p^{\alpha d}\sum_{I\subseteq\{1,\ldots,d-1\}} p^{\#Ip}\mathbb{E}_I\max_{\mathrm{i}_I}(\sum_{\mathrm{i}_{I^c}}\mathbb{E}_{I^c}g_{\mathrm{i}})^p. \end{aligned}$$

Ponadto dla każdego $I \subsetneq \{1,\ldots,d-1\}$ z Lematu 7 (zastosowanego z „$\alpha = \#I$”), faktu, że $p^{\#I} \leqslant p^d \leqslant K_d^p$ oraz Lematu 8, otrzymujemy

$$\begin{aligned} K_{d-1}^p p^{\alpha(d-1)}&\sum_{i_d}\mathbb{E}_d p^{\#Ip}\mathbb{E}_I\max_{\mathrm{i}_I}(\sum_{\mathrm{i}_{\tilde{I}^c}}\mathbb{E}_{\tilde{I}^c}g_{\mathrm{i}})^p \\ &= K_{d-1}^p p^{\alpha(d-1)}\mathbb{E}_I p^{\#Ip}\sum_{i_d}\mathbb{E}_d\max_{\mathrm{i}_I}(\sum_{\mathrm{i}_{\tilde{I}^c}}\mathbb{E}_{\tilde{I}^c}g_{\mathrm{i}})^p \\ &\leqslant K_d^p p^{\alpha(d-1)}p^{\#I}p^{\#Ip}\mathbb{E}_{I\cup\{d\}}\max_{\mathrm{i}_{I\cup\{d\}}}(\sum_{\mathrm{i}_{(I\cup\{d\})^c}}\mathbb{E}_{(I\cup\{d\})^c}g_{\mathrm{i}})^p \\ &\quad + K_d^p p^{\alpha(d-1)}p^{\#I}\mathbb{E}_I(\sum_{\mathrm{i}}\mathbb{E}_{I^c}g_{\mathrm{i}})^p \\ &\leqslant K_d^p p^{d\alpha}p^{\#Ip}\mathbb{E}_{I\cup\{d\}}\max_{\mathrm{i}_{I\cup\{d\}}}(\sum_{\mathrm{i}_{(I\cup\{d\})^c}}\mathbb{E}_{(I\cup\{d\})^c}g_{\mathrm{i}})^p \\ &\quad + K_d^p p^{\alpha d}\sum_{J\subseteq I} p^{\#Jp}\mathbb{E}_J\max_{\mathrm{i}_J}(\sum_{\mathrm{i}_{J^c}}\mathbb{E}_{J^c}g_{\mathrm{i}})^p. \end{aligned}$$

□

Dowód Twierdzenia 12. Dla $p = 2$, teza jest oczywista, ponieważ w przestrzeniach typu 2 zachodzi

$$\mathbb{E}\|\sum_{\mathrm{i}} h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\|^2 \leqslant K_d(B)\sum_{\mathrm{i}}\mathbb{E}\|h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\|^2.$$

Dla $p > 2$, wychodząc z Twierdzenia 11, a następnie stosując do $\sum_{\mathrm{i}_{I^c}}\mathbb{E}_{I^c}|\!|\!|(h_{\mathrm{i}})_{\mathrm{i}_I}|\!|\!|_{I,J}^p$ Lemat 9 z „$p = p/2$" i $\alpha = 2(\#I^c + \#J/2) + \#I^c \leqslant 4d$ (oraz korzystając z faktu, że $(p/2)^{\alpha d} \leqslant K_d^p$), otrzymujemy

$$\mathbb{E}\|Z\|^p \leqslant K_d^p\Big[\sum_{I\subseteq I_d}\sum_{J\subseteq I}\sum_{K\subsetneq I^c}\left(\frac{p}{2}\right)^{\#Kp/2-\#I^cp/2}\mathbb{E}_K\max_{\mathrm{i}_K}\Big(\sum_{\mathrm{i}_{I^c\setminus K}}\mathbb{E}_{I^c\setminus K}|\!|\!|(h_{\mathrm{i}})_{\mathrm{i}_I}|\!|\!|_{I,J}^2\Big)^{p/2}$$

$$+\sum_{I\subseteq I_d}\sum_{J\subseteq I}p^{p(\#J/2+\#I^c)}\mathbb{E}_{I^c}\max_{\mathrm{i}_{I^c}}|\!|\!|(h_{\mathrm{i}})_{\mathrm{i}_I}|\!|\!|_{I,J}^p\Big].$$

Lemat 6 daje

$$\mathbb{E}_K\max_{\mathrm{i}_K}\Big(\mathbb{E}_{I^c\setminus K}\sum_{\mathrm{i}_{I^c\setminus K}}|\!|\!|(h_{\mathrm{i}})_{\mathrm{i}_I}|\!|\!|_{I,J}^2\Big)^{p/2} \leqslant \mathbb{E}_K\max_{\mathrm{i}_K}\Big(\sum_{\mathrm{i}_{K^c}}\mathbb{E}_{K^c}\|h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\|^2\Big)^{p/2}.$$

Pozostaje zauważyć, że $(p/2)^{\#Kp/2-\#I^cp/2} \leqslant 1$. $\square$

Uwaga Z Lematu 8 wynika, że w przestrzeniach kotypu 2 (a zatem w szczególności w przestrzeniach Hilberta), wielkości $\max_{I\subseteq I_d}\mathbb{E}_{I^c}\max_{\mathrm{i}_{I^c}}(\sum_{\mathrm{i}_I}\mathbb{E}_I\|h_{\mathrm{i}}\|^2)^{p/2}$ są (przynajmniej z dokładnością do stałych) niezbędne, gdyż

$$\mathbb{E}\|Z\|^p \geqslant \frac{1}{2^{pd}}\mathbb{E}\|\sum_{|\mathrm{i}|\leqslant n}\epsilon_{\mathrm{i}}^{\mathrm{dec}}h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\|^p \geqslant \frac{1}{K_d(B)^p}\mathbb{E}(\sum_{\mathrm{i}}\|h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\|^2)^{p/2}$$

$$\geqslant \frac{1}{K_d(B)^p}\max_{I\subseteq I_d}\mathbb{E}_{I^c}\max_{\mathrm{i}_{I^c}}(\sum_{\mathrm{i}_I}\mathbb{E}_I\|h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\|^2)^{p/2},$$

gdzie $K_d(B)$ zależy tylko od d i stałej w definicji kotypu 2 dla przestrzeni B.

Na zakończenie dyskusji oszacowań momentów w przestrzeniach typu 2, podamy wniosek z Twierdzenia 12, który jest prostszy w sformułowaniu, a może okazać się wystarczający do zastosowań. Aby go otrzymać, wystarczy zastąpić $|\!|\!|(h_i)_{\mathrm{i}_I}|\!|\!|_{I,J}$ dla $I \neq I_d$ oraz $\mathbb{E}\|Z\| = |\!|\!|(h_{\mathrm{i}})|\!|\!|_{I_d,\emptyset}$ przez oszacowania podane w Lemacie 6.

Wniosek 4. *Jeśli B jest typu 2, to dla $p \geqslant 2$*

$$\mathbb{E}\|Z\|^p \leqslant K_d(B)^p\Bigg[(\sum_{\mathrm{i}}\mathbb{E}\|h_{\mathrm{i}}\|^2)^{p/2} + \sum_{J\subseteq I_d, J\neq\emptyset}p^{p\#J/2}|\!|\!|(h_{\mathrm{i}})|\!|\!|_{I_d,J}^p$$

$$+\sum_{I\subsetneq I_d}p^{p(d+\#I^c)/2}\mathbb{E}_{I^c}\max_{\mathrm{i}_{I^c}}\Big(\sum_{\mathrm{i}_I}\mathbb{E}_I\|h_{\mathrm{i}}\|^2\Big)^{p/2}\Bigg],$$

gdzie stałe $K_d(B)$ zależą tylko od d i stałej z definicji typu 2 dla przestrzeni B.

2.3 U-procesy indeksowane przez VC klasy

W analogiczny sposób jak dla przestrzeni typu 2, możemy z Twierdzenia 11 uzyskać nierówności dla U-procesów, indeksowanych przez regularne klasy funkcji. Schemat dowodu jest analogiczny jak dla Twierdzenia 12, jedyną różnicą jest użycie zamiast Lematu 6, Lematu 5 w połączeniu z następującą nierównością.

LEMAT 10 ([12], rozdział 5). *Niech $\mathcal{H}$ będzie przeliczalną klasą kanonicznych funkcji $h\colon \Sigma^d \to \mathbb{R}$, punktowo ograniczonych przez funkcję H. Wówczas*

$$\Big\|\sup_{h\in\mathcal{H}}\Big|\sum_{|\mathrm{i}|\leqslant n} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big|\Big\|_2 \leqslant \Big\|\sup_{h\in\mathcal{H}}\Big|\sum_{|\mathrm{i}|\leqslant n} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big|\Big\|_{\psi_{2/d}} \leqslant K_d A\Big(\sum_{|\mathrm{i}|\leqslant n}\mathbb{E}H^2(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big)^{1/2},$$

gdzie

$$A = \int_0^\infty \Big[\sup_Q \log N(\mathcal{H}, \|\cdot\|_{L^2(Q)}, \varepsilon(QH^2)^{1/2})\Big]^{d/2} d\varepsilon,$$

przy czym supremum jest wzięte po wszystkich miarach probabilistycznych Q na Σ^d, takich że $QH^2 = \int H^2 dQ < \infty$, zaś $N(T, d, \varepsilon)$ oznacza entropię metryczną przestrzeni metrycznej (T, d) (zob. np. [12], rozdział 5.1.2)

Uwaga Powyższy lemat, podobnie jak oszacowania momentów, które przedstawimy, można rozszerzyć na nieprzeliczalne klasy $\mathcal{H}$ o ile spełniają one odpowiednie, standardowe w teorii procesów empirycznych, założenia dotyczące mierzalności.

Stała A jest skończona na przykład dla klas o skończonym wymiarze VC.

TWIERDZENIE 13. *Przy oznaczeniach Lematu 10, jeśli $A < \infty$, to dla pewnej stałej $K_d(A)$, zależnej tylko od d i A oraz $p \geqslant 2$ zachodzi*

$$\begin{aligned}\mathbb{E}\sup_{h\in\mathcal{H}}\Big|\sum_{\mathrm{i}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big|^p \quad &\leqslant \quad K_d(A)^p\Big[\max_{I\subseteq I_d}\mathbb{E}_{I^c}\max_{\mathrm{i}_{I^c}}(n^{\#I}\mathbb{E}_I H^2)^{p/2} \\ &\quad + \sum_{I\subseteq I_d}\sum_{J\subseteq I} p^{p(\#J/2+\#I^c)}\mathbb{E}_{I^c}\max_{\mathrm{i}_{I^c}}\|(h_{\mathrm{i}})_{\mathrm{i}_I}\|_{I,J}^p\Big],\end{aligned}$$

gdzie $(h_{\mathrm{i}})_{|\mathrm{i}|\leqslant n}$ jest macierzą o równych współczynnikach $h_{\mathrm{i}} = (h)_{h\in\mathcal{H}}$, $h_{\mathrm{i}}\colon \Sigma^d \to L^\infty(\mathcal{H})$.

Uwaga Przestrzeń $L^\infty(\mathcal{H})$, w przypadku $\mathcal{H}$ nieskończonego, nie jest ośrodkowa, ale oczywiście istnieje w niej przeliczalny zbiór funkcjonałów, jak w Definicji 6 czy Lemacie 4. Dowód jest zatem analogiczny jak dla Twierdzenia 12, z tą różnicą, że Twierdzenie 11 stosuje się do supremum po skończonych podzbiorach $\mathcal{H}$. Podobnie, aby zdefiniować $\|(h_\mathbf{i})_{\mathbf{i}_I}\|_{I,J}$ wystarczy jako funkcjonały w Definicji 7 brać rzutowania na elementy $\mathcal{H}$.

2.4 Kilka uwag o potencjalnych zastosowaniach

Przedstawione w tej części pracy oszacowania U-statystyk i U-procesów zawierają wartości oczekiwane supremów pewnych procesów empirycznych. Podobnie jak w przypadku oszacowań chaosów w ogólnej przestrzeni Banacha, zastąpienie tych wartości oczekiwanych przez wielkości, w których kolejność całkowania i supremum jest odwrócona, wydaje się być trudnym zagadnieniem, zależnym w istotny sposób od geometrii konkretnej przestrzeni Banacha. W dalszej części pracy przedstawimy rozwiązanie tego problemu dla prostej i, ogólniej, dla przestrzeni Hilberta. Niemniej, warto zaznaczyć, że oszacowania przedstawione do tej pory, mogą mieć potencjalne zastosowania. Dla przykładu, nierówności Giné i Masona, implikowane przez Twierdzenie 11, wspomniane w uwadze po dowodzie twierdzenia, zostały użyte do zbadania centralnego twierdzenia granicznego i praw iterowanego logarytmu dla lokalnych U-statystyk, w teorii estymacji gęstości [19, 20]. Oszacowania dla U-procesów rzędu 2, korzystające z supremów procesów empirycznych, zostały także niedawno użyte w [10] przy analizie problemu rankingu w statystycznej teorii uczenia się. Wreszcie, oszacowania dla przestrzeni typu 2 (oraz VC klas), pozwalają osłabić znane dotychczas warunki dostateczne na prawo iterowanego logarytmu dla kanonicznych U-statystyk w tych przestrzeniach (oraz dla odpowiednich U-procesów).

Rozdział 3

Nierówności dla U-statystyk rzeczywistych i o wartościach w przestrzeniach Hilberta

W poprzednim rozdziale przedstawione zostały odpowiedniki nierówności Arconesa i Giné dla U-statystyk o wartościach w przestrzeniach Banacha oraz ich ograniczenia, wynikające z użycia jako wielkości szacujących momenty wartości oczekiwanych supremów procesów empirycznych. Celem niniejszego rozdziału jest wyeliminowanie wspomnianych trudności, w szczególnym przypadku U-statystyk o wartościach w przestrzeniach Hilberta. Otrzymane rezultaty, w przypadku rzeczywistym można traktować jako odpowiedniki wyników Latały [35], dot. oszacowań momentów rzeczywistych chaosów gaussowskich (przedstawionych powyżej w Twierdzeniu 9). Choć nie jest to explicite wspomniane w pracy [35], metoda Latały pozwala także na znalezienie dwustronnych oszacowań momentów chaosów gaussowskich w przestrzeni Hilberta.

Dla $d = 1$, U-statystyki sprowadzają się do sum niezależnych zmiennych losowych o średniej zero. Oszacowania momentów i odchyleń takich sum (także w przypadku przestrzeni Banacha) są klasycznym zagadnieniem w rachunku prawdopodobieństwa, obejmującym wiele wyników, uzyskanych m. in. przez Bernsteina, Hoeffdinga, Hoffmana-Joergensena, Pinelisa i Talagranda. Wyniki zaprezentowane w niniejszej sekcji można uznać za uogólnienie na U-statystyki dowolnego rzędu klasycznej nierówności Bernsteina. W przypadku U-statystyk rzędu 2 zostały one uzyskane przez Giné, Latałę i Zinna [18], alternatywny dowód został podany także przez Houdré i Reynaud-Bouret [29]. Trudności z przeniesieniem tych wyników na U-statystyki dowolnego rzędu związane są z koniecznością oszacowania wartości oczekiwanej sum niezależnych zmiennych losowych w przestrzeni operatorów wieloliniowych. Podobne problemy dla średnich gaussowskich w przestrzeni macierzy zostały przezwyciężone przez Latałę w pracy [35]. Okazuje się, że warunkowe zastosowanie tego wyniku, w połączeniu z nierównościami pomiędzy słabą i silną wariancją procesu empirycznego, przedstawionymi przez Boucherona, Bousqueta, Lugosi i Massarta [8] (implicite zawartymi już w [18]), pozwala na oszacowanie wartości oczekiwanej sumy dowolnych zmiennych niezależnych o wartościach w przestrzeni operatorów.

Metody dowodu dla U-statystyk rzeczywistych i o wartościach w przestrzeni Hilberta są bardzo zbliżone, jednak przypadek ogólny wymaga dodatkowego skomplikowania notacji, która staje się dość złożona już dla prostej rzeczywistej. Z tego względu, a także aby lepiej zobrazować najważniejszy przypadek, najpierw przedstawimy wyniki dla U-statystyk rzeczywistych, ze szczególnym naciskiem na U-statystyki rzędu 3, dopiero następnie pokażemy, w jaki sposób uogólniają się one na dowolne przestrzenie Hilberta. Część technicznych lematów, których treść w przypadku rzeczywistym jest intuicyjnie oczywista, zostanie jedynie sformułowana, natomiast ich dowody przedstawimy podczas omawiania ich uogólnionych wersji dla przestrzeni Hilberta.

Rozważmy zatem $(h_{\mathrm{i}})_{\mathrm{i}\in(\mathbb{N}^*)^d,|\mathrm{i}|\leqslant n}$, gdzie $h_{\mathrm{i}}\colon \Sigma^d \to \mathbb{R}$ są kanoniczne ze względu na $(X_i^{(j)})_{i\in I_n,j\in I_d}$ oraz zmienną Z, zdefiniowaną, jak poprzednio, wzorem

$$Z = \sum_{|\mathrm{i}|\leqslant n} h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}).$$

Oszacowania momentów zmiennej Z będą się wyrażać w terminach norm wprowadzonych w poniższej definicji. Korzysta ona z notacji wprowadzonej w Definicji 4. Podobnie jak w Definicji 7, zakładamy, że $\mathbb{E}|h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})|^2 < \infty$ dla każdego i.

Definicja 9. *Dla niepustego zbioru $I \subseteq I_d$ rozważmy $\mathcal{J} = \{J_1,\ldots,J_k\} \in \mathcal{P}_I$. Dla dowolnej d-wskaźnikowej macierzy $(h_{\mathrm{i}})_{|\mathrm{i}|\leqslant n}$ funkcji rzeczywistych na Σ^d oraz dowolnej ustalonej wartości i_{I^c}, zdefiniujmy*

$$\|(h_{\mathrm{i}})_{\mathrm{i}_I}\|_{\mathcal{J}} = \sup\Big\{|\sum_{\mathrm{i}_I} \mathbb{E}_I h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}) \prod_{j=1}^{\deg(\mathcal{J})} f^{(j)}_{\mathrm{i}_{J_j}}(\mathbf{X}^{\mathrm{dec}}_{\mathrm{i}_{J_j}})|\colon f^{(j)}_{\mathrm{i}_{J_j}}\colon \Sigma^{\# J_j} \to \mathbb{R}$$
$$\mathbb{E}\sum_{\mathrm{i}_{J_j}} |f^{(j)}_{\mathrm{i}_{J_j}}(\mathbf{X}^{\mathrm{dec}}_{\mathrm{i}_{J_j}})|^2 \leqslant 1 \text{ dla } j = 1,\ldots,\deg(\mathcal{J})\Big\}.$$

Niech ponadto $\|(h_{\mathrm{i}})_{\mathrm{i}_\emptyset}\|_\emptyset = |h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})|$.

Przykład Dla $d = 2$ mamy (upraszczając notację i pisząc np. $\|\cdot\|_{\{1\}\{2\}}$ zamiast $\|\cdot\|_{\{\{1\},\{2\}\}}$)

$$\begin{aligned}
\|(h_{ij})_{ij}\|_{\{1,2\}} &= \sup\Big\{ \sum_{i,j\leqslant n} \mathbb{E}h_{ij}(X_i,Y_j)f_{i,j}(X_i,Y_j) \colon \mathbb{E}\sum_{i,j\leqslant n} f_{i,j}(X_i,Y_j)^2 \leqslant 1\Big\} \\
&= \sqrt{\sum_{i,j\leqslant n} \mathbb{E}h_{ij}(X_i,Y_j)^2}, \\
\|(h_{ij})_{ij}\|_{\{1\}\{2\}} &= \sup\Big\{ \sum_{i,j\leqslant n} \mathbb{E}h_{ij}(X_i,Y_j)f_i(X_i)g_j(Y_j) \colon \\
&\qquad \mathbb{E}\sum_{i\leqslant n} f_i(X_i)^2 \leqslant 1, \mathbb{E}\sum_{j\leqslant n} g_j(Y_j)^2 \leqslant 1\Big\} \\
&= \sup\Big\{ \sqrt{\sum_{i\leqslant n} \mathbb{E}_X(\sum_{j\leqslant n} \mathbb{E}_Y h_{ij}(X_i,Y_j)g_j(Y_j))^2} \colon \mathbb{E}\sum_{i\leqslant n} g_j(Y_j)^2 \leqslant 1\Big\} \\
&= \sup\Big\{ \sqrt{\sum_{j\leqslant n} \mathbb{E}_Y(\sum_{i\leqslant n} \mathbb{E}_X h_{ij}(X_i,Y_j)f_i(X_i))^2} \colon \mathbb{E}\sum_{j\leqslant n} f_i(X_i)^2 \leqslant 1\Big\}
\end{aligned}$$

oraz przy ustalonym i,

$$\begin{aligned}
\|(h(X_i,Y_j))_j\|_{\{2\}} &= \sup\Big\{ \mathbb{E}_Y \sum_{j\leqslant n} h(X_i,Y_j)f_j(Y_j) \colon \sum_{j\leqslant n} \mathbb{E}f_j(Y_j)^2 \leqslant 1\Big\} \\
&= \sqrt{\sum_{j\leqslant n} \mathbb{E}_Y h(X_i,Y_j)^2}.
\end{aligned}$$

Uwagi Powyższa definicja jest pod pewnym względem zbliżona do Definicji 7. W szczególności, dla $I = I_d$, $\|(h_{\mathrm{i}})_{\mathrm{i}_I}\|_{\mathcal{J}}$ jest „deterministyczną” normą, zaś dla $I \neq I_d$ zmienną losową zależną od $(X_{i_j}^{(j)})_{j\in I^c}$, (jak łatwo zauważyć, jest to analogiczna norma, obliczona warunkowo dla podmacierzy niższego wymiaru). Dla uproszczenia notacji, zależności od $(X_{i_j}^{(j)})_{j\in I^c}$, podobnie jak w przypadku Definicji 7, nie będziemy zaznaczać.

Oczywiście istotną różnicą między Definicjami 7 i 9 jest fakt, że w tej drugiej nie występuje całkowanie supremów zmiennych losowych.

Gdy $\Sigma = \mathbb{R}$ oraz $h_{\mathrm{i}}(x_1,\ldots,x_d) = a_{\mathrm{i}}\prod_{k=1}^d x_k$, dla pewnych $a_{\mathrm{i}} \in \mathbb{R}$ (tzn. gdy U-statystyka jest uniezależnionym chaosem) oraz $(X_i^{(j)})$ to niezależne zmienne o rozkładzie $\mathcal{N}(0,1)$, jak łatwo sprawdzić, powyższa definicja sprowadza się do Definicji 5, dokładniej $\|(h_{\mathrm{i}})\|_{\mathcal{J}} = \|(a_{\mathrm{i}})\|_{\mathcal{J}}$. Normy tego typu nie są nowe, występują one w analizie funkcjonalnej w kontekście iniektywnych iloczynów tensorowych. Jak łatwo zauważyć, są to po prostu normy macierzy (macierzy funkcji) interpretowanej jako wieloliniowy funkcjonał na odpowiednich sumach prostych iloczynów tensorowych przestrzeni Hilberta.

Zauważmy ponadto, że gdy $\mathcal{J}$ jest rozbiciem zbioru $I \subseteq I_d$ na singletony, mamy $\|(h_{\mathrm{i}})_{\mathrm{i}_I}\|_{\mathcal{J}} = |\!|\!|(h_{\mathrm{i}})_{\mathrm{i}_I}|\!|\!|_{I,I}$. Z kolei, dla rozbicia trywialnego $\mathcal{J} = \{I\}$ oraz dowolnego $J \subseteq I_d$, zachodzi

$$\|(h_{\mathrm{i}})_{\mathrm{i}_I}\|_{\mathcal{J}} = \sqrt{\sum_{\mathrm{i}_I} \mathbb{E}_I h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})^2} \geqslant |\!|\!|(h_{\mathrm{i}})_{\mathrm{i}_I}|\!|\!|_{I,J}.$$

Zatem, stosując do prostej Twierdzenie 12, możemy rzeczywiście otrzymać oszacowania momentów rzeczywistych U-statystyk za pomocą wielkości wprowadzonych w Definicji 9. Dla $d = 1$, uzyskana nierówność jest z dokładnością do stałych równoważna nierówności Bernsteina

$$\mathbb{E}|\sum_i Z_i|^p \leqslant K^p \Big[p^{p/2}(\sum_i \mathbb{E}Z_i^2)^{p/2} + p^p \mathbb{E}\max_i |Z_i|^p\Big], \tag{3.1}$$

gdzie Z_i są niezależnymi zmiennymi losowymi o średniej 0 (dokładniej, nierówność ta jest równoważna oszacowaniom Bernsteina na odchylenia sum niezależnych zmiennych losowych, równoważność ta jest wykazana np. w [18]). Dla $d = 2$, uzyskujemy nierówność Giné, Latały, Zinna [18]

$$\begin{aligned} E\Big|\sum_{i,j} h_{i,j}(X_i^{(1)}, X_j^{(2)})\Big|^p \leqslant K^p \Big[& p^{p/2}\Big(\sum_{i,j} \mathbb{E}h_{i,j}(X_i^{(1)}, X_j^{(2)})^2\Big)^{p/2} + p^p \|(h_{i,j})_{i,j}\|^p_{L^2 \to L^2} \\ & + p^{3p/2}\mathbb{E}_1 \max_i \Big(\sum_j \mathbb{E}_2 h_{i,j}(X_i^{(1)}, X_j^{(2)})^2\Big)^{p/2} \\ & + p^{3p/2}\mathbb{E}_2 \max_j \Big(\sum_i \mathbb{E}_1 h_{i,j}(X_i^{(1)}, X_j^{(2)})^2\Big)^{p/2} \\ & + p^{2p}\mathbb{E} \max_{i,j} |h_{i,j}(X_i^{(1)}, X_j^{(2)})|^p\Big]. \end{aligned} \tag{3.2}$$

Dla $d \geqslant 3$ uzyskane nierówności, choć silniejsze niż znane dotychczas oszacowania momentów rzeczywistych U-statystyk, nie są niestety satysfakcjonujące, ze względu na nieoptymalną zależność od p. Zauważmy, że już dla $d = 3$, zastąpienie $|\!|\!| \cdot |\!|\!|^2_{\{1,2,3\},\{1,2\}}$ przez wariancję U-statystyki sprawia, że współczynnik przy wariancji wynosi $p^{\#\{1,2\}p/2} = p^p$. W konsekwencji, otrzymane oszacowania nie odzwierciedlają gaussowskiego zachowania momentów i odchyleń U-statystyk dla małych wartości odp. p i t. Takie zachowanie sugerowane jest zarówno przez przedstawione powyżej nierówności dla $d = 1, 2$, jak i oszacowania momentów chaosów gaussowskich dane w Twierdzeniu 9.

W dalszej części pracy potrzebna nam będzie jeszcze jedna definicja norm $\|\cdot\|_{\mathcal{J}}$.

Definicja 10. *Dla* $\mathrm{i} \in (\mathbb{N}^*)^{d-1} \times I_n$ *niech* $a_{\mathrm{i}} \colon \Sigma \to \mathbb{R}$ *będzie funkcją mierzalną. Dla podziału* $\mathcal{J} = \{J_1, \ldots, J_k\} \in \mathcal{P}_{I_d}$ *(załóżmy dla ustalenia uwagi, że* $d \in J_1$*), zdefiniujmy*

$$\|(a_{\mathrm{i}}(X_{i_d}^{(d)}))_{\mathrm{i}}\|_{\mathcal{J}} = \sup\Bigg\{ \sqrt{\sum_{\mathrm{i}_{J_1}} \mathbb{E}\Big(\sum_{\mathrm{i}_{I_d \setminus J_1}} a_{\mathrm{i}}(X_{i_d}^{(d)}) x_{\mathrm{i}_{J_2}}^{(1)} \cdots x_{\mathrm{i}_{J_k}}^{(k)} \Big)^2} : \\ \sum_{\mathrm{i}_{J_1}} (x_{\mathrm{i}_{J_2}}^{(1)})^2 \leqslant 1, \ldots, \sum_{\mathrm{i}_{J_k}} (x_{\mathrm{i}_{J_k}}^{(k)})^2 \leqslant 1 \Bigg\}.$$

Uwaga Zauważmy, że podobnie jak w Definicjach 5 oraz 9, $\|(a_{\mathrm{i}}(X_{i_d}^{(d)}))_{\mathrm{i}}\|_{\mathcal{J}}$ jest normą macierzy $(a_{\mathrm{i}}(X_{i_d}^{(d)}))_{\mathrm{i}}$ traktowanej jako k-liniowy operator na odpowiedniej sumie prostej iloczynów tensorowych (w tym wypadku na $\oplus_{j=1}^{\deg \mathcal{J}} (\otimes_{i \in J_j} H_i)$, gdzie $H_d = \oplus_{i_d=1}^n L^2(X_{i_d}^{(d)})$, zaś dla $i \neq d$, $H_i = \ell_2$). Ponadto normy te są dobrze zdefiniowane także dla skończonych macierzy $(a_{\mathrm{i}}(X_{i_d}^{(d)}))_{\mathrm{i}}$, gdyż możemy rozszerzyć je zerami do macierzy nieskończonych. Interpretacja w języku iloczynów tensorowych uzasadnia stosowanie takich samych oznaczeń, jak w Definicjach 5 i 9, niemniej w dalszej części rozprawy, będziemy do macierzy $(a_{\mathrm{i}}(X_{i_d}^{(d)}))_{\mathrm{i}}$ stosowali również normy z Definicji 5, warunkowo, przy ustalonych zmiennych $X_{i_d}^{(d)}$ (a zatem traktując $(a_{\mathrm{i}}(X_{i_d}^{(d)}))_{\mathrm{i}}$ jako d-wskaźnikową macierz o współczynnikach liczbowych). Aby uniknąć niejednoznaczności, otrzymaną w tę sposób liczbę (a właściwie zmienną losową, zależną od $(X_i^{(d)})_i$) będziemy oznaczać przez $\|(a_{\mathrm{i}}(X_{i_d}^{(d)}))_{\mathrm{i}}\|_{\mathcal{J},D}$ (litera D ma oznaczać, że stosujemy normę $\|\cdot\|_{\mathcal{J}}$ macierzy „deterministycznej”).

3.1 Rzeczywiste U-statystyki rzędu 3

Aby przeanalizować zachowanie momentów dla U-statystyk rzędu 3, uprośćmy nieco notację i zapiszmy

$$Z := \sum_{ijk=1}^{n} h_{ijk}(X_i, Y_j, Z_k),$$

przyjmując, że we wszystkich poprzednich definicjach Y_j i Z_k odpowiadają zmiennym $X_j^{(2)}, X_k^{(3)}$.

Analizując tezę Twierdzenia 12 dla $d = 3$ oraz korzystając z przedstawionych powyżej uwag, dotyczących związku między normami $\|\cdot\|_{\mathcal{J}}$ oraz $|\!|\!|\cdot|\!|\!|_{I,J}$, można zauważyć, że aby uzyskać oszacowania momentów z zależnością od p przy $\|\cdot\|_{\mathcal{J}}$ dla $\mathcal{J} \in \mathcal{P}_{\{1,2,3\}}$, odpowiadającą Twierdzeniu 9, wystarczy oszacować

$$\|\!|(h_{ijk})_{ijk}\|\!|_{\{1,2,3\},\{1,2\}} = \mathbb{E}_Z \sup\{ \sum_k \sum_{i,j} \mathbb{E}_{X,Y} h_{ijk}(X_i, Y_j, Z_k) f_i(X_i) g_j(Y_j) \colon$$
$$\sum_i \mathbb{E} f_i(X_i)^2, \sum_j \mathbb{E} g_j(Y_j)^2 \leqslant 1\}$$

oraz symetryczne wyrażenia $\|\!|(h_{ijk})_{ijk}\|\!|_{\{1,2,3\},\{1,3\}}$ i $\|\!|(h_{ijk})_{ijk}\|\!|_{\{1,2,3\},\{2,3\}}$.

Rozważmy $\mathbf{e}_m = (\mathbf{e}_m^1(X_1), \ldots, \mathbf{e}_m^n(X_n))$, $\mathbf{f}_m = (\mathbf{f}_m^1(Y_1), \ldots, \mathbf{f}_m^n(Y_n))$ – bazy ortonormalne odpowiednio w $L^2(X_1) \oplus \ldots \oplus L^2(X_n)$ oraz $L^2(Y_1) \oplus \ldots \oplus L^2(Y_n)$. Oznaczmy $a_{ijk}(Z_k) := \sum_{lm} \mathbb{E}_{XY} h_{lmk}(X_l, Y_m, Z_k) \mathbf{e}_i^l(X_l) \mathbf{f}_j^m(Y_m)$. W przypadku, gdy któraś z powyższych przestrzeni jest skończenie-wymiarowa, zakres indeksu i lub j może być tylko skończony, ale oczywiście możemy uzupełnić macierz $a_{ijk}(Z_k)_{ij}$ do macierzy nieskończonej kładąc na dalszych miejscach zera. Ten zabieg kosmetyczny pozwoli nam interpretować sumy po i, j zawsze jako sumy nieskończone.

Przy powyższych oznaczeniach, mamy

$$\begin{aligned} \|\!|(h_{ijk})_{ijk}\|\!|_{\{1,2,3\},\{1,2\}} &= \mathbb{E} \sup\{\sum_k \sum_{ij} a_{ijk}(Z_k) x_i y_j \colon \|x\|_2, \|y\|_2 \leqslant 1\} \\ &= \mathbb{E}\| \sum_k (a_{ijk}(Z_k))_{ij} \|_{\ell_2 \to \ell_2} = \mathbb{E}\| \sum_k (a_{ijk}(Z_k))_{ij} \|_{\{1\}\{2\}} \end{aligned}$$

gdzie $\sum_{ijk} \mathbb{E} a_{ijk}(Z_k)^2 < \infty$, przy czym wykorzystane tu zostały standardowe własności przestrzeni L^2 oraz ich iloczynów tensorowych (w szczególności np. równość $(\oplus_i L^2(X_i)) \otimes (\oplus_j L^2(Y_j)) \simeq \oplus_{i,j} L^2(X_i, Y_j)$).

Problem sprowadza się więc do oszacowania wartości oczekiwanej normy operatorowej sumy niezależnych macierzy losowych.

Aby kontynuować, zauważmy, że ponieważ normy $\|\cdot\|_{\mathcal{J}}$ zarówno dla macierzy funkcji, jak i macierzy losowych typu $(a_{ijk}(Z_k))_{ijk}$ mają interpretację jako normy funkcjonałów wieloliniowych na odpowiednich iloczynach tensorowych przestrzeni Hilberta, zachodzi następujący

Lemat 11. *Dla dowolnego rozbicia $\mathcal{J} \in \mathcal{P}_{\{1,2,3\}}$,*

$$\|(h_{ijk})_{ijk}\|_{\mathcal{J}} = \|(a_{ijk}(Z_k))_{ijk}\|_{\mathcal{J}}.$$

Aby oszacować $\mathbb{E}\| \sum_k (a_{ijk}(Z_k))_{ij} \|_{l_2 \to l_2}$, użyjemy warunkowo następującego wyniku Latały, kluczowego w dowodzie oszacowań momentów chaosów gaussowskich.

Lemat 12 ([35], Theorem 2). *Rozważmy 3-wskaźnikową macierz $A = (a_{ijk})$.*

Wówczas dla wszystkich $p \geqslant 2$,

$$\begin{aligned}\mathbb{E}\Big\|\Big(\sum_k a_{ijk}g_k\Big)_{ij}\Big\|_{\ell_2\to\ell_2} =&\mathbb{E}\Big\|(\sum_k a_{ijk}g_k)_{ij}\Big\|_{\{1\}\{2\}}\\ \leqslant& K\Big(\|A\|_{\{1\}\{2,3\}}+\|A\|_{\{2\},\{1,3\}}\\ &+\frac{1}{\sqrt{p}}\|A\|_{\{1,2,3\}}+\sqrt{p}\|A\|_{\{1\}\{2\}\{3\}}\Big).\end{aligned}$$

Istotny będzie też następujący fakt, wyrażający związki między silną i słabą wariancją procesu empirycznego. W nieco słabszej wersji został on udowodniony przez Talagranda, poniższe sformułowanie pochodzi od Boucherona, Bousqueta, Lugosi i Massarta [8], zaś dowód jest implicite zawarty także w pracy [18].

Lemat 13 ([8, Lemma 7]). *Niech $X_1,\ldots,X_n$ będą niezależnymi zmiennymi losowymi, o wartościach w $(\Sigma,\mathcal{B})$ zaś $\mathcal{T}$ – przeliczalną klasą funkcji mierzalnych $f\colon\Sigma\to\mathbb{R}$, takich że dla każdego i, $\mathbb{E}f(X_i)=0$. Wówczas*

$$\mathbb{E}\sup_{f\in\mathcal{T}}\sum_i f^2(X_i)\leqslant\sup_{f\in\mathcal{T}}\sum_i\mathbb{E}f(X_i)^2+32\sqrt{\mathbb{E}M^2}\mathbb{E}\sup_{f\in\mathcal{T}}|\sum_i f(X_i)|+8\mathbb{E}M^2,$$

gdzie $M:=\max_i\sup_{f\in\mathcal{T}}|f(X_i)|$.

Możemy teraz wrócić do oszacowań $\mathbb{E}\|(\sum_k a_{ijk}(Z_k))_{ij}\|_{\ell_2\to\ell_2}$. Rozważmy $\varepsilon_1,\ldots,\varepsilon_n$ - ciąg zmiennych Rademachera, niezależnych od $(X_i)_i,(Y_j)_j,(Z_k)_k$. Używając standardowych technik symetryzacyjnych ([39], Lemma 6.3), faktu, że średnie rademacherowe są zdominowane przez gaussowskie oraz Lematu 12 warunkowo na $(Z_k)_k$, otrzymujemy

$$\begin{aligned}\mathbb{E}\|(\sum_k a_{ijk}(Z_k))_{ij}\|_{\ell_2\to\ell_2}\leqslant&2\mathbb{E}\|(\sum_k a_{ijk}(Z_k)\varepsilon_k)_{ij}\|\leqslant 2\sqrt{\frac{\pi}{2}}\mathbb{E}\|(\sum_k a_{ijk}(Z_k)g_k)_{ij}\|\\ \leqslant& K\Big(\mathbb{E}\|(a_{ijk}(Z_k))\|_{\{1\}\{2,3\},D}+\mathbb{E}\|(a_{ijk}(Z_k))\|_{\{2\},\{1,3\},D}\\ &+\frac{1}{\sqrt{p}}\mathbb{E}\|(a_{ijk}(Z_k))\|_{\{1,2,3\},D}\\ &+\sqrt{p}\mathbb{E}\|(a_{ijk}(Z_k))\|_{\{1\}\{2\}\{3\},D}\Big),\end{aligned}\tag{3.3}$$

gdzie zgodnie z uwagą po Definicji 10, $\|(a_{ijk}(Z_k))\|_{\{1\}\{2\}\{3\},D}$ oznacza, normę z Definicji 5 zastosowaną do macierzy liczbowej przy ustalonych zmiennych Z_k.

Oczywiście

$$\frac{1}{\sqrt{p}}\mathbb{E}\|(a_{ijk}(Z_k))\|_{\{1,2,3\},D}\leqslant\frac{1}{\sqrt{p}}\sqrt{\mathbb{E}\sum_{ijk}a_{ijk}(Z_k)^2}=\frac{1}{\sqrt{p}}\|(a_{ijk}(Z_k))\|_{\{1,2,3\}}.\tag{3.4}$$

Zajmiemy się teraz oszacowaniem pozostałych składników po prawej stronie (3.3). Zacznijmy od $\mathbb{E}\|(a_{ijk}(Z_k))\|_{\{1\}\{2\}\{3\},D}$. Niech

$$M^2 = \max_k \sup_{\|x\|_2,\|y\|_2\leqslant 1} |\sum_{ij} a_{ijk}(Z_k)x_iy_j|^2 = \max_k \|(a_{ijk}(Z_k))_{ij}\|^2_{\ell_2\to\ell_2}.$$

Z Lematu 13 otrzymujemy

$$\begin{aligned}\mathbb{E}\|(a_{ijk}(Z_k))\|^2_{\{1\}\{2\}\{3\},D} =&\mathbb{E} \sup_{\|x\|_2,\|y\|_2\leqslant 1} \sum_k (\sum_{ij} a_{ijk}(Z_k)x_iy_j)^2 \\ \leqslant& \sup_{\|x\|_2,\|y\|_2\leqslant 1} \mathbb{E}\sum_k (\sum_{ij} a_{ijk}(Z_k)x_iy_j)^2 \\ &+ 32\sqrt{\mathbb{E}M^2}\mathbb{E}\|(\sum_k a_{ijk}(Z_k))_{ij}\|_{\ell_2\to\ell_2} + 8\mathbb{E}M^2 \\ =&\|(a_{ijk}(Z_k))\|^2_{\{1\}\{2\}\{3\}} + 32\sqrt{\mathbb{E}M^2}\mathbb{E}\|(\sum_k a_{ijk}(Z_k))_{ij}\|_{\ell_2\to\ell_2} \\ &+ 8\mathbb{E}M^2.\end{aligned}$$

Używając nierówności $\sqrt{ab} \leqslant \sqrt{p}a/\varepsilon + b\varepsilon/\sqrt{p}$, otrzymujemy ostatecznie dla $0 < \varepsilon < 1$, $p \geqslant 2$,

$$\begin{aligned}\mathbb{E}\|(a_{ijk}(Z_k))\|_{\{1\}\{2\}\{3\},D} \leqslant& K\Bigg(\|(a_{ijk}(Z_k))\|_{\{1\}\{2\}\{3\}} + \frac{\varepsilon}{\sqrt{p}}\mathbb{E}\|(\sum_k a_{ijk}(Z_k))_{ij}\|_{\ell_2\to\ell_2} \\ &+ \frac{\sqrt{p}}{\varepsilon}\sqrt{\mathbb{E}\max_k \|(a_{ijk}(Z_k))_{ij}\|^2}\Bigg). \qquad (3.5)\end{aligned}$$

Przejdźmy do oszacowań dla $\mathbb{E}\|(a_{ijk}(Z_k))\|_{\{2\}\{1,3\},D}$. Mamy

$$\begin{aligned}\mathbb{E}\|(a_{ijk}(Z_k))\|^2_{\{2\}\{1,3\},D} &= \mathbb{E}\sup_{\|y\|_2\leqslant 1} \sum_{i,k}(\sum_j a_{ijk}(Z_k)y_j)^2 \\ &= \mathbb{E}\sup_{\|y\|_2\leqslant 1} \sum_k \left(\varepsilon_k \sqrt{\sum_i(\sum_j a_{ijk}(Z_k)y_j)^2}\right)^2.\end{aligned}$$

Niech

$$S = \sup_{\|y\|_2\leqslant 1} \left|\sum_k \varepsilon_k \sqrt{\sum_i (\sum_j a_{ijk}(Z_k)y_j)^2}\right|.$$

Używając ponownie Lematu 13, tym razem do zmiennych $X_k = (Z_k, \varepsilon_k, k)$ oraz funkcji $f_y(Z_k, \varepsilon_k, k) = \varepsilon_k\sqrt{\sum_i(\sum_j a_{ijk}(Z_k)y_j)^2}$, otrzymujemy

$$\begin{aligned}\mathbb{E}\|(a_{ijk}(Z_k))\|^2_{\{2\}\{1,3\},D} =&\mathbb{E}\sup_{\|y\|_2\leqslant 1}\sum_k f_y(Z_k,\varepsilon_k,k)^2\\ \leqslant&\sup_{\|y\|_2\leqslant 1}\mathbb{E}\sum_{i,k}(\sum_j a_{ijk}(Z_k)y_j)^2+32\sqrt{\mathbb{E}M^2}\mathbb{E}S+8\mathbb{E}M^2\\ =&\|(a_{ijk}(Z_k))_{ijk}\|^2_{\{1,3\}\{2\}}+32\sqrt{\mathbb{E}M^2}\mathbb{E}S+8\mathbb{E}M^2. \qquad (3.6)\end{aligned}$$

Aby zakończyć, należy zatem oszacować $\mathbb{E}S$. Rozważmy w tym celu $g_1,\ldots,g_n$ – niezależne zmienne losowe o rozkładzie $\mathcal{N}(0,1)$, niezależne od zmiennych Z_k. Mamy, ponownie korzystając z dominacji średnich gaussowskich nad rademacherowymi,

$$\begin{aligned}\mathbb{E}S =&\mathbb{E}\sup_{\|y\|_2\leqslant 1}|\sum_k\varepsilon_k\sqrt{\sum_i(\sum_j a_{ijk}(Z_k)y_j)^2}|\\ \leqslant&\sqrt{\frac{\pi}{2}}\mathbb{E}\sup_{\|y\|_2\leqslant 1}|\sum_k g_k\sqrt{\sum_i(\sum_j a_{ijk}(Z_k)y_j)^2}|\\ =&\sqrt{\frac{\pi}{2}}\mathbb{E}\sup_{\|y\|_2\leqslant 1}|X_y|\leqslant\sqrt{\frac{\pi}{2}}\mathbb{E}(\sup_{\|y\|_2\leqslant 1}X_y^+ +\sup_{\|y\|_2\leqslant 1}X_y^-)\\ =&2\sqrt{\frac{\pi}{2}}\mathbb{E}\sup_{\|y\|_2\leqslant 1}X_y^+ = 2\sqrt{\frac{\pi}{2}}\mathbb{E}\sup_{\|y\|_2\leqslant 1}X_y, \qquad (3.7)\end{aligned}$$

gdzie $X_y=\sum_k g_k\sqrt{\sum_i(\sum_j a_{ijk}(Z_k)y_j)^2}$ jest procesem (warunkowo) gaussowskim indeksowanym przez kulę jednostkową w ℓ_2 (przy czym w przedostatniej równości użyliśmy faktu, że procesy $(X_y)_y$ i $(-X_y)_y$ mają te same rozkłady, zaś w ostatniej, że $X_0=0$).

Dla dalszych rozważań, załóżmy chwilowo, że dla pewnej liczby naturalnej N, z prawdopodobieństwem 1, liczby $a_{ijk}(Z_k)$ są równe 0, gdy $\max(i,j,k)>N$.

Struktura kowariancji procesu X indukuje metrykę na zbiorze indeksującym, daną wzorem

$$\begin{aligned}d_X(y,\tilde{y})^2 &= \mathbb{E}_g|X_y-X_{\tilde{y}}|^2\\ &= \sum_k\left(\sqrt{\sum_i(\sum_j a_{ijk}(Z_k)y_j)^2}-\sqrt{\sum_i(\sum_j a_{ijk}(Z_k)\tilde{y}_j)^2}\right)^2\\ &= \sum_k(\|(\sum_j a_{ijk}(Z_k)y_j)_i\|_2-\|(\sum_j a_{ijk}(Z_k)\tilde{y}_j)_i\|_2)^2\\ &\leqslant \sum_k\|(\sum_j a_{ijk}(Z_k)(y_j-\tilde{y}_j))_i\|_2^2\\ &= \sum_{ik}(\sum_j a_{ijk}(Z_k)(y_j-\tilde{y}_j))^2=d_{\tilde{X}}(y,\tilde{y})^2,\end{aligned}$$

gdzie $\tilde{X}_y = \sum_{ik} g_{ik} \sum_j a_{ijk}(Z_k) y_j$ jest również procesem (warunkowo) gaussowskim (g_{ik} są zmiennymi i.i.d. o rozkładzie $\mathcal{N}(0,1)$, niezależnymi od zmiennych Z_k).

Założenie o skończoności nośnika macierzy $(a_{ijk}(Z_k))_{ijk}$ implikuje, że wszystkie sumy w definicjach X_y i $\tilde{X}_y$ są skończone. W szczególności oba procesy są dobrze określone i mają ciągłe trajektorie ze względu na y (założenie to nie jest potrzebne, z warunków na zbieżność szeregów gaussowskich w przestrzeni Hilberta można bowiem łatwo wykazać, że oba procesy wyznaczają miary gaussowskie na odpowiednich przestrzeniach ℓ_2, tym niemniej upraszcza ono szczegóły techniczne dowodu, w szczególności pozwala „automatycznie" zmieniać kolejność sumowania).

Ciągłość trajektorii pozwala nam zastosować Lemat Slepiana (np. [39], Corollary 3.12) , który daje

$$\begin{aligned}\mathbb{E} \sup_{\|y\|_2 \leqslant 1} X_y \leqslant & \mathbb{E} \sup_{\|y\|_2 \leqslant 1} \tilde{X}_y = \mathbb{E} \sqrt{\sum_j \Big(\sum_{ik} a_{ijk}(Z_k) g_{ik}\Big)^2} \\ & \leqslant \sqrt{\sum_{ijk} \mathbb{E} a_{ijk}(Z_k)^2} = \|(a_{ijk}(Z_k))_{ijk}\|_{\{1,2,3\}}.\end{aligned}$$

Używając powyższej nierówności w połączeniu z (3.6) i (3.7) oraz korzystając z faktu $\sqrt{ab} \leqslant \sqrt{p} a + b/\sqrt{p}$, otrzymujemy dla $p \geqslant 2$,

$$\begin{aligned}\mathbb{E}\|(a_{ijk}(Z_k))\|_{\{2\}\{1,3\},D} \leqslant & \|(a_{ijk}(Z_k))_{ijk}\|_{\{1,3\}\{2\}} \\ & + \frac{K}{\sqrt{p}} \|(a_{ijk}(Z_k))_{ijk}\|_{\{1,2,3\}} + K\sqrt{p} \sqrt{\mathbb{E} \max_k \|(a_{ijk}(Z_k))_{ij}\|_{\ell_2 \to \ell_2}^2},\end{aligned} \tag{3.8}$$

przy dodatkowym założeniu o skończoności nośników macierzy. Łatwo widać, że powyższa nierówność rozszerza się na ogólny przypadek macierzy, o niekoniecznie ograniczonych nośnikach, takich że $\mathbb{E} \sum_{ijk} a_{ijk}(Z_k)^2 < \infty$ (wynika to w standardowy sposób z dominacji norm $\|\cdot\|_{\mathcal{J}}$ przez normę $\|\cdot\|_{\{1,2,3\}}$ oraz twierdzenia Lebesgue'a o zbieżności zmajoryzowanej).

Z symetrii

$$\begin{aligned}\mathbb{E}\|(a_{ijk}(Z_k))\|_{\{1\}\{2,3\},D} \leqslant & \|(a_{ijk}(Z_k))\|_{\{1\}\{2,3\}} \\ & + \frac{K}{\sqrt{p}} \|(a_{ijk}(Z_k))\|_{\{1,2,3\}} + K\sqrt{p} \sqrt{\mathbb{E} \max_k \|(a_{ijk}(Z_k))_{ij}\|_{\ell_2 \to \ell_2}^2}.\end{aligned} \tag{3.9}$$

Nierówności (3.3)–(3.5) (z ε dostatecznie małym) oraz (3.8),(3.9) razem z Lematem 11 implikują następujące twierdzenie, dotyczące wartości oczekiwanej normy sumy niezależnych macierzy losowych.

Twierdzenie 14. *Niech $Z_1, \ldots, Z_n$ będą niezależnymi zmiennymi losowymi o wartościach w $(\Sigma, \mathcal{F})$. Dla $i, j \in \mathbb{N}^*$, $k \in I_n$, niech $a_{ijk} \colon \Sigma \to \mathbb{R}$ będą funkcjami* *mierzalnymi takimi, że $\mathbb{E}_Z a_{ijk}(Z_k) = 0$. Wówczas, dla dowolnego $p \geqslant 2$,*

$$\begin{aligned}\mathbb{E}\|(\sum_k a_{ijk}(Z_k))_{ij}\|_{l_2\to l_2} \leqslant K\Big[&\frac{1}{\sqrt{p}}\|(a_{ijk}(Z_k))\|_{\{1,2,3\}} \\ &+ \|(a_{ijk}(Z_k))\|_{\{1\}\{2,3\}} + \|(a_{ijk}(Z_k))\|_{\{2\}\{1,3\}} \\ &+ \sqrt{p}\|(a_{ijk}(Z_k))\|_{\{1\}\{2\}\{3\}} \\ &+ p\sqrt{\mathbb{E}\max_k \|(a_{ijk}(Z_k))_{ij}\|^2_{\{1\}\{2\}}}\Big].\end{aligned}$$

Otrzymujemy stąd nierówność, pozwalającą na oszacowanie momentów U-statystyk.

WNIOSEK 5. *Dla dowolnego $p \geqslant 2$,*

$$\begin{aligned}\|(h_{ijk})\|_{\{1,2,3\}\{1,2\}} \leqslant K\Big[&\frac{1}{\sqrt{p}}\|(h_{ijk})\|_{\{1,2,3\}} + \|(h_{ijk})\|_{\{1\}\{2,3\}} + \\ &+ \|(h_{ijk})\|_{\{2\}\{1,3\}} + \sqrt{p}\|(h_{ijk})\|_{\{1\}\{2\}\{3\}} \\ &+ p\sqrt{\mathbb{E}_Z \max_k \|(h_{ijk})_{ij}\|^2_{\{1\}\{2\}}}\Big].\end{aligned}$$

Uwaga Porównując powyższe Twierdzenie z Lematem 12, możemy zauważyć, że pojawił się dodatkowy składnik $p\sqrt{\mathbb{E}_Z \max_k \|(a_{ijk}(Z_k))_{ij}\|^2_{\{1\}\{2\}}}$. Jego obecność jest do pewnego stopnia naturalna, gdyż z reguły w rozszerzeniach nierówności dla sum niezależnych zmiennych gaussowskich na ogólne niezależne zmienne losowe, pojawia się dodatkowy składnik, związany właśnie z maximum tych zmiennych. Jako przykład może tu służyć porównanie wzrostu momentów dla zmiennych gaussowskich z nierównością Bernsteina.

Z powyższego rezultatu oraz Twierdzenia 12 otrzymujemy następujące oszacowania momentów U-statystyk rzędu 3.

TWIERDZENIE 15. *Dla dowolnego $p \geqslant 2$,*

$$\mathbb{E}\Big|\sum_{ijk} h_{ijk}(X_i, Y_j, Z_k)\Big|^p \leqslant K^p\Big[\sum_{I\subseteq\{1,2,3\}}\sum_{\mathcal{J}\in\mathcal{P}_I} p^{p(\deg(\mathcal{J})/2+\#I^c)}\mathbb{E}_{I^c}\max_{\mathrm{i}_{I^c}} \|(h_{ijk})_{\mathrm{i}_I}\|^p_{\mathcal{J}}\Big].$$

3.2 Rzeczywiste U-statystyki dowolnego rzędu

Analiza nierówności Bernsteina (3.1) oraz Giné, Latały i Zinna (3.2) pokazuje, że mogą być one zapisane w takiej samej formie jak nierówność z Twierdzenia 15, poprzez zastąpienie zbioru $\{1, 2, 3\}$ odpowiednio przez $\{1\}$ i $\{1, 2\}$. Sugeruje to postać oszacowań momentów dla U-statystyk dowolnego rzędu d. Rzeczywiście, okazuje się, że zachodzi następujące twierdzenie.

TWIERDZENIE 16. *Istnieją stałe K_d takie, że dla $p \geqslant 2$,*

$$\mathbb{E}\Big|\sum_{|\mathrm{i}|\leqslant n} h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big|^p \leqslant K_d^p\Big[\sum_{I\subseteq I_d}\sum_{\mathcal{J}\in\mathcal{P}_I} p^{p(\#I^c+\deg(\mathcal{J})/2)}\mathbb{E}_{I^c}\max_{\mathrm{i}_{I^c}}\|(h_{\mathrm{i}})_{\mathrm{i}_I}\|_{\mathcal{J}}^p\Big].$$

Aby udowodnić powyższy wynik, będziemy potrzebować odpowiedniej wersji Twierdzenia 14.

TWIERDZENIE 17. *Niech $Z_1,\ldots,Z_n$ będą niezależnymi zmiennymi losowymi o wartościach w $(\Sigma,\mathcal{F})$. Dla $\mathrm{i}\in\mathbb{N}^{d-1}\times I_n$ niech $a_{\mathrm{i}}\colon\Sigma\to\mathbb{R}$ będą funkcjami mierzalnymi takimi, że $\mathbb{E}_Z a_{\mathrm{i}}(Z_{i_d})=0$. Wówczas istnieją stałe K_d takie, że dla dowolnego $p\geqslant 2$ zachodzi*

$$\begin{aligned}\mathbb{E}\|(\sum_{i_d} a_{\mathrm{i}}(Z_{i_d}))_{\mathrm{i}_{I_{d-1}}}\| \leqslant & K_d\sum_{\mathcal{J}\in\mathcal{P}_{I_d}} p^{(1+\deg(\mathcal{J})-d)/2}\|(a_{\mathrm{i}}(Z_{i_d}))_{\mathrm{i}}\|_{\mathcal{J}}\\ &+K_d\sum_{J\in\mathcal{P}_{I_{d-1}}} p^{1+(1+\deg(\mathcal{J})-d)/2}\sqrt{\mathbb{E}\max_{i_d}\|(a_{\mathrm{i}}(Z_{i_d}))_{\mathrm{i}_{I_{d-1}}}\|_{\mathcal{J}}^2},\end{aligned}$$

gdzie $\|\cdot\|$ oznacza normę $(d-1)$-wskaźnikowej macierzy, jako $(d-1)$-liniowego funkcjonału na $(\ell_2)^{d-1}$ (czyli normę $\|\cdot\|_{\{1\}\ldots\{d-1\}}$z Definicji 5). W szczególności,

$$\begin{aligned}\mathbb{E}_d\|(\sum_{i_d} h_{\mathrm{i}})_{\mathrm{i}_{I_{d-1}}}\|_{\{1\}\ldots\{d-1\}} \leqslant & K_d\sum_{\mathcal{J}\in\mathcal{P}_{I_d}} p^{(1+\deg(\mathcal{J})-d)/2}\|(h_{\mathrm{i}})_{\mathrm{i}}\|_{\mathcal{J}}\\ &+K_d\sum_{J\in\mathcal{P}_{I_{d-1}}} p^{1+(1+\deg(\mathcal{J})-d)/2}\sqrt{\mathbb{E}\max_{i_d}\|(h_{\mathrm{i}})_{\mathrm{i}_{I_{d-1}}}\|_{\mathcal{J}}^2}.\end{aligned}$$

Uwaga W Definicji 9 przyjęliśmy konwencję niezaznaczania zależności $\|(h_{\mathrm{i}})_{\mathrm{i}}\|_{\mathcal{J}}$ od zmiennych losowych $(X_i^{(j)})$. W powyższym twierdzeniu, $\|(\sum_{i_d} h_{\mathrm{i}})_{\mathrm{i}_{I_{d-1}}}\|_{\{1\}\ldots\{d-1\}}$, oznacza zmienną losową, zależną od $(X_i^{(d)})_{i\in I_n}$, zdefiniowaną jako $\|(\tilde{h}_{\mathrm{i}})_{\mathrm{i}\in I_{d-1}}\|_{\{1\}\ldots\{d-1\}}$, gdzie $(\tilde{h}_{\mathrm{i}})_{\mathrm{i}\in I_{d-1}}$ jest $(d-1)$-wskaźnikową losową macierzą funkcji, danych wzorem $\tilde{h}_{i_1,\ldots,i_{d-1}}(x_1,\ldots,x_{d-1}) := \sum_{i_d} h_{\mathrm{i}}(x_1,\ldots,x_{d-1},X_{i_d}^{(d)})$. Innymi słowy, dla ustalonej wartości $(X_{i_d}^{(d)})_{i_d\in I_n}$, warunkowo liczymy $\|(\tilde{h}_{\mathrm{i}})_{\mathrm{i}\in I_{d-1}}\|_{\{1\}\ldots\{d-1\}}$, traktując pozostałe zmienne $(X_i^{(k)})_{i\in I_n,k\in I_{d-1}}$ jako argumenty odpowiedniej U-statystyki.

Druga część twierdzenia wynika z pierwszej, jak w przypadku Twierdzenia 14, poprzez przejście do odpowiednich baz ortonormalnych. Kluczowe jest tu poniższe uogólnienie Lematu 11, wynikające z interpretacji norm $\|\cdot\|_{\mathcal{J}}$ jako norm funkcjonałów wieloliniowych.

LEMAT 14. *Niech dla $k\in I_{d-1}$, $\mathbf{e}_{k,m}=(e_{k,m}^1(X_1^{(k)}),\ldots,e_{k,m}^n(X_n^{(k)}))$, $m=1,2,\ldots$, będzie bazą ortonormalną w $L^2(X_1^{(k)})\oplus\ldots\oplus L^2(X_n^{(k)})$. Zdefiniujmy*

$$a_{\mathrm{i}}(X_{i_d}^{(d)}) := \sum_{|\mathrm{j}|\leqslant n, j_d=i_d}\mathbb{E}_{\{1,\ldots,d-1\}}[h_{\mathrm{j}}(\mathbf{X}_{\mathrm{j}}^{\mathrm{dec}})\prod_{k=1}^{d-1} e_{k,i_k}^{j_k}(X_{j_k}^{(k)})].$$

Wówczas, dla dowolnego rozbicia $\mathcal{J} \in \mathcal{P}_{I_d}$,

$$\|(h_{\mathrm{i}})_{\mathrm{i}}\|_{\mathcal{J}} = \|(a_{\mathrm{i}}(X^{(d)}_{i_d}))_{\mathrm{i}}\|_{\mathcal{J}}.$$

Powołamy się wreszcie na uogólnienie Lematu 12.

Lemat 15 (Latała, [35], Theorem 2). *Istnieją stałe K_d takie, że dla wszystkich $p \geqslant 2$ oraz dowolnej d-wskaźnikowej macierzy $A = (a_{\mathrm{i}})_{\mathrm{i} \in I_n^d}$,*

$$\mathbb{E}\|(\sum_{i_d} a_{\mathrm{i}} g_{i_d})_{\mathrm{i}_{I_{d-1}}}\|_{\{1\}\ldots\{d-1\}} \leqslant K_d \sum_{\mathcal{J} \in \mathcal{P}_{I_d}} p^{(1+\deg \mathcal{J}-d)/2} \|(a_{\mathrm{i}})\|_{\mathcal{J}}.$$

Dowód Twierdzenia 17. Podobnie jak w dowodzie Twierdzenia 14, możemy założyć, że nośnik macierzy jest ograniczony. Randomizując niezależnymi zmiennymi Rademachera, ograniczając średnie rademacherowe przez gaussowskie oraz stosując oszacowania dla średnich gaussowskich, warunkowo na Z (Lemat 15), otrzymujemy

$$\mathbb{E}\|(\sum_{i_d} a_{\mathrm{i}}(Z_{i_d}))_{\mathrm{i}_{I_{d-1}}}\| \leqslant K_d \sum_{\mathcal{J} \in \mathcal{P}_{I_d}} p^{(1+\deg \mathcal{J}-d)/2} \mathbb{E}\|(a_{\mathrm{i}}(Z_{i_d}))\|_{\mathcal{J},D}. \tag{3.10}$$

Rozważmy dowolny składnik z prawej strony (3.10), odpowiadający $\mathcal{J} = \{J_1, \ldots, J_k\}$ dla $\deg(\mathcal{J}) > 1$. Bez straty ogólności możemy założyć, że $d \in J_1$. Mamy (z Lematu 13, używając argumentów analogicznych jak w dowodzie Twierdzenia 14) dla $0 < \varepsilon \leqslant 1$,

$$\begin{aligned}
&p^{(1+k-d)/2}\mathbb{E}\|(a_{\mathrm{i}}(Z_{i_d}))_{\mathrm{i}}\|_{\mathcal{J},D} \\
&\quad\leqslant p^{(1+k-d)/2}\sqrt{\mathbb{E} \sup_{\|(x^{(j)}_{\mathrm{i}_{J_j}})\|_2 \leqslant 1:\, j=2,\ldots,k} \sum_{i_d} \sum_{\mathrm{i}_{J_1\setminus\{d\}}} (\sum_{\mathrm{i}_{I_d \setminus J_1}} a_{\mathrm{i}}(Z_{i_d}) \prod_{j=2}^{k} x^{(j)}_{\mathrm{i}_{J_j}})^2} \\
&\quad= p^{(1+k-d)/2} \times \\
&\quad\times \sqrt{\mathbb{E} \sup_{\|(x^{(j)}_{\mathrm{i}_{J_j}})\|_2 \leqslant 1:\, j=2,\ldots,k} \sum_{i_d} \Big\{\varepsilon_{i_d}[\sum_{\mathrm{i}_{J_1\setminus\{d\}}} (\sum_{\mathrm{i}_{I_d \setminus J_1}} a_{\mathrm{i}}(Z_{i_d}) \prod_{j=2}^{k} x^{(j)}_{\mathrm{i}_{J_j}})^2]^{1/2}\Big\}^2} \\
&\quad\leqslant K p^{(1+k-d)/2}\Big(\|(a_{\mathrm{i}}(Z_{i_d}))\|_{\mathcal{J}} \\
&\qquad + \frac{\varepsilon}{\sqrt{p}} \mathbb{E} \sup_{\|(x^{(j)}_{\mathrm{i}_{J_j}})\|_2 \leqslant 1:\, j=2,\ldots,k} \Big|\sum_{i_d} g_{i_d} \sqrt{\sum_{\mathrm{i}_{J_1\setminus\{d\}}} (\sum_{\mathrm{i}_{I_d \setminus J_1}} a_{\mathrm{i}}(Z_{i_d}) \prod_{j=2}^{k} x^{(j)}_{\mathrm{i}_{J_j}})^2)}\Big| \\
&\qquad + \frac{\sqrt{p}}{\varepsilon} \sqrt{\mathbb{E} \max_{i_d} \|(a_{\mathrm{i}}(Z_{i_d}))_{\mathrm{i}_{I_{d-1}}}\|^2_{J_1\setminus\{d\}, J_2, \ldots, J_k}}\Big),
\end{aligned}$$

gdzie dla $J_1 = \{d\}$ nadużywamy nieznacznie notacji, utożsamiając rodzinę zbiorów $\{\emptyset, J_2, \ldots, J_k\}$ z rozbiciem $\{J_2, \ldots, J_k\}$ zbioru I_{d-1}.

Podobnie jak w przypadku $d = 3$, Lemat Slepiana implikuje, że

$$\begin{aligned}
\mathbb{E} \sup_{\|(x^{(j)}_{\mathrm{i}_{J_j}})\|_2 \leqslant 1\colon j=2,\ldots,k} &\Big|\sum_{i_d} g_{i_d} \sqrt{\sum_{\mathrm{i}_{J_1\setminus\{d\}}} \Big(\sum_{\mathrm{i}_{I_d\setminus J_1}} a_{\mathrm{i}}(Z_{i_d}) \prod_{j=2}^{k} x^{(j)}_{\mathrm{i}_{J_j}}\Big)^2}\Big| \\
&\leqslant 2\mathbb{E} \sup_{\|(x^{(j)}_{\mathrm{i}_{J_j}})\|_2 \leqslant 1\colon j=2,\ldots,k} \Big|\sum_{\mathrm{i}_{J_1}} g_{\mathrm{i}_{J_1}} \sum_{\mathrm{i}_{I_d\setminus J_1}} a_{\mathrm{i}}(Z_{i_d}) \prod_{j=2}^{k} x^{(j)}_{\mathrm{i}_{J_j}}\Big| \\
&\leqslant K_d \sum_{\mathcal{K}\in\mathcal{P}_{I_d}, \deg(\mathcal{K})\leqslant k} p^{(1+\deg(\mathcal{K})-k)/2} \mathbb{E}\|(a_{\mathrm{i}}(Z_{i_d}))_{\mathrm{i}}\|_{\mathcal{K},D},
\end{aligned}$$

przy czym w ostatniej nierówności ponownie użyliśmy Lematu 15 (grupując indeksy odpowiadające zbiorom $J_1\backslash\{d\}, J_2, \ldots, J_k$ i traktując je jako nowe, pojedyncze indeksy).

Zauważmy że dla $J_1 = \{d\}$, użycie Lematu Slepiana niczego nie zmienia (podobnie jak dla $d = 3$), jednak nie wyróżniamy tego przypadku, aby skrócić i tak rozbudowany pod względem zapisu dowód.

Ostatecznie otrzymujemy

$$\begin{aligned}
p^{(1+\deg(\mathcal{J})-d)/2}&\mathbb{E}\|(a_{\mathrm{i}}(Z_{i_d}))_{\mathrm{i}}\|_{\mathcal{J},D} \\
&\leqslant K_d p^{(1+\deg(\mathcal{J})-d)/2}\|(a_{\mathrm{i}}(Z_{i_d}))_{\mathrm{i}}\|_{\mathcal{J}} \\
&+ K_d\varepsilon \sum_{\mathcal{K}\in\mathcal{P}_{I_d}, \deg(\mathcal{K})\leqslant k} p^{(1+\deg(\mathcal{K})-d)/2}\mathbb{E}\|(a_{\mathrm{i}}(Z_{i_d}))_{\mathrm{i}}\|_{\mathcal{K},D} \\
&+ K_d\varepsilon^{-1} p^{1+(1+\deg(J_1\backslash\{d\},J_2,\ldots,J_k)-d)/2} \sqrt{\mathbb{E}\max_{i_d} \|(a_{\mathrm{i}}(Z_{i_d}))_{\mathrm{i}_{I_{d-1}}}\|^2_{J_1\backslash\{d\},J_2,\ldots,J_k}}
\end{aligned}$$

Nierówność ta jest prawdziwa także dla $\deg(\mathcal{J}) = 1$ (tzn. $\mathcal{J} = \{I_d\}$), gdyż $\mathbb{E}\|(a_{\mathrm{i}}(Z_{i_d}))_{\mathrm{i}}\|_{\{I_d\},D} \leqslant \|(a_{\mathrm{i}}(Z_{i_d}))_{\mathrm{i}}\|_{\{I_d\}}$.

Sumując po wszystkich $\mathcal{J} \in \mathcal{P}_{I_d}$, otrzymujemy

$$\begin{aligned}
\sum_{\mathcal{J}\in\mathcal{P}_{I_d}} p^{(1+\deg(\mathcal{J})-d)/2}&\mathbb{E}\|(a_{\mathrm{i}}(Z_{i_d}))_{\mathrm{i}}\|_{\mathcal{J},D} \\
&\leqslant K_d \sum_{\mathcal{J}\in\mathcal{P}_{I_d}} p^{(1+\deg(\mathcal{J})-d)/2}\|(a_{\mathrm{i}}(Z_{i_d}))_{\mathrm{i}}\|_{\mathcal{J}} \\
&+ K_d\varepsilon \sum_{\mathcal{J}\in\mathcal{P}_{I_d}} p^{(1+\deg(\mathcal{J})-d)/2}\mathbb{E}\|(a_{\mathrm{i}}(Z_{i_d}))_{\mathrm{i}}\|_{\mathcal{J},D} \\
&+ \frac{K_d}{\varepsilon} \sum_{\mathcal{J}\in\mathcal{P}_{I_{d-1}}} p^{1+(1+\deg(\mathcal{J})-d)/2} \sqrt{\mathbb{E}\max_{i_d} \|(a_{\mathrm{i}}(Z_{i_d}))_{\mathrm{i}_{I_{d-1}}}\|^2_{\mathcal{J}}}.
\end{aligned}$$

Ustalając ε dostatecznie małe, otrzymujemy stąd górne ograniczenie na prawą stronę (3.10), co na mocy Lematu 14, pozwala zakończyć dowód. $\square$

Zmierzając w stronę dowodu Twierdzenia 16, wprowadźmy następującą definicję.

DEFINICJA 11. *Zdefiniujmy częściowy porządek $\prec$ na $\mathcal{P}_I$ jako*

$$\mathcal{I} \prec \mathcal{J}$$

wtedy i tylko wtedy, gdy dla wszystkich $I \in \mathcal{I}$, istnieje zbiór $J \in \mathcal{J}$ taki, że $I \subseteq J$.

Twierdzenie 17 daje następujący

WNIOSEK 6. *Niech $\mathcal{I} \in \mathcal{P}_{I_{d-1}}$. Wówczas*

$$\begin{aligned}\mathbb{E}_d \|(\sum_{i_d} h_{\mathrm{i}})_{\mathrm{i}_{I_{d-1}}}\|_{\mathcal{I}} \leqslant K_d \Big[& \sum_{\mathcal{J} \in \mathcal{P}_{I_d} \colon \mathcal{I} \cup \{\{d\}\} \prec \mathcal{J}} p^{(\deg(\mathcal{J}) - \deg(\mathcal{I}))/2} \|(h_{\mathrm{i}})_{\mathrm{i}}\|_{\mathcal{J}} \\ & + \sum_{\mathcal{J} \in \mathcal{P}_{I_{d-1}} \colon \mathcal{I} \prec \mathcal{J}} p^{1 + (\deg(\mathcal{J}) - \deg(\mathcal{I}))/2} \sqrt{\mathbb{E}_d \max_{i_d} \|(h_{\mathrm{i}})_{\mathrm{i}}\|_{\mathcal{J}}^2} \Big].\end{aligned}$$

Dość techniczny dowód powyższego wniosku przekładamy do części poświęconej U-statystykom o wartościach w przestrzeniach Hilberta, gdzie będziemy potrzebować nieco ogólniejszego faktu (Lemat 17).

Możemy teraz przystąpić do dowodu Twierdzenia 16. W przypadku $d \leqslant 3$ wywnioskowaliśmy je z Twierdzenia 12. Okazuje się jednak, że w ogólnym przypadku wygodniej jest naśladować dowód Twierdzenia 12 i użyć indukcji ze względu na d.

LEMAT 16. *Istnieją stałe K_d takie, że dla dowolnego $p \geqslant 2$,*

$$\mathbb{E}\Big|\sum_{\mathrm{i}} h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big|^p \leqslant K_d^p \sum_{I \subseteq I_d} \sum_{\mathcal{J} \in \mathcal{P}_I} \sum_{\mathrm{i}_{I^c}} p^{p(\# I^c + \deg(\mathcal{J})/2)} \mathbb{E}_{I^c} \|(h_{\mathrm{i}})_{\mathrm{i}_I}\|_{\mathcal{J}}^p. \tag{3.11}$$

DOWÓD. Prosta indukcja w duchu dowodu Twierdzenia 11. Dla $d = 1$, (3.11) jest natychmiastowym wnioskiem z (3.1), gdyż $\mathbb{E}|\sum h_i| \leqslant \sqrt{\mathbb{E}\sum h_i^2} = \|(h_i)_i\|_{\{1\}}$. W kroku indukcyjnym, stosujemy założenie indukcyjne (warunkowo na $X_{i_d}^{(d)}$) do funkcji $\tilde{h}_{i_1,\ldots,i_{d-1}}(x_1,\ldots,x_{d-1}) := \sum_{i_d} h_{\mathrm{i}}(x_1,\ldots,x_{d-1},X_{i_d}^{(d)})$, otrzymując

$$\mathbb{E}\Big|\sum_{\mathrm{i}} h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big|^p \leqslant K_d^p \sum_{I \subseteq I_{d-1}} \sum_{\mathcal{J} \in \mathcal{P}_I} \sum_{\mathrm{i}_{\tilde{I}^c}} p^{p(\# \tilde{I}^c + \deg(\mathcal{J})/2)} \mathbb{E}_{\tilde{I}^c} \mathbb{E}_d \|(\sum_{i_d} h_{\mathrm{i}})_{\mathrm{i}_I}\|_{\mathcal{J}}^p,$$

gdzie, podobnie jak w dowodzie Twierdzenia 11, $\tilde{I}^c = I^c \backslash \{d\}$.

Stosując Lemat 3, otrzymujemy dla $I \subseteq I_{d-1}$, $\mathcal{J} \in \mathcal{P}_I$,

$$\mathbb{E}_d \Big\|(\sum_{\mathrm{i}_d} h_{\mathrm{i}})_{\mathrm{i}_I}\Big\|_{\mathcal{J}}^p \leqslant K^p \Big[\Big(\mathbb{E}_d \|(\sum_{\mathrm{i}_d} h_{\mathrm{i}})_{\mathrm{i}_I}\|_{\mathcal{J}}\Big)^p + p^{p/2} \|(h_{\mathrm{i}})_{\mathrm{i}_{I \cup \{d\}}}\|_{\mathcal{J} \cup \{\{d\}\}} + p^p \mathbb{E} \max_{i_d} \|(h_{\mathrm{i}})_{\mathrm{i}_I}\|_{\mathcal{J}}^p\Big].$$

Aby zakończyć dowód, wystarczy zatem oszacować $\mathbb{E}_d \|(\sum_{\mathrm{i}_d} h_{\mathrm{i}})_{\mathrm{i}_I}\|_{\mathcal{J}}$. W tym celu używamy Wniosku 6, korzystając z faktu, że

$$\sqrt{\mathbb{E}_d \max_{i_d} \|(h_{\mathrm{i}})_{\mathrm{i}_{I_{d-1}}}\|_{\mathcal{J}}^2} \leqslant \Big(\mathbb{E}_d \sum_{i_d} \|(h_{\mathrm{i}})_{\mathrm{i}_{I_{d-1}}}\|_{\mathcal{J}}^p\Big)^{1/p}.$$

$\square$

Dowód Twierdzenia 16. Postępujemy podobnie jak w dowodzie Twierdzenia 12. Aby zastąpić sumy względem i_{I^c} po prawej stronie (3.11) przez odpowiednie maxima, wystarczy użyć Lematu 9 dla funkcji $g_{\mathrm{i}_{I^c}} = \|(h_\mathrm{i})_{\mathrm{i}_I}\|^2_{\mathcal{J}}$, z $p/2$ zamiast p oraz $\alpha = \deg(\mathcal{J}) + 2\#I^c + \#I^c$ oraz zauważyć, że z nierówności Cauchy'ego-Schwarza, dla $J \subseteq I^c$ oraz $\mathcal{J} \in \mathcal{P}_I$, mamy $\mathbb{E}_{I^c\setminus J} \sum_{\mathrm{i}_{I^c\setminus J}} \|(h_\mathrm{i})_{\mathrm{i}_I}\|^2_{\mathcal{J}} \leqslant \|(h_\mathrm{i})_{\mathrm{i}_{J^c}}\|^2_{\{J^c\}}$. □

3.2.1 Oszacowania odchyleń dla U-statystyk

Z nierówności Czebyszewa (Twierdzenie 3) oraz Twierdzenia 16 otrzymujemy następujący wniosek.

Twierdzenie 18. *Załóżmy, że wszystkie funkcje h_i są ograniczone. Wówczas dla pewnych stałych K_d oraz wszystkich $p \geqslant 2$,*

$$\mathbb{P}\left(\left|\sum_{\mathrm{i}} h_\mathrm{i}(\mathbf{X}_\mathrm{i}^{\mathrm{dec}})\right| > K_d \sum_{I\subseteq I_d} \sum_{\mathcal{J}\in\mathcal{P}_I} p^{\#I^c+\deg(\mathcal{J})/2} \max_{\mathrm{i}_{I^c}} \left\| \|(h_\mathrm{i})_{\mathrm{i}_I}\|_{\mathcal{J}} \right\|_\infty \right) \leqslant e^{-p}.$$

Równoważnie, dla $t \geqslant 0$,

$$\mathbb{P}\left(\left|\sum_{\mathrm{i}} h_\mathrm{i}(\mathbf{X}_\mathrm{i}^{\mathrm{dec}})\right| \geqslant t\right) \leqslant K_d \exp\left[-\frac{1}{K_d} \min_{I\subseteq I_d, \mathcal{J}\in\mathcal{P}_I} \left(\frac{t}{\left\| \|(h_\mathrm{i})_{\mathrm{i}_I}\|_{\mathcal{J}} \right\|_\infty}\right)^{2/(\deg(\mathcal{J})+2\#I^c)}\right].$$

Uwaga Powyższe twierdzenie jest w pewnym sensie optymalne. Nierówności Latały dla chaosów gaussowskich (wniosek z Twierdzenia 9), mówią, że gdy $h_\mathrm{i}(x_1, \ldots, x_d) = a_\mathrm{i} x_1 \ldots x_d$ oraz $(X_i^{(j)})$ są standardowymi zmiennymi normalnymi, to

$$\mathbb{P}(|\sum_{\mathrm{i}} h_\mathrm{i}(\mathbf{X}_\mathrm{i}^{\mathrm{dec}})| \geqslant k_d \sum_{\mathcal{J}\in\mathcal{P}_{I_d}} p^{\deg(\mathcal{J})/2} \|(h_\mathrm{i})_\mathrm{i}\|_{\mathcal{J}}) \geqslant k_d \wedge e^{-p},$$
$$\mathbb{P}(|\sum_{\mathrm{i}} h_\mathrm{i}| \geqslant K_d \sum_{\mathcal{J}\in\mathcal{P}_{I_d}} p^{\deg(\mathcal{J})/2} \|(h_\mathrm{i})_\mathrm{i}\|_{\mathcal{J}}) \leqslant e^{-p},$$

co, razem z centralnym twierdzeniem granicznym dla U-statystyk, pokazuje, że z dokładnością do stałych, składniki $p^{(\#I^c+\deg(\mathcal{J})/2)}\|(h_\mathrm{i})_{\mathrm{i}_I}\|_{\mathcal{J}}$ dla $I = I_d$ są poprawne i nie mogą być zaniedbane. Dla dyskusji nad optymalnością pozostałych składników, rozważmy $V = \mathbb{G}\prod_{i\in I} X_i$, gdzie X_i i $\mathbb{G}$ są niezależne, X_i są symetrycznymi zmiennymi Poissona z parametrem 1, zaś $\mathbb{G} = \sum_{\mathrm{i}_{I^c}} x_{\mathrm{i}_{I^c}} \prod_{j\in I^c} g_{i_j}^{(j)}$ to chaos gaussowski ($g_i^{(j)}$ są niezależnymi zmiennymi $\mathcal{N}(0,1)$). Wtedy V jest słabą granicą U-statystyk V_n zadanych przez funkcje $\prod_{i\in I} X_{n,i_j}^{(j)} a_{n,\mathrm{i}_{I^c}} \prod_{j\notin I} g_{i_j}^{(j)}$ ($\mathrm{i} \in I_n^d$), gdzie $X_{n,i_j}^{(j)}$ to scentrowane zmienne Bernoulliego z parametrem $p = 1/n$, a współczynniki $a_{n,\mathrm{i}_{I^c}}$ są odpowiednio dobrane z nieskończonej podzielności zmiennych gaussowskich. Wówczas $\mathbb{P}(V \geqslant k_d \alpha_p \sum_{\mathcal{J}\in\mathcal{P}_{I^c}} p^{\deg(\mathcal{J})/2} \|(x_{\mathrm{i}_{I^c}})\|_{\mathcal{J}}) \geqslant k_d \wedge e^{-p}$, gdzie $\alpha_p^{1/\#I} \log\alpha_p \sim p$, co pokazuje, że z dokładnością do czynnika rzędu $(\log p)^{\#I}$, także pozostałe składniki są poprawne.

Zauważmy także, że gdy $X_i^{(j)}$ mają ten sam rozkład oraz $h_{\mathbf{i}} = h$ dla pewnej funkcji h, wielkości z Definicji 9 upraszczają się. Dokładniej $\Big\|\|(h_{\mathbf{i}})_{i_I}\|_{\mathcal{J}}\Big\|_\infty = n^{\#I/2}\Big\|\|h\|_{\mathcal{J}}\Big\|_\infty$, gdzie $\|h\|_{\mathcal{J}}$ oznacza odpowiednią normę macierzy jednoelementowej $(h_{(1,\ldots,1)})$ z $h_{(1,\ldots,1)} = h$, np. dla $d = 3$:

$$\begin{aligned}\|h(X_1, X_2, X_3)\|_{\{1,2\}\{3\}} = \sup\{&\mathbb{E}h(X_1, X_2, X_3)f(X_1, X_2)g(X_3)\colon \\ &\mathbb{E}f(X_1, X_2)^2, \mathbb{E}g(X_3)^2 \leqslant 1\}\end{aligned}$$

Otrzymujemy zatem

Wniosek 7. *Jeśli $(X_i^{(j)})_{i\in I_n, j\in I_d}$ są zmiennymi i.i.d., to dla dowolnego $t \geqslant 0$,*

$$\mathbb{P}\left(\left|\sum_{\mathbf{i}} h(\mathbf{X}_{\mathbf{i}}^{\mathrm{dec}})\right| \geqslant t\right) \leqslant K_d \exp\left[-\frac{1}{K_d} \min_{I\subseteq I_d, \mathcal{J}\in\mathcal{P}_I} \left(\frac{t}{n^{\#I/2}\Big\|\|h\|_{\mathcal{J}}\Big\|_\infty}\right)^{2/(\deg(\mathcal{J})+2\#I^c)}\right].$$

Uwaga W szczególności widzimy, że odchylenia U-statystyki, generowanej przez ograniczoną kanoniczną funkcję h jest rzędu $n^{d/2}$, co zgadza się z centralnym twierdzeniem granicznym dla kanonicznych U-statystyk [44].

3.3 U-statystyki o wartościach w przestrzeni Hilberta

Przedstawione do tej pory metody pozwalają bez większych trudności rozszerzyć oszacowania momentów i odchyleń rzeczywistych U-statystyk na przypadek ogólnej przestrzeni Hilberta. Pewnym problemem staje się notacja, gdyż okazuje się, że same odpowiedniki norm z Definicji 9 nie są wystarczające. W niniejszej sekcji przedstawimy nierówności dla przestrzeni Hilberta, podając przy tym dowód ogólniejszej wersji Wniosku 6 z dowodu oszacowań dla przypadku rzeczywistego. Pominiemy natomiast szczegóły argumentów indukcyjnych, które są takie same jak w dowodach Twierdzeń 11, 12 czy Lematu 16 i Twierdzenia 16.

Rozważmy zatem ośrodkową przestrzeń Hilberta $(H, |\cdot|)$. Zauważmy, że jeśli w Definicji 9, zinterpretujemy $|\cdot|$ jako normę w H, otrzymamy poprawną definicję $\|\cdot\|_{\mathcal{J}}$ dla U-statystyk o wartościach w H. Niestety, jak już wspomnieliśmy, wielkości te nie wystarczają do uzyskania właściwych (np. z punktu widzenia prawa iterowanego logarytmu) oszacowań. Okazuje się jednak, że Definicję 9 można nieco zmodyfikować.

Definicja 12. *Dla niepustych zbiorów $K \subseteq I \subseteq I_d$ rozważmy dowolne $\mathcal{J} = \{J_1, \ldots, J_k\} \in \mathcal{P}_{I \setminus K}$. Dla d-wskaźnikowej macierzy $(h_{\mathrm{i}})_{\mathrm{i} \in I_n^d}$ funkcji o wartościach w H oraz dowolnej ustalonej wartości i_{I^c}, zdefiniujmy*

$$
\begin{aligned}
\|(h_{\mathrm{i}})_{\mathrm{i}_I}\|_{K,\mathcal{J}} = \sup\Big\{ |\sum_{\mathrm{i}_I} \mathbb{E}_I \langle h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}), g_{\mathrm{i}_K}(\mathbf{X}_{\mathrm{i}_K}^{\mathrm{dec}})\rangle \prod_{j=1}^{\deg(\mathcal{J})} f_{\mathrm{i}_{J_j}}^{(j)}(\mathbf{X}_{\mathrm{i}_{J_j}}^{\mathrm{dec}})| :& \\
f_{\mathrm{i}_{J_j}}^{(j)} \colon \Sigma^{J_j} \to \mathbb{R},\ g_{\mathrm{i}_K} \colon \Sigma^K \to H,& \\
\mathbb{E}\sum_{\mathrm{i}_{J_j}} |f_{\mathrm{i}_{J_j}}^{(j)}(\mathbf{X}_{\mathrm{i}_{J_j}}^{\mathrm{dec}})|^2 \leqslant 1 \text{ dla } j = 1, \ldots, \deg(\mathcal{J}),& \\
\mathbb{E}\sum_{\mathrm{i}_K} |g_{\mathrm{i}_K}(\mathbf{X}_{\mathrm{i}_K}^{\mathrm{dec}})|^2 \leqslant 1 \Big\}.&
\end{aligned}
$$

Jeżeli jeden ze zbiorów K, I jest pusty, oznaczmy $\|(h_{\mathrm{i}})_{\mathrm{i}_I}\|_{K,\mathcal{J}} = \|(h_{\mathrm{i}})_{\mathrm{i}_I}\|_{\mathcal{J}}$, gdzie $\|\cdot\|_{\mathcal{J}}$ jest normą z Definicji 9.

Uwaga Definicja 9 może być zatem interpretowana jako w pewnym sensie „zdegenerowany" przypadek Definicji 12. Jest to uzasadnione, gdyż oczywiście w przypadku przestrzeni Hilberta, Definicję 9 można zapisać jako

$$
\begin{aligned}
\|(h_{\mathrm{i}})_{\mathrm{i}_I}\|_{\mathcal{J}} = \sup\Big\{ |\sum_{\mathrm{i}_I} \mathbb{E}_I \langle h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}), g\rangle \prod_{j=1}^{\deg(\mathcal{J})} f_{\mathrm{i}_{J_j}}^{(j)}(\mathbf{X}_{\mathrm{i}_{J_j}}^{\mathrm{dec}})| :& \\
g \in H,\ |g| \leqslant 1,\ f_{\mathrm{i}_{J_j}}^{(j)} \colon \Sigma^{J_j} \to \mathbb{R},& \\
\mathbb{E}\sum_{\mathrm{i}_{J_j}} |f_{\mathrm{i}_{J_j}}^{(j)}(\mathbf{X}_{\mathrm{i}_{J_j}}^{\mathrm{dec}})|^2 \leqslant 1 \text{ dla } j = 1, \ldots, \deg(\mathcal{J}) \Big\},&
\end{aligned}
$$

czyli w formie analogicznej jak w Definicji 12, z deterministycznym g.

Oszacowania momentów przyjmują następującą formę.

Twierdzenie 19. *Dla dowolnego $p \geqslant 2$,*

$$
\mathbb{E}|\sum_{\mathrm{i}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})|^p \leqslant K_d^p \Big(\sum_{K \subseteq I \subseteq I_d} \sum_{\mathcal{J} \in \mathcal{P}_{I \setminus K}} p^{p(\#I^c + \deg \mathcal{J}/2)} \mathbb{E}_{I^c} \max_{\mathrm{i}_{I^c}} \|(h_{\mathrm{i}})_{\mathrm{i}_I}\|_{K,\mathcal{J}}^p \Big).
$$

Potrzebny nam będzie następujący lemat, będący zapowiedzianym w poprzedniej sekcji uogólnieniem Wniosku 6.

Lemat 17. *Dla dowolnego $K \subseteq I_{d-1}$ oraz $\mathcal{J} = \{J_1, \ldots, J_k\} \in \mathcal{P}_{I_{d-1} \setminus K}$ i każdego*

$p \geqslant 2$, *zachodzi*

$$\mathbb{E}_d \|(\sum_{i_d} h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}))_{\mathrm{i}_{I_{d-1}}}\|_{K,\mathcal{J}}$$
$$\leqslant K_d \Big(\sum_{\substack{K \subseteq L \subseteq I_d,\ \mathcal{K} \in \mathcal{P}_{I_d \setminus L}: \\ \mathcal{J} \cup \{K, \{d\}\} \prec \mathcal{K} \cup \{L\}}} p^{(\deg \mathcal{K} - \deg \mathcal{J})/2} \|(h_{\mathrm{i}})_{\mathrm{i}_{I_d}}\|_{L,\mathcal{K}}$$
$$+ \sum_{\substack{K \subseteq L \subseteq I_{d-1},\ \mathcal{K} \in \mathcal{P}_{I_{d-1} \setminus L}: \\ \mathcal{J} \cup \{K\} \prec \mathcal{K} \cup \{L\}}} p^{1+(\deg \mathcal{K} - \deg \mathcal{J})/2} \sqrt{\mathbb{E}_d \max_{i_d} \|(h_{\mathrm{i}})_{\mathrm{i}_{I_{d-1}}}\|_{L,\mathcal{K}}^2} \Big).$$

Dowód Lematu 17 jest nieskomplikowany i polega głównie na zmianie bazy, ze względu jednak na skomplikowanie oznaczeń, trudno jest zapisać go bezpośrednio i wygodniej będzie użyć notacji związanej z iloczynami tensorowymi. W tym celu potrzebny będzie następujący

LEMAT 18. *Niech H będzie ośrodkową przestrzenią Hilberta, zaś X zmienną losową o wartościach w Σ. Wówczas, $H \otimes L^2(X) \simeq L^2(X, H)$, gdzie $L^2(H)$ to przestrzeń całkowalnych z kwadratem zmiennych losowych postaci $f(Y)$, $f\colon \Sigma \to H$-mierzalne (iloczyn tensorowy rozumiany jest jako iloczyn przestrzeni Hilberta). Przy powyższym utożsamieniu, dla $h \in H$, $f(X) \in L^2(X)$, mamy $h \otimes f(X) = hf(X) \in L^2(X, H)$.*

DOWÓD. Lemat wynika w prosty sposób z przedstawienia H jako $L^2(Y)$, gdzie Y jest zmienną niezależną od X oraz standardowej równości $L^2(X) \otimes L^2(Y) \simeq L^2(X, Y)$. □

DOWÓD LEMATU 17. Argumentacja, którą przedstawimy nie jest skomplikowana, niemniej dość kłopotliwa w zapisie. W związku z tym, zamiast podawać wszystkie szczegóły techniczne, naszkicujemy jedynie główną ideę, opuszczając dość oczywiste rachunki, związane np. z utożsamieniem różnych iloczynów tensorowych przestrzeni Hilberta. Podobnie, rozważając funkcjonały liniowe na przestrzeni zadanej na kilka sposobów (np. explicite i jako iloczyn tensorowy), będziemy przechodzić bez szczegółowych wyjaśnień do zapisu odpowiadającego różnym reprezentacjom.

Niech
$$H_0 = H \otimes \Big[\otimes_{l \in K} (\oplus_{i=1}^n L^2(X_i^{(l)})\Big] \simeq \oplus_{|\mathrm{i}_K| \leqslant n} L^2(\mathbf{X}_{\mathrm{i}_K}^{\mathrm{dec}}, H)$$
oraz, dla $j = 1, \ldots, k$,
$$H_i = \otimes_{l \in J_j} (\oplus_{i=1}^n L^2(X_i^{(l)})) \simeq \oplus_{|\mathrm{i}_{J_j}| \leqslant n} L^2(\mathbf{X}_{\mathrm{i}_{J_j}}^{\mathrm{dec}}).$$

W sytuacji, gdy $K = \emptyset$, mamy zachowując naturalne konwencję dot. pustych iloczynów, $H_0 \simeq H$.

Dla $i_d = 1,\ldots,n$ oraz ustalonej wartości $X_{i_d}^{(d)}$, niech A_{i_d} będzie funkcjonałem liniowym na $\tilde{H} = \oplus_{|\mathrm{i}_{I_{d-1}}|\leqslant n} L^2(\mathbf{X}^{\mathrm{dec}}_{\mathrm{i}_{I_{d-1}}}, H) \simeq \otimes_{j=0}^k H_k$, zadanym przez $(h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}))_{|\mathrm{i}_{I_{d-1}}|\leqslant n} \in \tilde{H}$, wzorem

$$\begin{aligned} A_{i_d}((g_{\mathrm{i}_{I_{d-1}}}(\mathbf{X}^{\mathrm{dec}}_{\mathrm{i}_{I_{d-1}}}))_{\mathrm{i}_{I_{d-1}}}) &= \langle (g_{\mathrm{i}_{I_{d-1}}}(\mathbf{X}^{\mathrm{dec}}_{\mathrm{i}_{I_{d-1}}}))_{\mathrm{i}_{I_{d-1}}}, (h_{\mathrm{i}}(\mathbf{X}^{\mathrm{dec}}_{\mathrm{i}}))_{\mathrm{i}_{I_{d-1}}} \rangle_{\tilde{H}} \\ &= \sum_{|\mathrm{i}_{I_{d-1}}|\leqslant n} \mathbb{E}_{\{1,\ldots,d-1\}} \langle g_{\mathrm{i}_{I_{d-1}}}(\mathbf{X}^{\mathrm{dec}}_{\mathrm{i}_{I_{d-1}}}), h_{\mathrm{i}}(\mathbf{X}^{\mathrm{dec}}_{\mathrm{i}}) \rangle_H. \end{aligned}$$

Jako funkcje $X_{i_d}^{(d)}$, $A_{i_d} = A_{i_d}(X_{i_d}^{(d)})$ są niezależnymi losowymi funkcjonałami liniowymi. Wyznaczają zatem także losowe funkcjonały $(k+1)$-liniowe na $\oplus_{j=0}^k H_k$, dane wzorem

$$(h_0, h_1, \ldots, h_k) \mapsto A_{i_d}(h_0 \otimes h_1 \otimes \ldots \otimes h_k).$$

Jeśli przez $\|\cdot\|$ oznaczymy normę funkcjonału $(k+1)$-liniowego, prawa strona nierówności z Lematu, może być zapisana jako

$$\mathbb{E}\Big\| \sum_{i_d=1}^n A_{i_d}(X_{i_d}^{(d)}) \Big\|.$$

Ponadto, oznaczając przez $\|A_{i_d}\|_{HS}$ normę A_{i_d} jako operatora liniowego na $\otimes_{j=0}^k H_j$ (przez analogię z normą Hilberta-Schmidta macierzy), mamy

$$\sum_{i_d=1}^n \mathbb{E}\|A_{i_d}(X_{i_d}^{(d)})\|_{HS}^2 = \|(h_{\mathrm{i}})_{\mathrm{i}}\|_{I_d,\emptyset}^2 < \infty,$$

zatem ciąg zmiennych losowych $A_{i_d}(X_{i_d}^{(d)})$, wyznacza funkcjonał liniowy A na przestrzeni $\tilde{H} \otimes [\oplus_{i_d=1}^n L^2(X_{i_d}^{(d)})] \simeq \oplus_{|\mathrm{i}|\leqslant n} L^2(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}, H) \simeq \oplus_{i_d=1}^n L^2(X_{i_d}^{(d)}, \tilde{H})$, dany wzorem

$$A(g_1(X_1^{(d)}), \ldots, g_n(X_n^{(d)})) = \sum_{i_d=1}^n \mathbb{E}[A_{i_d}(X_{i_d}^{(d)})(g_{i_d}(X_{i_d}^{(d)}))],$$

przy czym łatwo wykazać, że jeśli zinterpretujemy dziedzinę tego funkcjonału jako iloczyn $\oplus_{|\mathrm{i}|\leqslant n} L^2(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}, H)$, to odpowiada mu macierz $(h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}))_{\mathrm{i}}$.

Wprowadźmy dodatkowo następującą notację, spójną z definicją norm $\|\cdot\|_{\mathcal{J}}$. Jeśli T jest funkcjonałem liniowym na $\otimes_{j=0}^m E_j$ dla pewnych przestrzeni Hilberta E_j, zaś $\mathcal{I} = \{L_1, \ldots, L_r\} \in \mathcal{P}_{I_m \cup \{0\}}$, to przez $\|T\|_{\mathcal{I}}$ oznaczmy normę T jako funkcjonału r-liniowego na $\oplus_{i=1}^r [\otimes_{j \in L_i} E_j]$, danego wzorem

$$(e_1, \ldots, e_r) \mapsto T(e_1 \otimes \ldots \otimes e_r).$$

Oznaczając ponadto $H_{k+1} = \oplus_{i_d=1}^n L^2(X_{i_d}^{(d)})$, możemy zastosować powyższą definicję do $\tilde{H} \otimes [\oplus_{i_d=1}^n L^2(X_{i_d}^{(d)})] = \otimes_{j=0}^{k+1} H_j$ oraz skorzystać z Twierdzenia 17, aby otrzymać

$$\mathbb{E}\Big\|\sum_{i_d=1}^{n} A_{i_d}(X_{i_d}^{(d)})\Big\| \leqslant K_d \sum_{\mathcal{I}\in\mathcal{P}_{I_{k+1}\cup\{0\}}} p^{(1+\deg(\mathcal{I})-(k+2))/2}\|A\|_{\mathcal{J}} + K_d \sum_{I\in\mathcal{P}_{I_k\cup\{0\}}} p^{1+(1+\deg(\mathcal{I})-(k+2))/2}\sqrt{\mathbb{E}\max_{i_d}\|A_{i_d}(X_{i_d}^{(d)})\|_{\mathcal{I}}^2}. \tag{3.12}$$

Powyższa nierówność to już inaczej zapisana teza lematu, co wynika z „łączności" iloczynu tensorowego oraz jego „rozdzielności" względem sumy prostej przestrzeni. Rzeczywiście, oznaczając dodatkowo $J_{k+1} = \{d\}$, mamy dla $0 \notin L_i$, oraz $U = \bigcup_{j\in L_i} J_j$,

$$\otimes_{j\in L_i} H_j \simeq \otimes_{j\in L_i}\otimes_{l\in J_j}(\oplus_{s=1}^n L^2(X_s^{(l)})) \simeq \otimes_{l\in U}(\oplus_{s=1}^n L^2(X_s^{(l)})) \simeq \oplus_{|\mathrm{i}_U|\leqslant n} L^2(\mathbf{X}_{\mathrm{i}_U}^{\mathrm{dec}}).$$

Podobnie, gdy $0 \in L_i$,

$$\begin{aligned}\otimes_{j\in L_i} H_j &\simeq [\oplus_{|\mathrm{i}_K|\leqslant n} L^2(\mathbf{X}_{\mathrm{i}_K}^{\mathrm{dec}}, H)] \times [\otimes_{0\neq j\in L_i}\otimes_{l\in J_j}(\oplus_{s=1}^n L^2(X_s^{(l)}))] \\ &\simeq \oplus_{|\mathrm{i}_U|\leqslant n} L^2(\mathbf{X}_{\mathrm{i}_U}^{\mathrm{dec}}, H),\end{aligned}$$

gdzie $U = (\bigcup_{0\neq j\in L_i} J_j) \cup K$. Korzystając z faktu, że dla ustalonych $X_{i_d}^{(d)}$, A_{i_d} odpowiadają macierzy $(h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}))_{|\mathrm{i}_{I_{d-1}}|\leqslant n}$, zaś A odpowiada macierzy $(h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}))_{|\mathrm{i}|\leqslant n}$, otrzymujemy, że każdy składnik $\|\cdot\|_{\mathcal{I}}$ po prawej stronie (3.12) jest równy pewnemu składnikowi $\|\cdot\|_{L,\mathcal{K}}$ z prawej strony tezy. Mówiąc nieformalnie i nadużywając nieco notacji (w sytuacji gdy $K = \emptyset$), „łączymy" elementy rozbicia $\{\{d\}, J_1, \ldots, J_k, K\}$ lub $\{J_1, \ldots, J_k, K\}$ w sposób zadany przez rozbicie $\mathcal{I}$, uzyskując rozbicie $\{L\} \cup \mathcal{K}$, gdzie L jest zbiorem odpowiadającym w nowym rozbiciu zbiorowi $L_i \in \mathcal{I}$, zawierającemu 0 (w szczególności, jeśli $K = \emptyset$ oraz $\{0\} \in \mathcal{I}$, to $L = \emptyset$). Zauważmy ponadto, że $\deg(\mathcal{I}) = \deg(\mathcal{K}) + 1$, a zatem

$$1 + \deg(\mathcal{I}) - (k+2) = \deg(\mathcal{K}) - \deg(\mathcal{J}),$$

co dowodzi, że także potęgi zmiennej p po prawej stronie tezy oraz (3.12) są zgodne, pokazując równość obu wyrażeń i kończąc dowód. □

Dowód Twierdzenia 19. Dowód jest analogiczny jak dla Twierdzenia 16. Najpierw dowodzimy indukcyjnie słabszej nierówności z sumami zamiast maksimów (odpowiednik Lematu 16), korzystając z Lematu 3 i zastępując Wniosek 6 Lematem 17. Następnie, używając Lematu 9, zastępujemy sumy przez maksima. □

Na zakończenie podamy jeszcze oszacowania odchyleń implikowane przez Twierdzenie 19 i nierówność Czebyszewa (Twierdzenie 3).

Twierdzenie 20. *Jeśli wszystkie funkcje h_{i} są ograniczone, to dla każdego $t \geqslant 0$,*

$$\mathbb{P}\Big(\Big|\sum_{|\mathrm{i}|\leqslant n} h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| \geqslant K_d\sqrt{\sum_{\mathrm{i}} \mathbb{E}h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})^2} + t\Big)$$
$$\leqslant K_d \exp\Big[-\frac{1}{K_d}\min_{\substack{K\subseteq I\subseteq I_d\\ K\neq I_d \text{ lub } I\neq I_d}}\min_{\mathcal{J}\in\mathcal{P}_{I\setminus K}}\Big(\frac{t}{\big\|\|(h_{\mathrm{i}})_{\mathrm{i}_I}\|_{K,\mathcal{J}}\big\|_\infty}\Big)^{2/(\deg(\mathcal{J})+2\#I^c)}\Big].$$

3.4 Zastosowanie. Wielokrotne całki stochastyczne

Pierwszą część pracy, dotyczącą nierówności dla U-statystyk zakończymy podaniem zastosowań do szacowania odchyleń wielokrotnych całek stochastycznych funkcji deterministycznych względem procesów o przyrostach niezależnych. Obiekty te są bezpośrednio związane z U-statystykami, gdyż mogą być zdefiniowane jako odpowiednie izometrie (całki Itô), poprzez przybliżenie ich chaosami jednorodnymi. Ponieważ naszym celem jest jedynie zilustrowanie zastosowań nierówności dla U-statystyk, nie będziemy przypominać standardowych definicji, związanych z całkami stochastycznymi. Można je znaleźć np. w książce [33]

Oszacowania odchyleń dla wielokrotnych całek stochastycznych względem procesu Wienera, wynikają oczywiście z nierówności Latały (Twierdzenie 9). Dla procesów ze skokami oraz jednokrotnych i dwukrotnych całek stochastycznych, odpowiednie nierówności zostały znalezione przez Reynaud-Bouret [43] oraz Houdré i Reynaud-Bouret [29]. Nierówności te mają potencjalne zastosowania statystyczne, w estymacji funkcji intensywności niejednorodnych procesów Poissona. Metody dowodu we wspomnianych pracach różnią się od przedstawionej poniżej i polegają na nierównościach logarytmicznych Sobolewa [43] lub nierównościach martyngałowych. Wersja drugiej z tych metod (dla czasu dyskretnego) została użyta przez autorów także do uzyskania dość dobrych stałych w nierówności Giné, Latały i Zinna dla U-statystyk rzędu 2. W naszym przypadku, postąpimy inaczej i wykorzystamy bezpośrednio oszacowania U-statystyk, by otrzymać nierówności dla całek.

Przedstawmy podstawowe założenia. Niech $(N_t^{(i)})_{t\in[0,T]}$ $(i \in I_d)$ będą niezależnymi procesami càdlàg o przyrostach niezależnych, $N_0^{(i)} = 0$. Oznaczmy $V^i(t) = \mathrm{Var} N_t^{(i)}$. Niech ponadto $\Lambda^i(t) = \mathbb{E}N_t^{(i)}$ będzie kompensatorem $N^{(i)}$, zaś $\tilde{N}^{(i)}(t) = N(t) - \Lambda(t)$.

Założymy także, że wszystkie skoki procesów N_i sa jednostajnie ograniczone (skąd wynika, że $V^i(t) < \infty$, [15]). Ponieważ jest to tylko kwestia normalizacji, a w przypadku nieciągłym, typowym procesem, jaki mamy na myśli, jest niejednorodny proces Poissona, przyjmiemy, że wszystkie skoki są nie większe niż 1.

Definicja 13. *Dla niepustego podzbioru $I \subseteq I_d$ oraz $\mathcal{J} = \{J_i\}_{i=1}^k \in \mathcal{P}_I$, zdefiniujmy*

$$\begin{aligned} \|h\|_{\mathcal{J}} &= \sup\{\int_{[0,T]^{\#I}} h(t_1, \ldots, t_d) \prod_{j=1}^{\deg(\mathcal{J})} f^{(j)}((t_i)_{i\in J_j}) \prod_{i\in I} dV^i(t_i) \colon \\ &\quad \int_{[0,T]^{\#J_j}} |f^j((t_i)_{i\in J_j})|^2 \prod_{i\in J_j} dV^i(t_i) \leqslant 1\}. \end{aligned}$$

Zdefiniujmy także $\|h\|_\emptyset = |h|$.

Uwaga Zauważmy, że podobnie jak w przypadku U-statystyk, dla $I = I_d$, $\|h\|_{\mathcal{J}}$ jest normą, zaś dla $I \neq I_d$ – funkcją $(t_i)_{i\in I^c}$.

Oznaczenie $\|h\|_{\mathcal{J}}$ stoi w niewielkim konflikcie z notacją użytą we Wniosku 7, gdzie w taki sam sposób oznaczaliśmy normę funkcji, definiującej U-statystykę, opartą o ciąg zmiennych i.i.d. Ponieważ jednak nie będziemy teraz korzystać z tamtego oszacowania, nie powinno to prowadzić do niejednoznaczności.

Udowodnimy następujące

Twierdzenie 21. *Niech $h\colon [0,T]^d \to \mathbb{R}$ będzie ograniczoną funkcją borelowską. Rozważmy całkę stochastyczną*

$$Z = \int_{[0,T]\times\ldots\times[0,T]} h(t_1,\ldots,t_d) d\tilde{N}_{t_1}^{(1)} \ldots d\tilde{N}_{t_d}^{(d)}.$$

Istnieją stałe K_d takie, że dla wszystkich $p \geqslant 2$,

$$\mathbb{P}\left(|Z| > K_d \sum_{I\subseteq I_d} \sum_{\mathcal{J}\in\mathcal{P}_I} p^{\#I^c+\deg(\mathcal{J})/2} \max_{\mathrm{i}_{I^c}} \Big\| \|h\|_{\mathcal{J}} \Big\|_\infty \right) \leqslant e^{-p}.$$

Metoda dowodu powyższego twierdzenia polegać będzie na klasycznym przybliżaniu funkcji h przez funkcje schodkowe, zaś całki stochastycznej przez odpowiednie U-statystyki. Jednak aproksymować funkcję h możemy w L^2, natomiast Twierdzenie 21 zawiera normy typu L^∞. Należy zatem skonstruować ciąg funkcji prostych, przybliżających h_n, dla których te normy są ograniczone przez odpowiadające im normy funkcji h. Do konstrukcji wykorzystamy następujący lemat.

Lemat 19. *Rozważmy przestrzenie probabilistyczne (Ω_i, μ_i), $i \leqslant d$ oraz $\Omega = \times_{i\in I_d}\Omega_i$, $\mu = \otimes_{i\in I_d}\mu_i$. Istnieją stałe K_d takie, że dla każdego $\varepsilon > 0$ oraz każdego mierzalnego zbioru $A \subseteq \Omega$, spełniającego $\mu(A) > 1-\varepsilon$, istnieje $B \subseteq A$ taki, że $\mu(B) > 1 - K_d\varepsilon^{1/2^{d-1}}$ oraz dla wszystkich $\emptyset \neq I \subsetneq I_d$ i $x_I \in \times_{i\in I}\Omega_i$ albo $B^I_{x_I} = \emptyset$ albo $\mu_{I^c}(B^I_{x_I}) > 1 - K_d\varepsilon^{1/2^{d-1}}$, gdzie $\mu_{I^c} = \otimes_{i\in I^c}\mu_i$, zaś $B^I_{x_I} = \{y_{I^c} \in \times_{i\in I^c}\Omega_i \colon y \in B, y_I = x_I\}$.*

Dowód. Zacznijmy od kilku uwag, dot. notacji. W dowodzie będziemy używać indukcji ze względu na d i w kroku indukcyjnym będziemy rozpatrywać pewne zbiory $C \subseteq \times_{i\in I}\Omega_i$ dla $I \subseteq I_d$. Dla takich zbiorów oraz $J \subsetneq I$ i $x_J \in \times_{i\in J}\Omega_i$ przez $C^J_{x_J}$ będziemy oznaczać zbiór $\{y_{I\setminus J} \in \times_{i\in I\setminus J}\Omega_i \colon y_I \in C, y_J = x_J\}$, co może być nieznacznie niespójne z notacją z w sformułowaniu lematu. Dodatkowo, rozpatrując iloczyny kartezjańskie wielu zbiorów nie będziemy zwracać uwagi na kolejność mnożenia (spojrzenie na iloczyn kartezjański jako na zbiór funkcji określonych na zbiorze indeksującym czyni go „przemiennym”).

Przejdźmy do dowodu. Dla $d = 1$, teza jest oczywista. Załóżmy zatem, że teza zachodzi dla wszystkich liczb mniejszych od $d > 1$. Dla $\emptyset \neq I \subsetneq I_d$, niech $A_I = \{x_I \in \times_{i\in I}\Omega_i \colon \mu_{I^c}(A^I_{x_I}) > 1 - \sqrt{\varepsilon}\}$. Z twierdzenia Fubiniego, $\mu_I(A_I) > 1 - \sqrt{\varepsilon}$, a zatem na mocy założenia indukcyjnego, istnieją zbiory $B_I \subseteq A_I$, spełniające $\mu_I(B_I) > 1 - K_{d-1}\varepsilon^{1/2^{d-1}}$, takie, że wszystkie ich cięcia są albo puste, albo miary większej niż $1 - K_{d-1}\varepsilon^{1/2^{d-1}}$. Zdefiniujmy

$$B = \bigcap_{\emptyset\neq I\subsetneq I_d} \bigcup_{z_I\in B_I} \{z_I\} \times A^I_{z_I} = \bigcap_{\emptyset\neq I\subsetneq I_d} (B_I \times (\times_{i\in I^c}\Omega_i)) \cap A.$$

Wówczas $\mu(B) > 1 - K_d\varepsilon^{1/2^{d-1}}$. Rozważmy dowolny zbiór $\emptyset \neq J \subsetneq I_d$ oraz dowolny element x_J. Mamy

$$B^J_{x_J} = \{y_{J^c} \colon y \in B, y_J = x_J\} = \bigcap_{\emptyset\neq I\subsetneq I_d} \{y_{J^c} \colon y \in \bigcup_{z_I\in B_I} \{z_I\} \times A^I_{z_I}, y_J = x_J\}. \quad (3.13)$$

Wykażemy, że $B^J_{x_J}$ jest pusty lub miary większej niż $1 - K_d\varepsilon^{1/2^{d-1}}$. Załóżmy zatem, że istnieje $x_{J^c} \in B^J_{x_J}$. Niech x będzie elementem $\times_{i\in I_d}\Omega_i$ otrzymanym przez konkatenację x_J i x_{J^c}. Wówczas $x \in B$, a zatem $x_I \in B_I$ dla wszystkich $I \subsetneq I_d$. Stąd dla $I \cap J^c \neq \emptyset$, mamy $x_{I\cap J^c} \in (B_I)^{I\cap J}_{x_{I\cap J}}$, a więc ten zbiór jest niepusty, czyli z definicji B_I, miary większej niż $1 - K_{d-1}\varepsilon^{1/2^{d-1}}$. Zdefiniujmy teraz

$$U = A^J_{x_J} \cap \bigcap_{I\subsetneq I_d, I\cap J^c\neq\emptyset} \Big(\times_{i\in J^c\setminus I} \Omega_i \times (B_I)^{I\cap J}_{x_{I\cap J}}\Big)$$

Oczywiście $\mu_{J^c}(U) > 1-K_d\varepsilon^{1/2^{d-1}}$ ponieważ wszystkie przecinane zbiory mają miarę większą niż $1 - K_{d-1}\varepsilon^{1/2^{d-1}}$ (włączając $A^J_{x_J}$, gdyż $x_J \in B_J \subseteq A_J$). Dla $x_{J^c} \in U$, mamy $x \in A$ (gdzie x jest ponownie konkatenacją x_J i x_{J^c}). Ponadto dla $I \cap J^c \neq \emptyset$, $x_{I\cap J^c} \in (B_I)^{I\cap J}_{x_{I\cap J}}$, a więc $x_I \in B_I$. Z dyskusji przeprowadzonej tuż po założeniu, że $B^J_{x_J}$ jest niepusty, jest to także prawdą dla $\emptyset \neq I \subseteq J$. Zatem dla każdego $\emptyset \neq I \subsetneq I_d$ mamy $x \in \{x_I\} \times A^I_{x_I}$, gdzie $x_I \in B_I$. Z (3.13) wynika, że $x_{J^c} \in B^J_{x_J}$. Wykazaliśmy zatem, że $U \subseteq B^J_{x_J}$, co oczywiście implikuje, że $\mu_{J^c}(B^J_{x_J}) > 1 - K_d\varepsilon^{1/2^{d-1}}$. $\square$

Lemat 20. *Istnieją funkcje schodkowe, tzn. funkcje postaci*

$$h_n = \sum_{\mathrm{i}\in I^d_{k_n}} a^{(n)}_{\mathrm{i}} \mathbf{1}_{(t^{(n)}_{i_1}, t^{(n)}_{i_1+1}]\times\ldots\times(t^{(n)}_{i_d}, t^{(n)}_{i_d+1}]},$$

takie, że $h_n \to h$ p.n. oraz w L^2 (względem miary produktowej na $[0,T]^d$ o rozkładach jednowymiarowych zadanych przez V^i) oraz $\Big\|\|h_n\|_{\mathcal{J}}\Big\|_\infty \leqslant 3\Big\|\|h\|_{\mathcal{J}}\Big\|_\infty$ dla wszystkich $I \subsetneq I_d$ oraz $\mathcal{J} \in \mathcal{P}_I$.

Dowód. Zauważmy najpierw, że jeśli zastąpimy $N^{(i)}$ przez $c_i N^{(i)}$, to $\|h\|_{\mathcal{J}}$ zostanie przemnożone przez $\prod_{i\in I} c_i$. Zatem, bez straty ogólności możemy założyć, że $V^{(i)}(T) = 1$, co pozwoli nam użyć Lematu 19. Rozważmy dowolny ciąg $\tilde{h}_n$ funkcji schodkowych zbiegający p.n. do h, taki, że $\|\tilde{h}_n\|_\infty \leqslant \|h\|_\infty$. Dla dowolnego $I \subsetneq I_d$ oraz $\mathcal{J} \in \mathcal{P}_I$, mamy $\|\tilde{h}_n\|_{\mathcal{J}} \to \|h\|_{\mathcal{J}}$ p.n., zatem możemy przejść do podciągu i założyć, że dla dużego zbioru $A_{I^c}^{(n)}$ (powiedzmy z $(A_{I^c}^{(n)})^c$ miary mniejszej niż $\varepsilon/2^{n2^{d-1}}$, gdzie ε zostanie dobrane później), zachodzi $\|\tilde{h}_n\|_{\mathcal{J}} \leqslant 2\Big\|\|h\|_{\mathcal{J}}\Big\|_\infty$. Niech wówczas $B^{(n)}$ będzie podzbiorem $\bigcap_I (A_{I^c}^{(n)} \times [0,T]^I)$ danym przez Lemat 19 zastosowany do σ-ciała generowanego przez zbiory postaci $(t_{i_1}^{(n)}, t_{i_1+1}^{(n)}] \times \ldots \times (t_{i_d}^{(n)}, t_{i_d+1}^{(n)}]$, gdzie $t_i^{(n)}$ odpowiada funkcji schodkowej $\tilde{h}_n$ jak w sformułowaniu Lematu. Zdefiniujmy dla $t = (t_1, \ldots, t_d)$, $h_n(t) = \tilde{h}_n(t)\mathbf{1}_{B^{(n)}}(t)$. Wówczas h_n jest funkcją schodkową i z lematu Borela-Cantelli wciąż mamy $h_n \to h$ p.n. i w L^2 (z twierdzenia Lebesgue'a o zbieżności zmajoryzowanej). Ponadto, dla wszystkich I oraz t_{I^c}, funkcja $g_n(t_I) = h_n(t)$ jest albo tożsamościowo równa 0 albo różni się od $\tilde{h}_n$ na zbiorze miary mniejszej niż $K_d \varepsilon^{1/2^{d-1}}/2^n$. Jeśli g_n nie jest równe 0, mamy $t_{I^c} \in A_{I^c}^{(n)}$, więc w punkcie t_{I^c} $\|\tilde{h}_n\|_{\mathcal{J}} \leqslant 2\Big\|\|h\|_{\mathcal{J}}\Big\|_\infty$. Zatem $\|h_n\|_{\mathcal{J}} = 0$ lub $\|h_n\|_{\mathcal{J}} \leqslant \|h_n - \tilde{h}_n\|_{\mathcal{J}} + \|\tilde{h}_n\|_{\mathcal{J}} \leqslant K_d(\|h_n\|_\infty + \|\tilde{h}_n\|_\infty)\varepsilon^{1/2^d}/2^{n/2} + 2\Big\|\|h\|_{\mathcal{J}}\Big\|_\infty \leqslant 3\Big\|\|h\|_{\mathcal{J}}\Big\|_\infty$ dla ε dostatecznie małych. □

Dowód Twierdzenia 21. Udowodnimy oszacowania momentów dla całek z funkcji ograniczonych, teza twierdzenia będzie z nich wynikać poprzez nierówność Czebyszewa. Rozważmy zatem funkcje h_n, dane przez Lemat 20. Możemy oczywiście założyć, że $\max_{i\leqslant k_n} |t_{i+1}^{(n)} - t_i^{(n)}| \to_n 0$. Niech Z_n będzie d-krotną całką stochastyczną h_n. Ponieważ $h_n \to h$ w L^2, mamy $Z_n \to Z$ w L^2 oraz możemy założyć, że $Z_n \to Z$ p.n. Oznaczmy (z drobnym nadużyciem notacji) przez $\|Z_n\|_{\mathcal{J}}$ normy $\|\cdot\|_{\mathcal{J}}$ macierzy funkcji które definiują jednorodny chaos Z_n traktowany jako U-statystyka (oznaczając w ten sposób, odróżnimy te normy od $\|h_n\|_{\mathcal{J}}$ danych w Definicji 13). Łatwo zobaczyć, że dla $\mathcal{J} \in \mathcal{P}_{I_d}$ zachodzi $\|Z_n\|_{\mathcal{J}} \leqslant \|h_n\|_{\mathcal{J}}$ oraz dla $I \subsetneq I_d$, $\mathcal{J} \in \mathcal{P}_I$, dowolnej wartości i_{I^c} i każdego $t_{I^c} \in \times_{k\in I^c}(t_{i_k}^{(n)}, t_{i_k+1}^{(n)}]$,

$$\|Z_n\|_{\mathcal{J}} \leqslant \|h_n\|_{\mathcal{J}} \prod_{k\in I^c} |\tilde{N}^{(k)}_{t^{(n)}_{i_k+1}} - \tilde{N}^{(k)}_{t^{(n)}_{i_k}}|,$$

gdzie $\|h_n\|_{\mathcal{J}}$ po prawej stronie jest wartością tej funkcji w punkcie t_{I^c}. Zatem z lematu Fatou, Twierdzenia 16 oraz definicji h_n, otrzymujemy

$$\begin{aligned}
\mathbb{E}|Z|^p =&\mathbb{E}\liminf_n |Z_n|^p \leqslant \liminf_n \mathbb{E}|Z_n|^p \\
\leqslant& \liminf_n K_d^p \Big(\sum_{\mathcal{J}\in\mathcal{P}_{I_d}} p^{p\deg(\mathcal{J})/2}\|h_n\|_{\mathcal{J}}^p \\
&+ \sum_{I\subsetneq I_d}\sum_{\mathcal{J}\in\mathcal{P}_I} p^{p(\# I^c+\deg(\mathcal{J})/2)}\mathbb{E}_{I^c}\max_{\mathrm{i}_{I^c}} \Big\|\|h_n\|_{\mathcal{J}}\Big\|_\infty^p \prod_{k\in I^c} |\tilde{N}^{(k)}_{t^{(n)}_{i_k+1}} - \tilde{N}^{(k)}_{t^{(n)}_{i_k}}|^p\Big) \\
\leqslant& \tilde{K}_d^p \sum_{I\subseteq I_d}\sum_{\mathcal{J}\in\mathcal{P}_I} p^{p(\# I^c+\deg(\mathcal{J})/2)}\Big\|\|h\|_{\mathcal{J}}\Big\|_\infty^p \\
&\times \mathbb{E}_{I^c}\limsup_n \max_{\mathrm{i}_{I^c}} \prod_{k\in I^c} |\tilde{N}^{(k)}_{t^{(n)}_{i_k+1}} - \tilde{N}^{(k)}_{t^{(n)}_{i_k}}|^p \\
\leqslant& \tilde{K}_d^p \sum_{I\subseteq I_d}\sum_{\mathcal{J}\in\mathcal{P}_I} p^{\# I^c+\deg(\mathcal{J})/2}\Big\|\|h\|_{\mathcal{J}}\Big\|_\infty^p,
\end{aligned}$$

przy czym w dwóch ostatnich nierównościach użyliśmy założenia, że skoki $N^{(k)}$ są ograniczone przez 1 (wówczas lim sup jest także ograniczone przez 1, a ponadto procesy $N^{(k)}$ mają wszystkie momenty [15], co razem z nierównością Dooba i faktem, że $\tilde{N}^{(k)}$ to martyngały, pozwala dokonać przejścia granicznego). □

Część II

Prawo iterowanego logarytmu dla U-statystyk

Rozdział 4

Wstępne fakty

W niniejszym rozdziale przedstawimy wstępne intuicje związane z prawem iterowanego logarytmu dla U-statystyk oraz wykażemy równoważność prawa iterowanego logarytmu dla różnych wersji U-statystyk (zwyczajnych, uniezależnionych, dodatkowo zrandomizowanych). Wprowadzimy także klasyczne narzędzia, pomocne w dowodzie głównego wyniku tej części pracy, jakim jest charakterystyka prawa iterowanego logarytmu w terminach jedynie funkcji h oraz rozkładu zmiennych losowych generujących U-statystykę.

Od tej pory będziemy zajmować się U-statystyką opartą o jedną funkcję $h\colon \Sigma^d \to \mathbb{R}$ oraz $(X_i)_{i\in\mathbb{N}^*}$ – ciąg niezależnych zmiennych losowych o tym samym rozkładzie. Podobnie jak w poprzednich twierdzeniach, wygodniej będzie pracować z U-statystykami uniezależnionymi. Dlatego rozważać będziemy także $(X_i^{(j)})_{j\in I_d, i\in\mathbb{N}^*}$, macierz niezależnych kopii zmiennej X_1. Miejscami będziemy potrzebować ogólniejszej macierzy $(X_i^{(j)})_{j\in I_d, i\in\mathbb{N}^*}$ niezależnych zmiennych losowych, niekoniecznie o tym samym rozkładzie. Będziemy wtedy zaznaczać, że odpowiednie fakty dotyczą ogólniejszej sytuacji.

Zdefiniujmy także, dla uniknięcia problemów technicznych, $\mathrm{LL}\, x = \operatorname{loglog}(x \vee e^e)$.

Klasyczne prawo iterowanego logarytmu, pochodzące od Kołmogorowa, Hartmana, Wintnera i Strassena (zob. np. [39]) mówi, że w przypadku $d = 1$, jeśli $\mathbb{E}h(X_1) = 0$, $\mathbb{E}h(X_1)^2 < \infty$, to

$$\limsup_{n\to\infty} \frac{1}{\sqrt{2n\mathrm{LL}\, n}} \Big| \sum_{i=1}^{n} h(X_i) \Big| = \sqrt{\mathbb{E}h(X_1)^2} \text{ p.n.}$$

Ponadto, jeśli powyższe warunki na całkowalność h nie zachodzą, rozpatrywana granica górna jest nieskończona.

Fakt ten pozwala już na ustalenie naturalnego normowania w prawie iterowanego logarytmu dla U-statystyk. Rzeczywiście, jeśli np. $d = 2$, $h(x, y) = xy$, zaś X_i są rzeczywistymi scentrowanymi zmiennymi losowymi o skończonej wariancji mamy

$$\sum_{\substack{i,j\leqslant n\\ i\neq j}} h(X_i,X_j) = (\sum_{i=1}^n X_i)^2 - \sum_{i=1}^n X_i^2.$$

Zatem z prawa iterowanego logarytmu oraz prawa wielkich liczb dla sum zmiennych i.i.d., mamy

$$\limsup_{n\to\infty} \frac{1}{n\mathrm{LL}\,n}\Big|\sum_{\substack{i,j\leqslant n\\ i\neq j}} h(X_i,X_j)\Big| = 2\mathbb{E}X_1^2.$$

W niniejszej rozprawie będziemy się zajmować prawem iterowanego logarytmu dla U-statystyk, z normowaniem sugerowanym przez powyższą równość, tzn. $(n\mathrm{LL}\,n)^{d/2}$, gdzie d jest rzędem U-statystyki. Pokażemy w szczególności, że prawo iterowanego logarytmu z takim normowaniem, implikuje kanoniczność (całkowite zdegenerowanie) jądra h. W przypadku U-statystyk o innym stopniu zdegenerowania, mogą pojawić się inne normowania. Można to łatwo zaobserwować, analizując opisany poniżej rozkład Hoeffdinga U-statystyki, my nie będziemy jednak poruszać tych zagadnień.

Zanim przejdziemy do prezentacji nowych wyników, opiszmy dotychczasowy stan wiedzy, dot. prawa iterowanego logarytmu dla kanonicznych U-statystyk. Arcones i Giné [5], wykazali, że warunek $\mathbb{E}h(X_1,\ldots,X_d)^2<\infty$, implikuje

$$\limsup_{n\to\infty} \frac{1}{(2n\mathrm{LL}\,n)^{d/2}}\Big|\sum_{\mathrm{i}\in I_n^d} h(\mathbf{X}_{\mathrm{i}})\Big| = \sup\{|\mathbb{E}h(X_1,\ldots,X_d)g(X_1)\cdots g(X_d)|\colon \mathbb{E}g^2(X_1)\leqslant 1\}. \tag{4.1}$$

Następnie, Giné i Zhang [21] skonstruowali przykłady, pokazujące, że całkowalność z kwadratem funkcji generującej U-statystykę nie jest warunkiem koniecznym dla skończoności granicy w prawie iterowanego logarytmu.

Dla $d=2$, warunki konieczne i dostateczne zostały znalezione przez Giné, Kwapienia, Latałę i Zinna [24]. Okazuje się, że skończoność granicy górnej po lewej stronie (4.1) jest równoważna kanoniczności funkcji h oraz następującym warunkom:

$$\mathbb{E}(h^2(X_1,X_2)\wedge u)\leqslant C\mathrm{LL}\,u$$

i

$$\sup\{\mathbb{E}h(X_1,X_2)f(X_1)g(X_2)\colon \|f(X_1)\|_2,\|g(X_2)\|_2\leqslant 1, \|f\|_\infty,\|g\|_\infty<\infty\}\leqslant C$$

dla pewnej stałej $C<\infty$, przy czym granica górna szacuje się przez KC, gdzie K jest pewną stałą uniwersalną. Dokładna wartość granicy górnej nie jest w ogólności znana. Kwapień, Latała, Oleszkiewicz i Zinn [32] znaleźli jej postać w pewnych szczególnych przypadkach, gdy funkcja h nie jest całkowalna z kwadratem, ale ma odpowiednią blokową postać. W niniejszej pracy nie będziemy zajmować się problemem dokładnego wyznaczenia granicy, uogólnimy natomiast powyższe twierdzenie Giné, Kwapienia, Latały i Zinna na U-statystyki dowolnego rzędu. Prezentowane wyniki pochodzą ze wspólnej pracy autora i Rafała Latały [3], przy czym znaczna część dowodu została uproszczona.

4.1 Podstawowe narzędzia

4.1.1 Rozkład Hoeffdinga

Zaczniemy od przypomnienia klasycznego rozkładu U-statystyki, generowanej przez jądro o średniej 0 na sumę całkowicie zdegenerowanych U-statystyk różnych rzędów, wprowadzonego przez Hoeffdinga [27] i będącego jednym z podstawowych narzędzi w analizie U-statystyk. Zgodnie ze wstępnymi uwagami pracujemy z ustalonym ciągiem $X_1, X_2, \ldots$ zmiennych losowych i.i.d. oraz oznaczamy przez X, wektor $(X_1, \ldots, X_d)$.

Definicja 14 (Projekcje Hoeffdinga). *Dla funkcji $h\colon \Sigma^d \to \mathbb{R}$, spełniającej warunek $\mathbb{E}|h(X)| < \infty$, oraz $k = 0, 1, \ldots, d$, zdefiniujmy $\pi_k h\colon \Sigma^k \to \mathbb{R}$ wzorem*

$$\pi_k h(x_1, \ldots, x_k) = (\delta_{x_1} - \mathbf{P}) \otimes (\delta_{x_2} - \mathbf{P}) \otimes \ldots \otimes (\delta_{x_k} - \mathbf{P}) \otimes \mathbf{P}^{d-k} h,$$

gdzie $\mathbf{P}$ to rozkład zmiennej X_1, zaś dla miary μ i funkcji f, $\mu f := \int f d\mu$.

W szczególności $\pi_0 h = \mathbb{E}h, \pi_1 h(x_1) = \mathbb{E}_{\{2,\ldots,d\}} h(x_1, X_2, \ldots, X_d) - \mathbb{E}h$.

W dalszej części pracy będziemy potrzebować naturalnego odpowiednika powyższej definicji (dla $k = d$) dla uniezależnionej U-statystyki, opartej o ciąg zmiennych o niekoniecznie tym samym rozkładzie.

Definicja 15. *Niech $h\colon \Sigma_1 \times \ldots \times \Sigma_d \to \mathbb{R}$ będzie funkcją mierzalną. Rozważmy niezależne ciągi $(X_j^{(1)})_j, \ldots, (X_j^{(d)})_j$ zmiennych i.i.d. o wartościach odpowiednio w $\Sigma_1, \ldots, \Sigma_d$. Załóżmy, że $\mathbb{E}|h(X_1^{(1)}, \ldots, X_1^{(d)})| < \infty$. Zdefiniujmy $\pi_d h\colon \Sigma_1 \times \ldots \times \Sigma_d \to \mathbb{R}$ wzorem*

$$\pi_d h(x_1, \ldots, x_d) = (\delta_{x_1} - \mathbf{P}_{X_1^{(1)}}) \otimes \ldots \otimes (\delta_{x_d} - \mathbf{P}_{X_1^{(d)}}) h,$$

gdzie $\mathbf{P}_{X_1^{(i)}}$ jest rozkładem zmiennej $X_1^{(i)}$.

Uwaga Oczywiście, gdy $\Sigma_1 = \ldots = \Sigma_d$ oraz $(X_i^{(j)})_{i \in \mathbb{N}^*}$ są niezależnymi kopiami ciągu $(X_i)_{i \in \mathbb{N}^*}$, obie definicje $\pi_d h$ pokrywają się.

Bezpośrednio z definicji wynika, że w przypadku i.i.d., dla $k \geqslant 1$, $\pi_k h$ jest kanoniczna względem rozkładu X_1. Ponadto zachodzi równość $\pi_0 h = \mathbb{E}h$.

Udowodnimy teraz prosty fakt, dotyczący porównywania momentów dla U-statystyk.

LEMAT 21. *Rozważmy dowolną d-wskaźnikową macierz całkowalnych funkcji $h_{\mathrm{i}}\colon \Sigma_1 \times \ldots \times \Sigma_d \to \mathbb{R}$, $|\mathrm{i}| \leqslant n$. Dla dowolnej macierzy $(X_i^{(j)})_{i,j}$ jak w Definicji 15 oraz $p \geqslant 1$ zachodzi*

$$\Big\| \sum_{|\mathrm{i}|\leqslant n} \pi_d h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big\|_p \leqslant 2^d \Big\| \sum_{|\mathrm{i}|\leqslant n} \epsilon_{\mathrm{i}}^{\mathrm{dec}} h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big\|_p,$$

gdzie $(\varepsilon_i^{(j)})_{i,j}$ jest macierzą niezależnych zmiennych Rademachera, niezależnych od $(X_i^{(j)})_{i,j}$.

DOWÓD. Dla $d = 1$, teza jest klasyczną nierównością symetryzacyjną dla sum niezależnych zmiennych losowych (np. [39], Lemma 6.3). Dowód ogólnego przypadku przeprowadzimy przez indukcję ze względu na d. Dla uproszczenia notacji, oznaczmy przez $\bar{\pi}_{d-1} h_{\mathrm{i}}$ odpowiednią projekcję Hoeffdinga funkcji h_{i} traktowanej jako funkcja zmiennych $x_2, \ldots, x_d$, przy ustalonym pierwszym argumencie. Dokładniej, niech

$$\bar{\pi}_{d-1} h_{\mathrm{i}}(x) = \delta_{x_1} \otimes (\delta_{x_2} - \mathbf{P}_{X_1^{(2)}}) \otimes \ldots \otimes (\delta_{x_d} - \mathbf{P}_{X_1^{(d)}}) h_{\mathrm{i}}.$$

Załóżmy zatem, że lemat zachodzi dla wszystkich U-statystyk stopnia mniejszego od d. Rozważmy $(\tilde{X}_i^{(k)})_{i\in\mathbb{N}^*, k\leqslant d}$, niezależną kopię $(X_i^{(k)})_{i\in\mathbb{N}^*, k\leqslant d}$. Oznaczmy przez $\tilde{\mathbb{E}}_1$ wartość oczekiwaną ze względu na $\tilde{X}^{(1)}$. Wówczas, kanoniczność $\pi_d h_{\mathrm{i}}$ oraz nierówność Jensena, implikują

$$\begin{aligned}
&\mathbb{E}_1 \Big| \sum_{|\mathrm{i}|\leqslant n} \pi_d h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big|^p \\
&= \mathbb{E}_1 \Big| \sum_{|\mathrm{i}|\leqslant n} (\pi_d h_{\mathrm{i}}(X_{i_1}^{(1)}, \ldots, X_{i_d}^{(d)}) - \tilde{\mathbb{E}}_1 \pi_d h_{\mathrm{i}}(\tilde{X}_{i_1}^{(1)}, X_{i_2}^{(2)}, \ldots, X_{i_d}^{(d)}))\Big|^p \\
&\leqslant \mathbb{E}_1 \tilde{\mathbb{E}}_1 \Big| \sum_{|\mathrm{i}|\leqslant n} (\pi_d h_{\mathrm{i}}(X_{i_1}^{(1)}, \ldots, X_{i_d}^{(d)}) - \pi_d h_{\mathrm{i}}(\tilde{X}_{i_1}^{(1)}, X_{i_2}^{(2)}, \ldots, X_{i_d}^{(d)}))\Big|^p \\
&= \mathbb{E}_1 \tilde{\mathbb{E}}_1 \Big| \sum_{|\mathrm{i}|\leqslant n} \varepsilon_{i_1}^{(1)} (\pi_d h_{\mathrm{i}}(X_{i_1}^{(1)}, \ldots, X_{i_d}^{(d)}) - \pi_d h_{\mathrm{i}}(\tilde{X}_{i_1}^{(1)}, X_{i_2}^{(2)}, \ldots, X_{i_d}^{(d)}))\Big|^p \\
&= \mathbb{E}_1 \tilde{\mathbb{E}}_1 \Big| \sum_{|\mathrm{i}|\leqslant n} \varepsilon_{i_1}^{(1)} \Big(\bar{\pi}_{d-1} h_{\mathrm{i}}(X_{i_1}^{(1)}, \ldots, X_{i_d}^{(d)}) \\
&\qquad\qquad - \bar{\pi}_{d-1} h_{\mathrm{i}}(\tilde{X}_{i_1}^{(1)}, X_{i_2}^{(2)}, \ldots, X_{i_d}^{(d)})\Big)\Big|^p.
\end{aligned}$$

Z nierówności trójkąta, otrzymujemy

$$\Big\| \sum_{|\mathrm{i}|\leqslant n} \pi_d h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big\|_p \leqslant 2 \Big\| \sum_{|\mathrm{i}|\leqslant n} \varepsilon_{i_1}^{(1)} \bar{\pi}_{d-1} h_{\mathrm{i}}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big\|_p.$$

Twierdzenie Fubiniego oraz założenie indukcyjne zastosowane do macierzy funkcji $\tilde{h}_{(i_2,\ldots,i_d)}(x_2,\ldots,x_d) = \sum_{i_1 \leqslant n} \varepsilon_{i_1}^{(1)} h_{\mathrm{i}}(X_{i_1}^{(1)}, x_2, \ldots, x_d)$ dla ustalonych wartości zmiennych $(X_i^{(1)})_i, (\varepsilon_i^{(1)})_i$, kończy dowód. $\square$

Będziemy także potrzebować wspomnianego już twierdzenia Hoeffdinga, dającego rozkład U-statystyki na sumę nieskorelowanych, kanonicznych U-statystyk różnych rzędów.

Lemat 22 (Rozkład Hoeffdinga, np. [12], str. 137). *Niech $h\colon \Sigma^d \to \mathbb{R}$, będzie niezmiennicza ze względu na permutację argumentów. Zdefiniujmy*

$$U_n(h) = \frac{(n-d)!}{n!} \sum_{\mathrm{i} \in I_n^d} h(\mathbf{X}_{\mathrm{i}}).$$

Wówczas

$$U_n(h) = \sum_{k=0}^{d} \binom{d}{k} U_n(\pi_k h).$$

4.1.2 Dolne oszacowania odchyleń dla chaosów rademacherowych

Zaprezentujemy teraz nieco inne niż w Rozdziale 1 nierówności dla chaosów generowanych przez niezależne zmienne Rademachera. Oszacowania te posłużą nam w dalszej części pracy do znalezienia warunków koniecznych na prawo iterowanego logarytmu dla U-statystyk.

Definicja 16. *Dla d-wskaźnikowej macierzy $(a_{\mathrm{i}})_{|\mathrm{i}| \leqslant n}$ oraz dla każdego rozbicia $\mathcal{J} = \{J_1, \ldots, J_m\} \in \mathcal{P}_{I_d}$, zdefiniujmy*

$$\|(a_{\mathrm{i}})\|^*_{\mathcal{J},p} := \sup\Big\{ \Big|\sum_{|\mathrm{i}| \leqslant n} a_{\mathrm{i}} \prod_{k=1}^{m} \alpha^{(k)}_{\mathrm{i}_{J_k}}\Big| \colon \sum_{\mathrm{i}_{J_k}} (\alpha^{(k)}_{\mathrm{i}_{J_k}})^2 \leqslant p, \\ \forall_{i_{\max J_k} \in I_n} \sum_{\mathrm{i}_{\diamond J_k}} (\alpha^{(k)}_{\mathrm{i}_{J_k}})^2 \leqslant 1,\ k = 1, \ldots, m \Big\},$$

gdzie $\diamond J_k = J_k \backslash \{\max J_k\}$ (przyjmujemy konwencję $\sum_{\mathrm{i}_\emptyset} a_{\mathrm{i}} = a_{\mathrm{i}}$).

Uwaga Powyższa definicja jest niesymetryczna, gdyż wyróżniamy dla każdego zbioru J_k, indeks $i_{\max J_k}$. Czynimy tak tylko dla ustalenia uwagi i moglibyśmy oczywiście wyróżnić dowolny inny indeks i_j, gdzie $j \in J_k$.

Przykład Dla $d = 1$, powyższa definicja sprowadza się do wzoru

$$\|(a_i)_{i\leqslant n}\|^*_{\{1\},p} = \sup\Big\{\Big|\sum_{i=1}^n a_i\alpha_i\Big|\colon\ \sum_{i=1}^n \alpha_i^2 \leqslant p, |\alpha_i| \leqslant 1,\ i=1,\ldots,n\Big\}.$$

Z kolei dla $d = 2$, otrzymujemy

$$\begin{aligned}\|(a_{ij})_{i,j\leqslant n}\|^*_{\{1\},\{2\},p} &= \sup\Big\{\Big|\sum_{i,j=1}^n a_{ij}\alpha_i\beta_j\Big|\colon\ \sum_{i=1}^n \alpha_i^2 \leqslant p, \sum_{j=1}^n \beta_j^2 \leqslant p,\\ &\qquad \forall_{i\in I_n}|\alpha_i| \leqslant 1, \forall_{j\in I_n}|\beta_j| \leqslant 1\Big\},\\ \|(a_{ij})_{i,j\leqslant n}\|^*_{\{1,2\},p} &= \sup\Big\{\Big|\sum_{i,j=1}^n a_{ij}\alpha_{ij}\Big|\colon\ \sum_{i,j=1}^n \alpha_{ij}^2 \leqslant p, \forall_{j\in I_n}\sum_{i=1}^n \alpha_{ij}^2 \leqslant 1\Big\}.\end{aligned}$$

Lemat 23. *Niech $(a_{\mathrm{i}})_{|\mathrm{i}|\leqslant n}$ będzie d-wskaźnikową macierzą liczb rzeczywistych. Rozważmy zmienną losową*

$$S := |\sum_{|\mathrm{i}|\leqslant n} a_{\mathrm{i}} \prod_{k=1}^d \varepsilon^{(k)}_{i_k}| = |\sum_{|\mathrm{i}|\leqslant n} a_{\mathrm{i}}\epsilon^{\mathrm{dec}}_{\mathrm{i}}|.$$

Wówczas, dla $p \geqslant 1$,

$$\|S\|_p \geqslant \frac{1}{L_d}\sum_{\mathcal{J}\in\mathcal{P}_{I_d}} \|(a_{\mathrm{i}})\|^*_{\mathcal{J},p}.$$

W szczególności dla pewnej stałej c_d,

$$\mathbb{P}(S \geqslant c_d \sum_{\mathcal{J}\in\mathcal{P}_{I_d}} \|(a_{\mathrm{i}})\|^*_{\mathcal{J},p}) \geqslant c_d \wedge e^{-p}.$$

Dowód. Druga część lematu wynika z pierwszej analogicznie jak we Wniosku 2. Aby udowodnić oszacowania momentów, użyjemy indukcji ze względu na d. Dla $d = 1$ odpowiednia nierówność została udowodniona przez Hitczenkę [26], dla $d = 2$ przez Latałę [35] (jako część precyzyjnych oszacowań dwustronnych). Załóżmy zatem, że oszacowania momentów są prawdziwe dla chaosów stopnia mniejszego niż $d \geqslant 3$.

Rozważmy najpierw rozbicie $\mathcal{J} = \{I_d\}$. Mamy

$$\begin{aligned}
\mathbb{E}S^p &= \mathbb{E}_d \mathbb{E}_{I_{d-1}} \Big| \sum_{i_d} \varepsilon_{i_d}^{(d)} \Big| \sum_{\mathrm{i}_{I_{d-1}}} a_{\mathrm{i}} \prod_{k=1}^{d-1} \varepsilon_{i_k}^{(k)} \Big|\Big|^p \\
&\geqslant \mathbb{E}_d \Big| \sum_{i_d} \varepsilon_{i_d}^{(d)} \mathbb{E}_{I_{d-1}} \big| \sum_{\mathrm{i}_{I_{d-1}}} a_{\mathrm{i}} \prod_{k=1}^{d-1} \varepsilon_{i_k}^{(k)} \big| \Big|^p \\
&\geqslant \frac{1}{\tilde{L}_d^p} \mathbb{E}_d \Big| \sum_{i_d} \varepsilon_{i_d}^{(d)} (\sum_{\mathrm{i}_{I_{d-1}}} a_{\mathrm{i}}^2)^{1/2} \Big|^p \\
&\geqslant \frac{1}{\tilde{L}_d^p L_1^p} \sup \Big\{ \sum_{i_d} \alpha_{i_d} (\sum_{\mathrm{i}_{I_{d-1}}} a_{\mathrm{i}}^2)^{1/2} \colon \sum_{i_d} \alpha_{i_d}^2 \leqslant p, |\alpha_i| \leqslant 1 \Big\}^p,
\end{aligned}$$

gdzie pierwsza nierówność wynika z nierówności Jensena, druga z własności hiperkontracji dla chaosów rademacherowych (zob. [12], Theorem 3.2.5, p. 114) oraz zasady kontrakcji dla średnich rademacherowych (zob. np. [39], Theorem 4.4., str. 95), trzecia wreszcie z założenia indukcyjnego.

Pozostaje pokazać, że

$$\sup \Big\{ \sum_{i_d} \alpha_{i_d} (\sum_{\mathrm{i}_{I_{d-1}}} a_{\mathrm{i}}^2)^{1/2} \colon \sum_{i_d} \alpha_{i_d}^2 \leqslant p, |\alpha_i| \leqslant 1 \Big\} \geqslant \|(a_{\mathrm{i}})\|_{\{I_d\},p}^*.$$

Niech zatem (γ_{i}) będzie d-wskaźnikową macierzą taką, że $\sum_{\mathrm{i}} \gamma_{\mathrm{i}}^2 \leqslant p$, $\sum_{\mathrm{i}_{I_{d-1}}} \gamma_{\mathrm{i}}^2 \leqslant 1$ dla wszystkich i_d. Wówczas

$$\begin{aligned}
|\sum_{\mathrm{i}} \gamma_{\mathrm{i}} a_{\mathrm{i}}| &\leqslant \sum_{\mathrm{i}} |\gamma_{\mathrm{i}}||a_{\mathrm{i}}| \leqslant \sum_{i_d} (\sum_{\mathrm{i}_{I_{d-1}}} \gamma_{\mathrm{i}}^2)^{1/2} (\sum_{\mathrm{i}_{I_{d-1}}} a_{\mathrm{i}}^2)^{1/2} \\
&\leqslant \sup \Big\{ \sum_{i_d} \alpha_{i_d} (\sum_{\mathrm{i}_{I_{d-1}}} a_{\mathrm{i}}^2)^{1/2} \colon \sum_{i_d} \alpha_{i_d}^2 \leqslant p, |\alpha_i| \leqslant 1 \Big\}.
\end{aligned}$$

Niech teraz $\mathcal{J} = \{J_1, \ldots, J_m\}$, $m \geqslant 2$. Z założenia indukcyjnego i nierówności Jensena

$$\begin{aligned}
\|S\|_p &\geqslant \frac{1}{L_{d-\#J_1}} \Big(\mathbb{E}_{J_1} \Big(\|(\sum_{\mathrm{i}_{J_1}} a_{\mathrm{i}} \prod_{k \in J_1} \varepsilon_{i_k}^{(k)})_{\mathrm{i}_{I_d \setminus J_1}} \|_{\mathcal{J} \setminus \{J_1\},p}^* \Big)^p \Big)^{1/p} \\
&\geqslant \frac{1}{L_{d-\#J_1} L_{\#J_1}} \|(a_{\mathrm{i}})\|_{\mathcal{J},p}^*.
\end{aligned}$$

□

4.1.3 Oszacowania odchyleń U-statystyk

Przedstawimy teraz oszacowania odchyleń dla U-statystyk, będące prostą konsekwencją Twierdzenia 16 oraz Lematu 21.

Wniosek 8. *Niech $h\colon \Sigma^d \to \mathbb{R}$ będzie dowolną (niekoniecznie kanoniczną) funkcją ograniczoną. Wówczas dla każdego $t \geqslant 0$,*

$$\mathbb{P}\Big(\Big|\sum_{|\mathrm{i}|\leqslant n}\pi_d h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| \geqslant t\Big)$$
$$\leqslant L_d \exp\Big[-\frac{1}{L_d}\Big(\min_{\mathcal{J}\in\mathcal{P}_{I_d}}\Big(\frac{t}{n^{d/2}\|h\|_{\mathcal{J}}}\Big)^{2/deg(\mathcal{J})}\Big) \wedge \Big(\min_{I\subsetneq I_d}\Big(\frac{t}{n^{\#I/2}\|(\mathbb{E}_I h^2)^{1/2}\|_\infty}\Big)^{2/(d+\#I^c)}\Big)\Big].$$

Dowód. Korzystamy z Twierdzenia 16, aby oszacowań momenty U-statystyki

$$\sum_{|\mathrm{i}|\leqslant n}\epsilon_{\mathrm{i}}^{\mathrm{dec}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}),$$

które z Lematu 21 szacują momenty U-statystyki z tezy wniosku. Ponieważ nasze zmienne mają identyczny rozkład, uwaga przed Wnioskiem 7 pozwala wyrazić normy $\|\cdot\|_{\mathcal{J}}$ macierzy funkcji, poprzez n i normy pojedynczej funkcji h (zauważmy, że normy $\|\cdot\|_{\mathcal{J}}$ macierzy $(\epsilon_{\mathrm{i}}^{\mathrm{dec}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}))_{|\mathrm{i}|\leqslant n}$ są takie same jak macierzy $(h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}))_{|\mathrm{i}|\leqslant n}$). Zatem, z podanego na początku rozprawy Twierdzenia 3, dającego związek między oszacowaniami momentów i odchyleń, otrzymujemy oszacowanie na U-statystykę generowaną przez $\pi_d h$, jak we Wniosku 7. Pozostaje oszacować dla $I \subsetneq I_d$, normy $\|h\|_{\mathcal{J}}$ przez odpowiednie drugie momenty. □

4.1.4 Wnioski z całkowalności funkcji h

Przypomnijmy, że zgodnie z rozdziałem Ogólne założenia i notacja, w sytuacji gdy pracujemy z ciągiem $X_1, X_2, \ldots$ zmiennych i.i.d. oraz U-statystyką rzędu d, przez X oznaczamy wektor $(X_1, \ldots, X_d)$. Niniejsza sekcja poświęcona jest konsekwencjom warunku

$$\mathbb{E}(h(X)^2 \wedge u) = \mathcal{O}((\operatorname{loglog} u)^{d-1}),$$

który, jak udowodnili Giné i Zhang [21], jest konieczny dla prawa iterowanego logarytmu dla U-statystyk rzędu d.

Lemat 24. *Jeśli $\mathbb{E}(h(X)^2 \wedge u) = \mathcal{O}((\operatorname{loglog} u)^{d-1})$ to dla $I \subsetneq I_d$ oraz $a > 0$*

$$\sum_{l=0}^{\infty}\sum_{n=3}^{\infty} 2^{l+\#I^c n}\mathbb{P}_{I^c}(\mathbb{E}_I(h(X)^2 \wedge 2^{an}) \geqslant 2^{2l+\#I^c n}\log^d n) < \infty.$$

W konsekwencji

$$\sum_n 2^{\#I^c n}(\log n)^{-k}\mathbb{P}_{I^c}(\mathbb{E}_I(h(X)^2 \wedge 2^{2nd}) \geqslant 2^{\#I^c n}(\log n)^{d-k}) < \infty$$

dla wszystkich $k \geqslant 0$.

Dowód. Dla ustalonych l oraz k, mamy

$$\begin{aligned}
&\sum_{2^k<\log n\leqslant 2^{k+1}} 2^{l+\#I^c n}\mathbb{P}_{I^c}(\mathbb{E}_I(h(X)^2\wedge 2^{an})\geqslant 2^{2l+\#I^c n}\log^d n)\\
&\leqslant \sum_{2^k<\log n\leqslant 2^{k+1}} 2^{l+\#I^c n}\mathbb{P}_{I^c}\Big(\mathbb{E}_I\Big(h(X)^2\wedge 2^{ae^{2^{k+1}}}\Big)\geqslant 2^{2l+\#I^c n+dk}\Big)\\
&\leqslant 2^l\mathbb{E}_{I^c}\sum_n 2^{\#I^c n}\mathbf{1}_{\{\mathbb{E}_I(h(X)^2\wedge 2^{ae^{2^{k+1}}})\geqslant 2^{2l+\#I^c n+dk}\}}\\
&\leqslant 2^l\mathbb{E}_{I^c}\Big[2\frac{\mathbb{E}_I(h(X)^2\wedge 2^{ae^{2^{k+1}}})}{2^{2l+dk}}\Big]\leqslant 2^{-l}K\frac{(\log ae^{2^{k+1}})^{d-1}}{2^{dk}}\\
&\leqslant 2^{-l}\tilde{K}\Big(\frac{\log^{d-1}a}{2^{dk}}+2^{-k}\Big),
\end{aligned}$$

gdzie $\tilde{K}$ zależy tylko od h, co po wysumowaniu względem $l\geqslant 0$ i $k\geqslant 0$, dowodzi pierwszej części lematu. Druga nierówność może być otrzymana przez przybliżoną zamianę zmiennych $2^{\#I^c m}(\log m)^{-k}\simeq 2^{\#I^c n}$ i użycie zbieżności wewnętrznej sumy dla $l=0$ w pierwszej nierówności, przy $a>2d$.

□

Lemat 25. *Jeśli $\xi\geqslant 0$ jest zmienną losową taką, że $\mathbb{E}(\xi^2\wedge u)=\mathcal{O}((\operatorname{loglog} u)^\beta)$, to*

$$\mathbb{E}\xi\mathbf{1}_{\{\xi\geqslant s\}}=\mathcal{O}(\frac{(\operatorname{loglog} s)^\beta}{s}).$$

Dowód. Rzeczywiście, dla $u\geqslant e^e$, mamy $\mathbb{P}(\xi\geqslant u)\leqslant K\frac{(\operatorname{loglog} u)^\beta}{u^2}$, a zatem dla odpowiednio dużych s,

$$\begin{aligned}
\mathbb{E}\xi\mathbf{1}_{\{\xi\geqslant s\}}&=\sum_{k=0}^\infty\mathbb{E}\xi\mathbf{1}_{\{2^ks\leqslant\xi<2^{k+1}s\}}\leqslant K\sum_{k=0}^\infty 2^{k+1}s\frac{(\operatorname{loglog} 2^ks)^\beta}{2^{2k}s^2}\\
&\leqslant 2K\sum_{k=0}^\infty\frac{(\operatorname{loglog} 2^ks)^\beta}{2^ks}=\mathcal{O}(\frac{(\operatorname{loglog} s)^\beta}{s}).
\end{aligned}$$

□

Lemat 26. *Jeśli $\mathbb{E}(h(X)^2\wedge u)=\mathcal{O}((\operatorname{loglog} u)^\beta)$ to*

$$\mathbb{E}\frac{|h(X)|^2}{(\mathrm{LL}\,|h(X)|)^{\beta+\varepsilon}}<\infty$$

Dla każdego $\varepsilon>0$.

Dowód. Dla dużych n

$$\begin{aligned}
\mathbb{E}\frac{|h(X)|^2}{(\mathrm{LL}\,|h(X)|)^{\beta+\varepsilon}}\mathbf{1}_{\{2^{2^{2^n}}\leqslant|h(X)|<2^{2^{2^{n+1}}}\}}&\leqslant K\frac{\mathbb{E}(|h(X)|^2\wedge 2^{2^{2^{n+1}}})}{2^{n(\beta+\varepsilon)}}\leqslant\tilde{K}\frac{2^{(n+1)\beta}}{2^{n(\beta+\varepsilon)}}\\
&=\tilde{K}2^\beta 2^{-n\varepsilon}.
\end{aligned}$$

$\square$

LEMAT 27. *Załóżmy, że* $h\colon \Sigma^d \to \mathbb{R}$ *spełnia* $\mathbb{E}(h(X)^2 \wedge u) = \mathcal{O}((\operatorname{loglog} u)^{d-1})$. *Niech dla* $I \subsetneq I_d$

$$h_{I,n} = h\mathbf{1}_{A_n^I},$$

gdzie

$$A_n^I = \{x\colon h(x)^2 \leqslant 2^{nd}\log^d n,\ 2^{\#I^c n}(\log n)^{-2d} \leqslant \mathbb{E}_I(h^2 \wedge 2^{2nd}) \leqslant 2^{\#I^c n}(\log n)^d\}$$

Wówczas

$$\sum_n \Big(\frac{\mathbb{E}h_{I,n}(X)^2}{\log^d n}\Big)^2 < \infty.$$

DOWÓD. Zauważmy, że

$$\begin{aligned}
\mathbb{E}h_{I,n}^2 &\leqslant \mathbb{E}(h^2 \wedge 2^{2nd})\mathbf{1}_{\{(\log n)^{-2d} \leqslant 2^{-\#I^c n}\mathbb{E}_I(h^2\wedge 2^{2nd}) \leqslant (\log n)^d\}} \\
&\leqslant \sum_{k=1}^{3d} \mathbb{E}_{I^c}\mathbb{E}_I(h^2 \wedge 2^{2nd})\mathbf{1}_{\{(\log n)^{d-k} \leqslant 2^{-\#I^c n}\mathbb{E}_I(h^2\wedge 2^{2nd}) \leqslant (\log n)^{d+1-k}\}} \\
&\leqslant \sum_{k=1}^{3d} 2^{\#I^c n}(\log n)^{d+1-k}\mathbb{P}_{I^c}(\mathbb{E}_I(h^2 \wedge 2^{2nd}) \geqslant 2^{\#I^c n}(\log n)^{d-k}).
\end{aligned}$$

Zatem

$$\begin{aligned}
\sum_n \frac{(\mathbb{E}h_{I,n}^2)^2}{(\log n)^{2d}} &\leqslant K\sum_n \frac{\mathbb{E}h_{I,n}^2}{(\log n)^{d+1}} \\
&\leqslant K\sum_{k=1}^{3d}\sum_n \frac{2^{\#I^c n}}{(\log n)^k}\mathbb{P}_{I^c}(\mathbb{E}_I(h^2 \wedge 2^{2nd}) \geqslant 2^{\#I^c n}(\log n)^{d-k}) \\
&< \infty
\end{aligned}$$

z Lematu 24. $\square$

4.2 Równoważność różnych sformułowań prawa iterowanego logarytmu

W dalszej kolejności przedstawimy kilka ogólnych faktów, dotyczących równoważności różnych wersji prawa iterowanego logarytmu. Większość przytoczonych tu faktów jest rozwinięciem pomysłów z prac [21, 18].

Podstawową wersją prawa iterowanego logarytmu, którą będziemy rozważać, jest

$$\limsup_{n\to\infty} \frac{1}{(n\mathrm{LL}\,n)^{d/2}}\Big|\sum_{\mathrm{i}\in I_n^d} h(\mathbf{X}_{\mathrm{i}})\Big| < \infty \text{ p.n.}$$

Zauważmy, że z prawa zerojedynkowego Hewitta-Savage'a wynika, że powyższa granica górna jest z prawdopodobieństwem 1 stała. Zatem, prawo iterowanego logarytmu może być równoważnie zapisane jako

$$\limsup_{n\to\infty} \frac{1}{(n\mathrm{LL}\, n)^{d/2}} \Big| \sum_{\mathrm{i}\in I_n^d} h(\mathbf{X}_{\mathrm{i}}) \Big| \leqslant C \text{ p.n.}, \tag{4.2}$$

dla pewnej stałej $C < \infty$.

Prezentację podstawowych faktów dotyczących prawa iterowanego logarytmu zaczniemy od następującego lematu, pochodzącego od Giné i Zhanga [21]. Dla pełności rozprawy, lemat przytaczamy wraz z pochodzącym z tej samej pracy dowodem.

Lemat 28. *Niech $h\colon \Sigma^d \to \mathbb{R}$ będzie funkcją symetryczną. Wówczas istnieją stałe L_d takie, że jeśli*

$$\limsup_{n\to\infty} \frac{1}{(n\mathrm{LL}\, n)^{d/2}} \Big| \sum_{\mathrm{i}\in I_n^d} h(\mathbf{X}_{\mathrm{i}}) \Big| < C \text{ p.n.},$$

to

$$\sum_{n=1}^{\infty} \mathbb{P}\Big(\Big| \sum_{|\mathrm{i}|\leqslant 2^n} \epsilon_{\mathrm{i}}^{\mathrm{dec}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}) \Big| > D 2^{nd/2} \log^{d/2} n \Big) < \infty \tag{4.3}$$

dla $D = L_d C$.

Dowód. Zdefiniujmy $I = \{\mathrm{i} = (i_1, \ldots, i_d) \in (\mathbb{N}^*)^d \colon \forall_{j\neq k, j,k\in I_d}\ i_j \neq i_k\}$.

Dla skończonego zbioru $A \subseteq \mathbb{N}^*$, oznaczmy $S_A = \sum_{\mathrm{i}\in I\cap A^d} h(\mathbf{X}_{\mathrm{i}})$ i zauważmy, że jeśli $A_0, A_1, \ldots, A_d$ są parami rozłącznymi, skończonymi podzbiorami $\mathbb{N}^*$, to

$$d! \sum_{\mathrm{i}\in A_1\times\ldots\times A_d} h(\mathbf{X}_{\mathrm{i}}) = \sum_{r=0}^{d} (-1)^{d-r} \sum_{1\leqslant s_1<\ldots<s_r\leqslant d} S_{A_0\cup A_{s_1}\cup\ldots\cup A_{s_r}}, \tag{4.4}$$

co wynika z równości

$$\sum_{r=k}^{d} (-1)^{d-r} \binom{d-k}{r-k} = \delta_{kd}.$$

Dla $n \in \mathbb{N}$ oraz $k = 1, \ldots, d$, zdefiniujmy zbiory

$$J_{n,k} = [(2^n-1)d + 2^n(k-1) + 1, (2^n-1)d + 2^n k] \cap \mathbb{N}$$

i dla każdego $k = 1, \ldots, d$, przenumerujmy ciąg $(X_i)_{i\in\bigcup_{j=1}^{\infty} J_{j,k}}$, oznaczając go w naturalnej kolejności przez $(X_i^{(k)})_{i\in\mathbb{N}^*}$. Innymi słowy, jeśli $i = d(2^n-1)+2^n(k-1)+t$, $0 < t \leqslant 2^n$, to $X_i = X_{2^n-1+t}^{(k)}$. Wówczas ciągi $(X_i^{(k)})_{i\in\mathbb{N}^*}$ są niezależnymi kopiami ciągu $(X_i)_{i\in\mathbb{N}^*}$. Możemy zatem napisać $\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}} = (X_{i_1}^{(1)}, \ldots, X_{i_d}^{(d)})$. Rozważmy także $(\varepsilon_i)_{i\in\mathbb{N}^*}$ – ciąg niezależnych zmiennych Rademachera, niezależny od $(X_i)_{i\in\mathbb{N}^*}$ oraz uniezależnioną tablicę $(\varepsilon_i^{(k)})_{i\in\mathbb{N}^*,k\in I_d}$, otrzymaną z $(\varepsilon_i)_i$ w analogiczny sposób, jak $(X_i^{(k)})_{i,k}$ z $(X_i)_i$.

Zauważmy, że dla $n \in \mathbb{N}$, zmienna losowa

$$\sum_{2^n \leqslant i_1,\ldots,i_d \leqslant 2^{n+1}-1} \epsilon_{\mathrm{i}}^{\mathrm{dec}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})$$

jest sumą 2^d wyrażeń, postaci

$$\pm \sum_{\mathrm{i}\in A_{n,1}\times\ldots\times A_{n,d}} h(\mathbf{X}_{\mathrm{i}}),$$

gdzie dla $k = 1,\ldots,d$, $A_{n,k} = \{i \in J_{n,k} \colon \varepsilon_i = 1\}$ lub $A_{n,k} = \{i \in J_{n,k} \colon \varepsilon_i = -1\}$.

Dla ustalonej wartości ciągu $(\varepsilon_i)_i$, rozważmy teraz dowolną z 2^d rodzin zbiorów $(A_{n,k})_{n\in\mathbb{N},k\in I_d}$ (zadanych przez niezależny od n wybór znaków w powyższej równości, definiującej zbiory $A_{n,k}$ dla $k = 1,\ldots,d$). Niech $A_{n,0} = (0,(2^n-1)d] \cap \mathbb{N}$.

Wracając do (4.4), zauważmy, że dla dowolnego ciągu zbiorów $B_n \subseteq B_{n+1}$, takiego, że $B_n \subseteq (0,(2^{n+1}-1)d] \cap \mathbb{N}$,

$$\limsup_{n\to\infty} \frac{1}{2^{(n+1)d/2}\log^{d/2}(n+1)} |S_{B_n}| < Cd^{d/2} \text{ p.n.},$$

gdyż $(X_i)_{i\in\bigcup_j B_j}$ ma taki sam rozkład jak $(X_i)_{i\in\mathbb{N}^*}$. Zatem, korzystając z inkluzji $A_{n,k} \subseteq J_{n,k} \subseteq ((2^n-1)d,(2^{n+1}-1)d]$ dla $k = 1,\ldots,d$ i stosując powyższy fakt do $B_j = A_{n,0} \cup A_{n,s_1} \cup \ldots \cup A_{n,s_r}$ przy dowolnych ustalonych $1 \leqslant s_1 < \ldots < s_r \leqslant d$ i korzystając z (4.4), otrzymujemy

$$\limsup_{n\to\infty} \frac{1}{2^{d(n+1)/2}\log^{d/2}(n+1)} \Big| \sum_{\mathrm{i}\in A_{n,1}\times\ldots\times A_{n,d}} h(\mathbf{X}_{\mathrm{i}}) \Big| < 2^d d^{d/2} C \text{ p.n.},$$

warunkowo na zmienne Rademachera. Zatem, z twierdzenia Fubiniego,

$$\limsup_{n\to\infty} \frac{1}{2^{nd/2}\log^{d/2} n} \Big| \sum_{2^n \leqslant i_1,\ldots,i_d \leqslant 2^{n+1}-1} \epsilon_{\mathrm{i}}^{\mathrm{dec}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}) \Big| < L_d C \text{ p.n.}$$

Zauważmy, że zmienne losowe $\left|\sum_{2^n \leqslant i_1,\ldots,i_d \leqslant 2^{n+1}-1} \epsilon_{\mathrm{i}}^{\mathrm{dec}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\right|$, $n \in \mathbb{N}^*$, są niezależne. Zatem, powyższa nierówność wraz z lematem Borela-Cantellego implikują, że

$$\sum_{n=1}^{\infty} \mathbb{P}\Big(\Big| \sum_{2^n \leqslant i_1,\ldots,i_d \leqslant 2^{n+1}-1} \epsilon_{\mathrm{i}}^{\mathrm{dec}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}) \Big| > L_d C 2^{nd/2} \log^{d/2} n \Big) < \infty,$$

co jest równoważne tezie. $\square$

LEMAT 29. *Jeśli dla pewnego $D < \infty$,*

$$\lim_{n\to\infty} \mathbb{P}\Big(\Big|\sum_{|\mathrm{i}|\leqslant 2^n} \epsilon_{\mathrm{i}}^{\mathrm{dec}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| > D2^{nd/2}\log^{d/2} n\Big) = 0,$$

to

$$\limsup_{u\to\infty} \frac{\mathbb{E}(h(X)^2 \wedge u)}{\operatorname{loglog}^d u} < \infty.$$

DOWÓD. Zauważmy, że z nierówności Paleya-Zygmunda oraz własności hiperkontrakcji dla chaosu rademacherowego (np. [12], Theorem 3.2.2), otrzymujemy dla dowolnego n

$$\mathbb{P}_\varepsilon\bigg(\Big|\sum_{|\mathrm{i}|\leqslant n} \epsilon_{\mathrm{i}}^{\mathrm{dec}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| \geqslant L_d^{-1}\big(\sum_{|\mathrm{i}|\leqslant n} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})^2\big)^{1/2}\bigg) \geqslant \frac{1}{L_d}. \tag{4.5}$$

Ponadto, jeśli $\mathbb{E}(h(X)^2 \wedge n) \geqslant 1$, to

$$\begin{aligned}
\mathbb{E}\Big(\sum_{|\mathrm{i}|\leqslant n} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})^2 \wedge n\Big)^2 &= \sum_{I\subseteq I_d}\sum_{|\mathrm{i}|\leqslant n}\sum_{\substack{|\mathrm{j}|\leqslant n\\ \{k:\, i_k=j_k\}=I}} \mathbb{E}([h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})^2\wedge n][h(\mathbf{X}_{\mathrm{j}}^{\mathrm{dec}})^2\wedge n])\\
&\leqslant n^{2d}[\mathbb{E}(h(X)^2\wedge n)]^2 + \sum_{I\subseteq I_d, I\neq\emptyset} n^{d+\#I^c} n\mathbb{E}(h(X)^2\wedge n)\\
&\leqslant n^{2d}[\mathbb{E}(h(X)^2\wedge n)]^2 + (2^d-1)n^{2d}\mathbb{E}(h(X)^2\wedge n)\\
&\leqslant 2^d n^{2d}[\mathbb{E}(h(X)^2\wedge n)]^2 = 2^d\big(\mathbb{E}\big(\sum_{|\mathrm{i}|\leqslant n} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})^2\wedge n\big)\big)^2.
\end{aligned}$$

Zatem, znów z nierówności Paleya-Zygmunda

$$\mathbb{P}\bigg(\sum_{|\mathrm{i}|\leqslant n} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})^2 \geqslant \frac{1}{2} n^d \mathbb{E}(h(X)^2\wedge n)\bigg) \geqslant \frac{1}{L_d},$$

co razem z (4.5) daje

$$\mathbb{P}\bigg(\Big|\sum_{|\mathrm{i}|\leqslant n} \epsilon_{\mathrm{i}}^{\mathrm{dec}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| \geqslant \tilde{L}_d^{-1} n^{d/2}\sqrt{\mathbb{E}(h(X)^2\wedge n)}\bigg) \geqslant \frac{1}{\tilde{L}_d}. \tag{4.6}$$

Z drugiej strony z założenia, dla odpowiednio dużych n,

$$\mathbb{P}\bigg(\Big|\sum_{|\mathrm{i}|\leqslant 2^n} \epsilon_{\mathrm{i}}^{\mathrm{dec}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| > D2^{nd/2}\log^{d/2} n\bigg) < \frac{1}{\tilde{L}_d}.$$

Łącząc powyższą nierówność oraz (4.6) z n zastąpionym przez 2^n, dostajemy dla dużych n,

$$\mathbb{E}(h(X)^2 \wedge 2^n) \leqslant D^2 \tilde{L}_d^2 \log^d n,$$

co już jest równoważne tezie. □

Łącząc powyższy fakt z Lematem 28, otrzymujemy

WNIOSEK 9. *Prawo iterowanego logarytmu (4.2) implikuje, że* $\mathbb{E}(h^2(X) \wedge u) = \mathcal{O}((\text{loglog}\, u)^d)$.

Uwaga Jak już wspomnieliśmy, Giné i Zhang [21] wykazali, że prawo iterowanego logarytmu (4.2) implikuje silniejszą całkowalność, tj. $\mathbb{E}(h(X)^2 \wedge u) = \mathcal{O}((\text{loglog}\, u)^{d-1})$. W dalszej części pracy, rozpatrując warunki konieczne i dostateczne dla (4.2) podamy nowy dowód tego faktu, sprowadzając go do jednego z przedstawionych warunków. Dowód będzie jednak korzystał z prostszego oszacowania, danego w powyższym Wniosku. Okazuje się też, że dla wykazania równoważności różnych form prawa iterowanego logarytmu, powyższe, osłabione oszacowanie jest wystarczające.

LEMAT 30. *Dla symetrycznej funkcji* $h\colon \Sigma^d \to \mathbb{R}$, *prawo iterowanego logarytmu (4.2) jest równoważne z* uniezależnionym *prawem iterowanego logarytmu*

$$\limsup_{n\to\infty} \frac{1}{(n\text{LL}\, n)^{d/2}} \Big| \sum_{\mathrm{i}\in I_n^d} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}) \Big| \leqslant D \text{ p.n.} \tag{4.7}$$

Dokładniej, (4.2) implikuje (4.7) z $D = L_d C$ *oraz odwrotnie, z (4.7) wynika (4.2) dla* $C = L_d D$.

DOWÓD. Warunek (4.2) może być równoważnie zapisany jako

$$\lim_{k\to\infty} \mathbb{P}\Big(\sup_{n\geqslant k} \frac{1}{(n\text{LL}\, n)^{d/2}} \Big| \sum_{\mathrm{i}\in I_n^d} h(\mathbf{X}_{\mathrm{i}}) \Big| \geqslant C + \varepsilon \Big) = 0,$$

dla wszystkich $\varepsilon > 0$, co może być z kolei sformułowane w postaci

$$\lim_{k\to\infty} \mathbb{P}\Big(\Big\| \sum_{\substack{|\mathrm{i}|<\infty \\ l\neq j \Rightarrow i_l \neq i_j}} h_{|\mathrm{i}|,k}(\mathbf{X}_{\mathrm{i}}) \Big\|_\infty \geqslant C + \varepsilon \Big) = 0, \tag{4.8}$$

gdzie dla $i, k \in \mathbb{N}$, $h_{i,k}$ jest funkcją o wartośiach w ℓ_∞, zdefiniowaną wzorem

$$h_{i,k} = \Big(\frac{h}{(k\text{LL}\, k)^{d/2}}, \frac{h}{((k+1)\text{LL}\,(k+1))^{d/2}}, \ldots, \frac{h}{(n\text{LL}\, n)^{d/2}}, \ldots, \Big)$$

dla $i \leqslant k$ oraz

$$h_{i,k} =$$

$$\Big(\underbrace{0, \ldots, 0}_{i-k}, \frac{h}{(i\text{LL}\, i)^{d/2}}, \frac{h}{((i+1)\text{LL}\,(i+1))^{d/2}}, \ldots, \frac{h}{(n\text{LL}\, n)^{d/2}}, \ldots \Big)$$

w przeciwnym przypadku. Łatwo zauważyć, że uniezależniona wersji warunku (4.8) jest w analogiczny sposób równoważna (4.7). Zatem, gdybyśmy mogli skorzystać z twierdzenia o uniezależnianiu, teza zostałaby dowiedziona. Pojawiają się jednak dwa techniczne problemy, powodujące, że nie możemy się na to twierdzenie powołać bezpośrednio. Mianowicie, przestrzeń ℓ_∞ nie jest ośrodkowa, a dodatkowo zbieżność wielokrotnego szeregu w warunku (4.8) nie jest zbieżnością w normie, a jedynie po współrzędnych. Są to jednak trudności pozorne, gdyż na dowolnej ustalonej współrzędnej mamy tylko skończoną U-statystykę, co pozwala nam zastosować twierdzenie o uniezależnianiu do rzutu na pierwsze m współrzędnych (czyli w ośrodkowej przestrzeni ℓ_∞^m), a następnie przejść z m do nieskończoności, aby dostać tezę. □

Lemat 31. *Istnieje stała uniwersalna $L < \infty$ taka, że dla każdej funkcji $h\colon \Sigma_1 \times \ldots \times \Sigma_d \to B$, gdzie $(B, \|\cdot\|)$ jest przestrzenią Banacha oraz macierzy $(X_i^{(j)})$, jak w Definicji 15,*

$$\mathbb{P}\Big(\max_{|\mathrm{j}|\leqslant n}\Big\|\sum_{\mathrm{i}\colon i_k\leqslant j_k, k=1\ldots d} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big\| \geqslant t\Big) \leqslant L^d\mathbb{P}\Big(\Big\|\sum_{|\mathrm{i}|\leqslant n} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big\| \geqslant t/L^d\Big).$$

Dowód. Tezę udowodnimy indukcyjnie. Dla $d = 1$, jest to wynik Montgomery'ego-Smitha [41]. Załóżmy zatem, że teza zachodzi dla U-statystyk rzędu mniejszego niż d i rozważmy funkcję $h\colon \Sigma_1 \times \ldots \times \Sigma_d \to B$. Wówczas, warunkując na $X^{(d)}$, stosując założenie indukcyjne do przestrzeni $\ell_\infty^n(B)$ oraz funkcji $g(x_1, \ldots, x_{d-1}) = (\sum_{i_d\leqslant l} h(x_1, \ldots, x_{d-1}, X_{i_d}^{(d)})\colon l \leqslant n)$, a następnie używając twierdzenia Fubiniego, otrzymujemy

$$\begin{aligned}
&\mathbb{P}(\max_{|\mathrm{j}|\leqslant n}\Big\|\sum_{\mathrm{i}\colon i_k\leqslant j_k, k=1\ldots d} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big\| \geqslant t\Big) \\
&\leqslant L^{d-1}\mathbb{P}(\max_{j\leqslant n}\Big\|\sum_{|\mathrm{i}|\leqslant n\colon i_d\leqslant j} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big\| \geqslant t/L^{d-1}\Big).
\end{aligned}$$

Aby zakończyć dowód, wystarczy ponownie zastosować wynik Montgomery'ego-Smitha, tym razem warunkowo na $(X^{(1)}, \ldots, X^{(d-1)})$. □

Wniosek 10. *Rozważmy funkcję $h\colon \Sigma_1 \times \ldots \times \Sigma_d \to \mathbb{R}$, macierz zmiennych $(X_i^{(j)})$, jak w Definicji 15 oraz $\alpha > 0$. Jeśli*

$$\sum_{n=1}^{\infty}\mathbb{P}(|\sum_{|\mathrm{i}|\leqslant 2^n} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})| \geqslant C2^{n\alpha}\log^\alpha n) < \infty,$$

to

$$\limsup_{n\to\infty}\frac{1}{(n\mathrm{LL}\, n)^\alpha}\Big|\sum_{|\mathrm{i}|\leqslant n} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| \leqslant L_{d,\alpha}C \text{ p.n.}$$

Dowód. Dla $0 < D < \infty$ mamy

$$\mathbb{P}\Big(\sup_{n\geqslant N}\frac{1}{(n\mathrm{LL}\,n)^{\alpha}}\Big|\sum_{|\mathrm{i}|\leqslant n}h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| > D\Big)$$

$$\leqslant \mathbb{P}\Big(\sup_{k>\lfloor \log N/\log 2\rfloor}\max_{2^{k-1}\leqslant n\leqslant 2^k}\frac{L_\alpha}{(2^k\log k)^{\alpha}}\Big|\sum_{|\mathrm{i}|\leqslant n}h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| > D\Big)$$

$$\leqslant \sum_{k>\lfloor \log N/\log 2\rfloor}\mathbb{P}\Big(\max_{2^{k-1}\leqslant n\leqslant 2^k}\frac{L_\alpha}{(2^k\log k)^{\alpha}}\Big|\sum_{|\mathrm{i}|\leqslant n}h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| > D\Big),$$

więc teza wynika z Lematu 31. □

Następny lemat pokazuje, że przy pewnych założeniach, dot. całkowalności, wkład w U-statystykę uniezależnioną, pochodzący od „przekątnej", czyli indeksów $\mathrm{i} \notin I_n^d$, jest zaniedbywalny. Pozwoli to, w zależności od potrzeb, dodawać lub usuwać „przekątną" w prawie iterowanego logarytmu.

Lemat 32. *Jeśli $h\colon \Sigma^d \to \mathbb{R}$ jest kanoniczna i spełnia*

$$\mathbb{E}(h(X)^2\wedge u) = \mathcal{O}((\mathrm{loglog}\,u)^{\beta}),$$

dla pewnego β, to

$$\limsup_{n\to\infty}\frac{1}{(n\mathrm{LL}\,n)^{d/2}}\Big|\sum_{\substack{|\mathrm{i}|\leqslant n\\ \exists_{j\neq k} i_j=i_k}}h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| = 0 \text{ p.n.} \tag{4.9}$$

Dowód. Podzielmy sumę po „przekątnej" na części, zależnie od „poziomic" indeksu i. Dla podziału $\mathcal{J}\in\mathcal{P}_{I_d}$ niech $A_{\mathcal{J}}(n)$ będzie zbiorem tych $|\mathrm{i}|\leqslant n$, dla których i jest stałe na J dla każdego $J\in\mathcal{J}$ (i traktujemy jako funkcję określoną na I_d). Zauważmy, że suma

$$U_{\mathcal{J}}(n) := \sum_{\mathrm{i}\in A_{\mathcal{J}}(n)}h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}),$$

może być traktowana jako uniezależniona U-statystyka rzędu $\deg\mathcal{J}$, o ile potraktujemy zmienne $\mathbf{X}_{\mathrm{i}_J}^{\mathbf{dec}}$ dla każdego $J\in\mathcal{J}$ jako jedną nową zmienną. Oznaczmy ponadto dla $j<k$, $j,k\in I_d$, $A_{jk} = \{\mathrm{i}\colon |\mathrm{i}|\leqslant n, i_j = i_k\}$ oraz $\Delta = \{(j,k)\subseteq I_d^2\colon j<k\}$. Z wzoru włączeń i wyłączeń otrzymujemy dla dowolnego $|\mathrm{i}|\leqslant n$

$$\mathbf{1}_{\{\exists_{j\neq k}i_j=i_k\}} = \mathbf{1}_{\bigcup_{(j,k)\in\Delta}A_{jk}} = \sum_{l=1}^{\binom{d}{2}}\sum_{\substack{(j_1,k_1),\ldots,(j_l,k_l)\in\Delta\\ \forall_{r\neq s}(j_r,k_r)\neq(j_s,k_s)}}(-1)^{l-1}\mathbf{1}_{A_{j_1k_1}\cap\ldots\cap A_{j_lk_l}}.$$

Dostajemy stąd

$$\sum_{\substack{|\mathrm{i}|\leqslant n\\ \exists_{j\neq k}i_j=i_k}}h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}) = \sum_{\substack{\mathcal{J}\in\mathcal{P}_{I_d}\\ \deg\mathcal{J}<d}}a_{\mathcal{J}}U_{\mathcal{J}}(n),$$

dla pewnych liczb $a_{\mathcal{J}}$, co do modułu nie przekraczających stałej, zależnej tylko od d. Ponieważ liczba składników po prawej stronie również nie zależy od n, wystarczy pokazać, że

$$\limsup_{n\to\infty} \frac{|U_{\mathcal{J}}(n)|}{(n\mathrm{LL}\, n)^{d/2}} = 0$$

dla dowolnego $\mathcal{J}$ takiego, że $\deg \mathcal{J} < d$. Na mocy Wniosku 10, wystarczy wykazać, że

$$\sum_{n=1}^{\infty} \mathbb{P}\Big(\Big| \sum_{\mathrm{i}\in A_{\mathcal{J}}(2^n)} \pi_{\deg \mathcal{J}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| \geqslant C 2^{nd/2} \log^{d/2} n\Big) < \infty \tag{4.10}$$

dla dowolnego $C > 0$ (tutaj $\pi_{\deg \mathcal{J}}$ oznacza projekcję Hoeffdinga jądra h jako funkcji $\deg \mathcal{J}$ zmiennych, jak wyjaśniono powyżej. Mamy zatem $\pi_{\deg \mathcal{J}} h = h$). Nierówność (4.10) nie jest trudna do udowodnienia, gdyż liczba składników w U-statystyce jest rzędu istotnie mniejszego niż 2^{nd}. Oczywiście $\#A_{\mathcal{J}}(2^n) = 2^{n \deg \mathcal{J}} \leqslant 2^{n(d-1)}$. Niech I będzie dowolnym podzbiorem I_d takim, że dla każdego $J \in \mathcal{J}$, $\#(I \cap J) = 1$. Stosując do $h_n = h\mathbf{1}_{\{|h| > 2^{nd}\}}$ Lemat 25, otrzymujemy

$$\mathbb{E}\Big| \sum_{\mathrm{i}\in A_{\mathcal{J}}(2^n)} \epsilon_{\mathrm{i}_I}^{\mathrm{dec}} h_n(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| \leqslant 2^{n(d-1)} \mathbb{E}|h_n| \leqslant K 2^{n(d-1)} \frac{\log^{\beta} n}{2^{nd}} = K \frac{\log^{\beta} n}{2^n},$$

zatem zbieżność (4.10) z h zastąpionym przez h_n wynika z Lematu 21 oraz nierówności Czebyszewa. Z drugiej strony, dla $\tilde{h}_n = h\mathbf{1}_{\{|h| < 2^{nd}\}}$,

$$\begin{aligned} \frac{\mathbb{E}\Big| \sum_{\mathrm{i}\in A_{\mathcal{J}}(2^n)} \epsilon_{\mathrm{i}_I}^{\mathrm{dec}} \tilde{h}_n(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big|^2}{C^2 2^{nd} \log^d n} &\leqslant \frac{\#A_{\mathcal{J}}(2^n) \mathbb{E}\tilde{h}_n^2}{C^2 2^{nd} \log^d n} \leqslant \frac{2^{n(d-1)} \mathbb{E}\tilde{h}_n^2}{C^2 2^{nd} \log^d n} \\ &\leqslant K C^{-2} 2^{-n} \log^{\beta - d} n, \end{aligned}$$

co (ponownie z Lematu 21 i nierówności Czebyszewa) pozwala zakończyć dowód. □

Wniosek 11. *Dla symetrycznej, kanonicznej funkcji $h\colon \Sigma^d \to \mathbb{R}$, prawo iterowanego logarytmu (4.2) jest równoważne uniezależnionemu prawu iterowanego logarytmu „z przekątną”*

$$\limsup_{n\to\infty} \frac{1}{(n\mathrm{LL}\, n)^{d/2}} \Big| \sum_{|\mathrm{i}| \leqslant n} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| \leqslant D, \tag{4.11}$$

tzn. istnieją stałe L_d takie, że (4.2) zachodzi dla pewnego D, to zachodzi także (4.11) dla $D = L_d C$, i odwrotnie, (4.11) implikuje (4.2) dla $C = L_d D$.

Dowód. Aby wykazać, że (4.2) implikuje (4.11), wystarczy użyć Lematu 30, a następnie Lematu 32, aby dodać „przekątną” (odpowiednia całkowalność h wynika z Wniosku 9).

Aby otrzymać drugą implikację, wystarczy udowodnić warunek $\mathbb{E}(h(X)^2 \wedge u) = \mathcal{O}((\operatorname{loglog} u)^d)$, gdyż wówczas możemy skasować przekątną (Lemat 32), a następnie zastosować Lemat 30, aby uzyskać (4.2).

Z założenia wynika, że dla dowolnego $\varepsilon > 0$ oraz odpowiednio dużych n

$$\mathbb{P}(\Big(\Big|\sum_{|\mathrm{i}|\leqslant n} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| > Dn^{d/2}\operatorname{loglog}^{d/2} n\Big) < \varepsilon.$$

Zauważmy, że z Lematu 31, dla dowolnych zbiorów $A_1, \ldots, A_d \subseteq I_n$

$$\mathbb{P}\Big(\Big|\sum_{\mathrm{i}\in A_1\times\ldots\times A_d} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| > L^d Dn^{d/2}\operatorname{loglog}^{d/2} n\Big) \leqslant L^d\varepsilon.$$

Ponadto, dla ustalonych wartości zmiennych $(\varepsilon_i^{(j)})$, wyrażenie$\sum_{|\mathrm{i}|\leqslant n} \epsilon_{\mathrm{i}}^{\mathrm{dec}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})$ jest sumą 2^d wyrażeń postaci $\pm\sum_{\mathrm{i}\in A_1\times\ldots\times A_d} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})$, gdzie $A_k = \{i\colon \varepsilon_i^{(k)} = \pm 1\}$. Zatem, stosując warunkowo powyższe oszacowanie oraz twierdzenie Fubiniego, dostajemy dla dostatecznie dużych n

$$\mathbb{P}\Big(\Big|\sum_{|\mathrm{i}|\leqslant n} \epsilon_{\mathrm{i}}^{\mathrm{dec}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| > 2^d L^d Dn^{d/2}\operatorname{loglog}^{d/2} n\Big) \leqslant 2^d L^d \varepsilon.$$

Żądana całkowalność wynika zatem z Lematu 29 □

WNIOSEK 12. *Zrandomizowane i uniezależnione prawo iterowanego logarytmu*

$$\limsup_{n\to\infty} \frac{1}{(n\operatorname{loglog} n)^{d/2}}\Big|\sum_{|\mathrm{i}|\leqslant n} \epsilon_{\mathrm{i}}^{\mathrm{dec}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| < C \tag{4.12}$$

jest równoważne (4.3). Dokładniej jeśli zachodzi (4.12), to prawdziwe jest również (4.3) z $D = L_d C$ oraz (4.3) implikuje (4.12) z $C = L_d D$.

DOWÓD. Implikacja (4.3) $\Rightarrow$ (4.12) wynika z Wniosku 10. Implikacja (4.12) $\Rightarrow$ (4.3) wynika z Wniosku 11 oraz Lematu 28 (zauważmy, że jeśli $(\varepsilon_i)_i, (\varepsilon_i^{(j)})_{i,j}$, są niezależnymi zmiennymi Rademachera, to są nimi również $(\varepsilon_i\varepsilon_i^{(j)})_{i,j}$). □

Rozdział 5

Prawo iterowanego logarytmu dla rzeczywistych U-statystyk

Jak pokazuje wynik Giné, Kwapienia, Latały i Zinna, już dla U-statystyk rzędu 2, warunki konieczne i dostateczne na prawo iterowanego logarytmu są dużo bardziej skomplikowane niż dla sum niezależnych zmiennych losowych. W przypadku ogólnym, sytuacja jest jeszcze bardziej złożona i aby scharakteryzować prawo iterowanego logarytmu, wygodnie będzie wprowadzić nową notację.

Przypomnijmy, że przez X oznaczamy wektor $(X_1, X_2, \ldots, X_d)$, zaś dla $J \subseteq I_d$, $X_J = (X_i)_{i \in J}$.

Definicja 17. *Dla funkcji $h\colon \Sigma^d \to \mathbb{R}$, podziału $\mathcal{J} = \{J_1, \ldots, J_k\} \in \mathcal{P}_{I_d}$ oraz dodatniej liczby u, zdefiniujmy*

$$\|h\|_{\mathcal{J},u} = \|h(X)\|_{\mathcal{J},u} = \sup\Big\{\mathbb{E}[h(X)\prod_{i=1}^{k} f_i(X_{J_i})]\colon \\ \|f_i(X_{J_i})\|_2 \leqslant 1, \|f_i(X_{J_i})\|_\infty \leqslant u,\ i = 1, \ldots, k\Big\}.$$

Przykład Dla $d = 3$, powyższa definicja daje

$$\begin{aligned}
\|h(X_1,X_2,X_3)\|_{\{1,2,3\},u} &= \sup\{\mathbb{E}h(X_1,X_2,X_3)f(X_1,X_2,X_3)\colon \\
&\qquad \mathbb{E}f(X_1,X_2,X_3)^2 \leqslant 1, \|f\|_\infty \leqslant u\},\\
\|h(X_1,X_2,X_3)\|_{\{1,2\}\{3\},u} &= \sup\{\mathbb{E}h(X_1,X_2,X_3)f(X_1,X_2)g(X_3)\colon \\
&\qquad \mathbb{E}f(X_1,X_2)^2, \mathbb{E}g(X_3)^2 \leqslant 1\\
&\qquad \|f\|_\infty, \|g\|_\infty \leqslant u\},\\
\|h(X_1,X_2,X_3)\|_{\{1\}\{2\}\{3\},u} &= \sup\{\mathbb{E}h(X_1,X_2,X_3)f(X_1)g(X_2)k(X_3)\colon \\
&\qquad \mathbb{E}f(X_1)^2, \mathbb{E}g(X_2)^2, \mathbb{E}k(X_3)^2 \leqslant 1\\
&\qquad \|f\|_\infty, \|g\|_\infty, \|k\|_\infty \leqslant u\}.
\end{aligned}$$

Mamy więc do czynienia ze sparametryzowaną rodziną norm. Okazuje się, że warunki konieczne i dostateczne na prawo iterowanego logarytmu, związane są ze wzrostem funkcji $u \mapsto \|h\|_{\mathcal{J},u}$. Zanim przejdziemy do sytuacji ogólnej, zauważmy, że dla $d = 1$, warunek $\mathbb{E}h(X)^2 < \infty$, może być wyrażony jako ograniczoność funkcji $u \mapsto \|h\|_{\{1\},u}$. Podobnie dla $d = 2$, $\|h\|_{L^2\to L^2} < \infty$ jest równoważne ograniczoności$\|h\|_{\{1\}\{2\},u}$. Poniższy prosty lemat pokazuje, że podobna sytuacja ma miejsce w przypadku drugiego z warunków podanych w twierdzeniu Giné, Kwapienia, Latały i Zinna.

Lemat 33. *Dla dowolnej funkcji* $h\colon \Sigma^d \to \mathbb{R}$

$$\limsup_{u\to\infty} \frac{(\mathbb{E}(h(X)^2 \wedge u))^{1/2}}{(\operatorname{loglog} u)^{(d-1)/2}} = \limsup_{u\to\infty} \frac{\|h\|_{\{I_d\},u}}{(\operatorname{loglog} u)^{(d-1)/2}}.$$

Dowód. Oznaczmy granice górne po lewej i prawej stronie odpowiednio przez a oraz b. Załóżmy także bez straty ogólności, że $h \geqslant 0$. Wykażemy najpierw, że $a \leqslant b$. Rzeczywiście, albo $\mathbb{E}(h^2\wedge u) \leqslant 1$, albo możemy użyć jako funkcji testowej w definicji $\|h\|_{\{I_d\},u}$ funkcji

$$f := \frac{h \wedge \sqrt{u}}{(\mathbb{E}(h^2 \wedge u))^{1/2}},$$

otrzymując dla $u \geqslant 1$,

$$\|h\|_{\{I_d\},u} \geqslant \mathbb{E}hf = \frac{\mathbb{E}(h^2 \wedge \sqrt{u}h)}{(\mathbb{E}(h^2 \wedge u))^{1/2}} \geqslant (\mathbb{E}(h^2 \wedge u))^{1/2}.$$

Mamy zatem $(\mathbb{E}(h^2 \wedge u))^{1/2} \leqslant 1 + \|h\|_{\{I_d\},u}$, co natychmiast implikuje nierówność $a \leqslant b$. Aby udowodnić nierówność w drugą stronę, zauważmy, że jeśli $a < \infty$, to dla u dostatecznie dużych i dowolnej funkcji f takiej, że $\|f\|_2 \leqslant 1, \|f\|_\infty \leqslant u$, Lemat 25 implikuje

$$\mathbb{E}hf \leqslant \sqrt{\mathbb{E}h^2\mathbf{1}_{\{h\leqslant u^2\}}} + u\mathbb{E}|h|\mathbf{1}_{\{h\geqslant u^2\}} \leqslant (\mathbb{E}(h^2 \wedge u^4))^{1/2} + u\frac{K(\operatorname{loglog} u^2)^{d-1}}{u^2},$$

skąd już wynika, że $b \leqslant a$, gdyż

$$\lim_{u\to\infty} \frac{\operatorname{loglog} u^4}{\operatorname{loglog} u} = 1.$$

□

5.1 Prawo iterowanego logarytmu dla uniezależnionych U-statystyk

Sformułujemy teraz pierwszy z głównych wyników tej części pracy, charakteryzujący prawo iterowanego logarytmu dla kanonicznych i uniezależnionych U-statystyk.

TWIERDZENIE 22. *Niech h będzie kanoniczną symetryczną funkcją d zmiennych. Wtedy uniezależnione prawo iterowanego logarytmu*

$$\limsup_{n\to\infty} \frac{1}{n^{d/2}(\log\log n)^{d/2}}\Big|\sum_{|\mathrm{i}|\leqslant n} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| \leqslant C < \infty \text{ p.n.} \tag{5.1}$$

zachodzi wtedy i tylko wtedy, gdy dla każdego rozbicia $\mathcal{J} \in \mathcal{P}_{I_d}$

$$\limsup_{u\to\infty} \frac{1}{(\log\log u)^{(d-\deg \mathcal{J})/2}} \|h\|_{\mathcal{J},u} \leqslant D < \infty. \tag{5.2}$$

Dokładniej, jeśli zachodzi (5.1), to (5.2) jest spełnione dla $D = L_d C$ i odwrotnie, (5.2) implikuje (5.1) z $C = L_d D$, gdzie L_d jest stałą zależną tylko od d.

Uwaga Zauważmy, że zgodnie z Lematem 33, warunek (5.2) dla $\mathcal{J} = \{I_d\}$, jest równoważny podwyższonej całkowalności $\mathbb{E}(h(X)^2 \wedge u) \leqslant L_d C^2 (\log\log u)^{d-1}$. Dowodząc powyższego twierdzenia, wykazujemy więc także warunek konieczny na prawo iterowanego logarytmu, pochodzący od Giné i Zhanga, zgodnie z zapowiedzią z uwagi po Wniosku 9.

DOWÓD KONIECZNOŚCI. Udowodnimy najpierw następujący lemat.

LEMAT 34. *Niech $g\colon \Sigma^d \to \mathbb{R}$ będzie funkcją całkowalną z kwadratem (względem rozkładu wektora X). Wówczas*

$$\mathrm{Var}\Big(\sum_{|\mathrm{i}|\leqslant n} g(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big) \leqslant (2^d - 1) n^{2d-1} \mathbb{E} g(X)^2.$$

DOWÓD. Mamy

$$\begin{aligned}
&\mathrm{Var}\Big(\sum_{|\mathrm{i}|\leqslant n} g(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big) = \mathbb{E}\Big(\sum_{|\mathrm{i}|\leqslant n} (g(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}) - \mathbb{E} g(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}))\Big)^2 \\
&= \sum_{I\subseteq I_d} \sum_{|\mathrm{i}|\leqslant n} \sum_{\substack{|\mathrm{j}|\leqslant n:\\ \{k\colon i_k = j_k\} = I}} \mathbb{E}[(g(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}) - \mathbb{E} g(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}))(g(\mathbf{X}_{\mathrm{j}}^{\mathbf{dec}}) - \mathbb{E} g(\mathbf{X}_{\mathrm{j}}^{\mathbf{dec}}))] \\
&= \sum_{I\subseteq I_d, I\neq\emptyset} \sum_{|\mathrm{i}|\leqslant n} \sum_{\substack{|\mathrm{j}|\leqslant n:\\ \{k\colon i_k = j_k\} = I}} \mathbb{E}[(g(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}) - \mathbb{E} g(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}))(g(\mathbf{X}_{\mathrm{j}}^{\mathbf{dec}}) - \mathbb{E} g(\mathbf{X}_{\mathrm{j}}^{\mathbf{dec}}))] \\
&\leqslant \sum_{I\subseteq I_d, I\neq\emptyset} n^d n^{d-\#I} \mathrm{Var}(g(X)) \leqslant (2^d - 1) n^{2d-1} \mathbb{E} g(X)^2.
\end{aligned}$$

$\square$

Przechodząc do dowodu (5.2), zauważmy najpierw, że wystarczy jeśli wykażemy tezę dla $C > 0$. Z Wniosku 11 i Lematu 28, szereg (4.3) jest zbieżny. Ponadto z Wniosku 9,

$$\mathbb{E}(h(X)^2 \wedge u) \leqslant K(\operatorname{loglog} u)^d. \tag{5.3}$$

Ponieważ $\lim_{n\to\infty} \sum_{k=n}^{2n} \frac{1}{k} = \log 2$, istnieje N_0 takie, że dla wszystkich $N > N_0$, istnieje $N \leqslant n \leqslant 2N$, spełniające

$$\mathbb{P}(|\sum_{|\mathrm{i}|\leqslant 2^n} \epsilon_{\mathrm{i}}^{\mathrm{dec}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})| > L_d C 2^{nd/2} \log^{d/2} n) < \frac{1}{10n}. \tag{5.4}$$

Rozważmy dowolne $N > N_0$ i liczbę n, spełniającą powyższy warunek. Niech $\mathcal{J} = \{J_1, \ldots, J_k\} \in \mathcal{P}_{I_d}$. Ustalmy także funkcje $f_j \colon \Sigma^{\# J_j} \to \mathbb{R}$, $j = 1, \ldots, k$ takie, że

$$\|f_j(X_{J_j})\|_2 \leqslant 1,$$
$$\|f_j(X_{J_j})\|_\infty \leqslant 2^{n/(2k+1)}.$$

Z nierówności Czebyszewa mamy

$$\mathbb{P}(\sum_{|\mathrm{i}_{J_j}|\leqslant 2^n} f_j(\mathbf{X}_{\mathrm{i}_{J_j}}^{\mathrm{dec}})^2 \log n \leqslant 10 \cdot 2^d 2^{\# J_j n} \log n) \geqslant 1 - \frac{1}{10 \cdot 2^d}. \tag{5.5}$$

Ponadto dla odpowiednio dużych N,

$$\begin{aligned}\sum_{|\mathrm{i}_{\diamond J_j}|\leqslant 2^n} \frac{1}{2^{n\# J_j}} f_j(\mathbf{X}_{\mathrm{i}_{J_j}}^{\mathrm{dec}})^2 \cdot \log n &\leqslant \frac{2^{n\#\diamond J_j} 2^{2n/(2k+1)} \log n}{2^{n\# J_j}} \\ &\leqslant \frac{2^{2n/(2k+1)} \log n}{2^n} \leqslant 1.\end{aligned}$$

Bez straty ogólności możemy założyć, że ciągi $(X_i^{(j)})_{i,j}$ oraz $(\varepsilon_i^{(j)})_{i,j}$ są zdefiniowane jako współrzędne na produktowej przestrzeni probabilistycznej. Jeśli dla każdego $j = 1, \ldots, k$ oznaczymy zbiór z (5.5) przez A_j, mamy $\mathbb{P}(\bigcap_{j=1}^k A_j) \geqslant 0.9$. Użyjemy teraz Lematu 23. Na zbiorze $\bigcap_{j=1}^k A_j$ możemy oszacować normy $\|\cdot\|_{\mathcal{J},\log n}^*$ macierzy $(h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}))_{|\mathrm{i}|\leqslant 2^n}$, używając jako ciągów testowych

$$\alpha_{\mathrm{i}_{J_j}} = \frac{f_j(\mathbf{X}_{\mathrm{i}_{J_j}}^{\mathrm{dec}})\sqrt{\log n}}{10^{1/2} 2^{d/2} 2^{n\# J_j/2}}.$$

Zatem, z prawdopodobieństwem przynajmniej 0.9 mamy

$$\|(h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}))_{|\mathrm{i}|\leqslant 2^n}\|^*_{\mathcal{J},\log n} \geqslant \frac{(\log n)^{k/2}}{2^{dk/2}10^{k/2}2^{(\sum_j \#J_j)n/2}} \Big| \sum_{|\mathrm{i}|\leqslant 2^n} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}) \prod_{j=1}^k f_j(\mathbf{X}_{\mathrm{i}_{J_j}}^{\mathrm{dec}})\Big|$$
$$= \frac{(\log n)^{k/2}}{2^{dk/2}10^{k/2}2^{dn/2}} \Big| \sum_{|\mathrm{i}|\leqslant 2^n} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}) \prod_{j=1}^k f_j(\mathbf{X}_{\mathrm{i}_{J_j}}^{\mathrm{dec}})\Big|. \tag{5.6}$$

Naszym celem jest znalezienie dalszego ograniczenia z dołu na prawą stronę powyższej nierówności, aby poprzez Lemat 23, kontrolować odchylenie zmiennej $\sum_{|\mathrm{i}|\leqslant 2^n} \epsilon_{\mathrm{i}}^{\mathrm{dec}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})$, pod warunkiem $(X_i^{(j)})$.

Aby kontynuować, załóżmy, że

$$|\mathbb{E}h(X) \prod_{j=1}^k f_j(X_{J_j})| > 1. \tag{5.7}$$

Nierówność Markowa i Lemat 25 dają

$$\mathbb{P}\Big(\Big| \sum_{|\mathrm{i}|\leqslant 2^n} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}) \mathbf{1}_{\{|h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})|>2^n\}} \prod_{j=1}^k f_j(\mathbf{X}_{\mathrm{i}_{J_j}}^{\mathrm{dec}})\Big| \geqslant \frac{2^{nd}|\mathbb{E}h\prod_{j=1}^k f_j|}{4}\Big)$$
$$\leqslant 4\frac{2^{nd}(\prod_{j=1}^k \|f_j\|_\infty)\cdot \mathbb{E}|h|\mathbf{1}_{\{|h|>2^n\}}}{2^{nd}|\mathbb{E}h\prod_{j=1}^k f_j|} \leqslant 4\frac{2^{nd}2^{nk/(2k+1)}\mathbb{E}|h|\mathbf{1}_{\{|h|>2^n\}}}{2^{nd}|\mathbb{E}h\prod_{j=1}^k f_j|}$$
$$\leqslant 4K\frac{(\log n)^d}{2^{\frac{n(k+1)}{2k+1}}}. \tag{5.8}$$

Niech teraz $h_n = h\mathbf{1}_{\{|h|\leqslant 2^n\}}$. Z nierówności Czebyszewa, Lematu 34 i (5.3) dostajemy

$$\mathbb{P}\Big(\Big| \sum_{|\mathrm{i}|\leqslant 2^n} h_n(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}) \prod_{j=1}^k f_j(\mathbf{X}_{\mathrm{i}_{J_j}}^{\mathrm{dec}}) - 2^{nd}\mathbb{E}(h_n\prod_{j=1}^k f_j)\Big| \geqslant \frac{2^{nd}}{5}|\mathbb{E}h_n\prod_{j=1}^k f_j|\Big)$$
$$\leqslant 25\frac{\mathrm{Var}(\sum_{|\mathrm{i}|\leqslant 2^n} h_n(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\prod_{j=1}^k f_j(\mathbf{X}_{\mathrm{i}_{J_j}}^{\mathrm{dec}}))}{2^{2nd}|\mathbb{E}h_n\prod_{j=1}^k f_j|^2}$$
$$\leqslant 25\frac{(2^d-1)2^{n(2d-1)}}{2^{2nd}|\mathbb{E}h_n\prod_{j=1}^k f_j|^2}\mathbb{E}|h_n\prod_{j=1}^k f_j|^2$$
$$\leqslant 25(2^d-1)\frac{2^{2nk/(2k+1)}\mathbb{E}h_n^2}{2^n|\mathbb{E}h_n\prod_{j=1}^k f_j|^2}$$
$$\leqslant 25K(2^d-1)\frac{\log^d n}{2^{n/(2k+1)}|\mathbb{E}h_n\prod_{j=1}^k f_j|^2}. \tag{5.9}$$

Zauważmy także, że dla dużych n, z (5.3), Lematu 25 oraz (5.7)

$$
\begin{aligned}
|\mathbb{E}h_n \prod_{j=1}^k f_j| &\geqslant |\mathbb{E}h \prod_{j=1}^k f_j| - |\mathbb{E}h\mathbf{1}_{\{|h|>2^n\}} \prod_{j=1}^k f_j| \\
&\geqslant |\mathbb{E}h \prod_{j=1}^k f_j| - 2^{nk/(2k+1)} K \frac{\log^d n}{2^n} \geqslant \frac{5}{8}|\mathbb{E}h \prod_{j=1}^k f_j| \geqslant \frac{5}{8}. \qquad (5.10)
\end{aligned}
$$

Nierówności (5.8), (5.9) oraz (5.10) implikują, że dla dużych n z prawdopodobieństwem przynajmniej 0.9, mamy

$$
\begin{aligned}
&| \sum_{|\mathrm{i}|\leqslant 2^n} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}) \prod_{j=1}^k f_j(\mathbf{X}_{\mathrm{i}_{J_j}}^{\mathrm{dec}})| \\
&\geqslant | \sum_{|\mathrm{i}|\leqslant 2^n} h_n(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}) \prod_{j=1}^k f_j(\mathbf{X}_{\mathrm{i}_{J_j}}^{\mathrm{dec}})| - | \sum_{|\mathrm{i}|\leqslant 2^n} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\mathbf{1}_{\{|h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})|>2^n\}} \prod_{j=1}^k f_j(\mathbf{X}_{\mathrm{i}_{J_j}}^{\mathrm{dec}})| \\
&\geqslant 2^{nd}\Big(\frac{4}{5}|\mathbb{E}h_n \prod_{j=1}^n f_j| - \frac{1}{4}|\mathbb{E}h \prod_{j=1}^k f_j|\Big) \\
&\geqslant 2^{nd}\Big(\frac{4}{5}\cdot\frac{5}{8}|\mathbb{E}h \prod_{j=1}^n f_j| - \frac{1}{4}|\mathbb{E}h \prod_{j=1}^k f_j|\Big) \geqslant \frac{2^{nd}}{4}|\mathbb{E}h \prod_{j=1}^k f_j|.
\end{aligned}
$$

Łącząc ten fakt z (5.6), otrzymujemy, że dla dużych n, z prawdopodobieństwem przynajmniej 0.8

$$
\|(h_{\mathrm{i}})_{|\mathrm{i}|\leqslant 2^n}\|_{\mathcal{J},\log n}^* \geqslant \frac{2^{nd/2}\log^{k/2} n}{4\cdot 2^{dk/2}10^{k/2}}|\mathbb{E}h \prod_{j=1}^k f_j|.
$$

Zatem, z Lematu 23, dla dużych n

$$
\mathbb{P}\Big(\Big| \sum_{|\mathrm{i}|\leqslant 2^n} \epsilon_{\mathrm{i}}^{\mathrm{dec}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| \geqslant c_d \frac{2^{nd/2}\log^{k/2} n}{4\cdot 2^{dk/2}10^{k/2}}|\mathbb{E}h \prod_{j=1}^k f_j|\Big) \geqslant \frac{8}{10n},
$$

co razem z (5.4) daje

$$
|\mathbb{E}h \prod_{j=1}^k f_j| \leqslant L_d C \frac{4\cdot 2^{dk/2}10^{k/2}}{c_d} \log^{(d-k)/2} n.
$$

W szczególności, dla dostatecznie dużych N oraz dowolnych funkcji $f_j \colon \Sigma^{\# J_j} \to \mathbb{R}$, $j = 1, \ldots, k$ takich, że

$$
\begin{aligned}
\|f_j(X_{J_j})\|_2 &\leqslant 1, \\
\|f_j(X_{J_j})\|_\infty &\leqslant 2^{N/(2k+1)},
\end{aligned}
$$

mamy

$$|\mathbb{E}h\prod_{j=1}^{k} f_j| \leqslant L_d C\frac{4\cdot 2^{dk/2}10^{k/2}}{c_d}\log^{(d-k)/2} n \leqslant \tilde{L}_d C\log^{(d-k)/2} N.$$

Stąd, dla dużych u ($u \geqslant u_0$)

$$\begin{aligned}\sup\{|\mathbb{E}h(X)\prod_{j=1}^{k} f_j(X_{J_j})|\colon \|f_j(X_{J_j})\|_2 \leqslant 1, \|f_j(X_{J_j})\|_\infty \leqslant u^{1/(2k+1)}\}\\ \leqslant \bar{L}_d C(\operatorname{loglog} u)^{(d-k)/2},\end{aligned}$$

a więc

$$\begin{aligned}\sup\{|\mathbb{E}h(X)\prod_{j=1}^{k} f_j(X_{J_j})|\colon \|f_j(X_{J_j})\|_2 \leqslant 1, \|f_j(X_{J_j})\|_\infty \leqslant u\}\\ \leqslant \hat{L}_d C(\operatorname{loglog} u)^{(d-k)/2},\end{aligned}$$

dla wszystkich $u \geqslant u_0^{1/(2k+1)}$, co kończy dowód konieczności warunków z Twierdzenia 22 □

Dowód dostateczności. Przedstawiony poniżej dowód jest zmodyfikowaną (uproszczoną) wersją dowodu z pracy [3]. Składa się on z kilku kroków. W każdym z nich wykażemy, że

$$\sum_{n=1}^{\infty}\mathbb{P}\Big(\Big|\sum_{|\mathrm{i}|\leqslant 2^n}\pi_d h_n(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| \geqslant C 2^{nd/2}\log^{d/2} n\Big) < \infty, \tag{5.11}$$

gdzie $h_n = h\mathbf{1}_{A_n}$ dla pewnego ciągu zbiorów A_n.

Zanim przejdziemy do właściwej części dowodu, zauważmy, że z naszych założeń dla $\mathcal{J} = \{I_d\}$ oraz Lematu 33, mamy

$$\limsup_{u\to\infty}\frac{\mathbb{E}(h^2(X)\wedge u)}{(\operatorname{loglog} u)^{d-1}} \leqslant D^2.$$

Krok 1 Nierówność (5.11) zachodzi dla dowolnego $C > 0$, jeśli

$$A_n \subseteq \Big\{x\colon h^2(x) \geqslant 2^{nd}\log^d n\Big\}.$$

Z nierówności Czebyszewa i nierówności $\mathbb{E}|\pi_d h_n| \leqslant 2^d\mathbb{E}|h_n|$ (wynikającej bezpośrednio z definicji π_d i będącej trywialnym przypadkiem Lematu 21), otrzymujemy

$$\sum_n \mathbb{P}\Big(\Big|\sum_{|\mathrm{i}|\leqslant 2^n} \pi_d h_n(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| \geqslant C 2^{nd/2}\log^{d/2} n\Big)$$
$$\leqslant \sum_n \frac{\mathbb{E}\Big|\sum_{|\mathrm{i}|\leqslant 2^n} \pi_d h_n(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big|}{C 2^{nd/2}\log^{d/2} n} \leqslant 2^d \sum_n \frac{2^{nd}\mathbb{E}|h|\mathbf{1}_{\{|h|\geqslant 2^{nd/2}\log^{d/2} n\}}}{C 2^{nd/2}\log^{d/2} n}$$
$$= 2^d C^{-1}\mathbb{E}|h|\sum_n \frac{2^{nd/2}}{\log^{d/2} n}\mathbf{1}_{\{|h|\geqslant 2^{nd/2}\log^{d/2} n\}} \leqslant L_d C^{-1}\mathbb{E}\frac{|h|^2}{(\mathrm{LL}\,|h|)^d} < \infty,$$

przy czym ostatnia nierówność wynika z Lematu 26.

Krok 2 Nierówność (5.11) zachodzi dla dowolnego $C > 0$, jeśli

$$A_n \subseteq \Big\{x \in \Sigma^d \colon h^2(x) \leqslant 2^{2nd},\ \exists_{I\neq\emptyset, I_d}\ \mathbb{E}_I(h^2 \wedge 2^{2nd}) \geqslant 2^{\#I^c n}\log^d n\Big\}.$$

Z Lematu 21 oraz nierówności Czebyszewa, wystarczy udowodnić, że

$$\sum_n \frac{\mathbb{E}\Big|\sum_{|\mathrm{i}|\leqslant 2^n} \epsilon_{\mathrm{i}}^{\mathrm{dec}} h_n(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big|}{2^{nd/2}\log^{d/2} n} < \infty.$$

Zbiory A_n mogą być zapisane jako

$$\bigcup_{I\subseteq I_d, I\neq I_d,\emptyset} A_n(I),$$

gdzie zbiory $A_n(I)$ są parami rozłączne oraz

$$A_n(I) \subseteq \{x \colon h^2(x) \leqslant 2^{2nd},\ \mathbb{E}_I(h^2 \wedge 2^{2nd}) \geqslant 2^{\#I^c n}\log^d n\}.$$

Wystarczy zatem udowodnić, że

$$\sum_n \frac{\mathbb{E}\Big|\sum_{|\mathrm{i}|\leqslant 2^n} \epsilon_{\mathrm{i}}^{\mathrm{dec}} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\mathbf{1}_{A_n(I)}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big|}{2^{nd/2}\log^{d/2} n} < \infty. \tag{5.12}$$

Niech dla $l \in \mathbb{N}$

$$A_{n,l}(I) := \{x \colon h^2(x) \leqslant 2^{2nd},$$
$$2^{2l+2+\#I^c n}\log^d n > \mathbb{E}_I(h^2 \wedge 2^{2nd}) \geqslant 2^{2l+\#I^c n}\log^d n\} \cap A_n(I).$$

Wówczas $h_n\mathbf{1}_{A_n(I)} = \sum_{l=0}^{\infty} h_{n,l}$, gdzie $h_{n,l} := h_n\mathbf{1}_{A_{n,l}(I)}$.

Mamy

$$
\begin{aligned}
&\mathbb{E}|\sum_{|\mathrm{i}|\leqslant 2^n}\epsilon_{\mathrm{i}}^{\mathrm{dec}}h_{n,l}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})| \leqslant \sum_{|\mathrm{i}_{I^c}|\leqslant 2^n}\mathbb{E}_{I^c}\mathbb{E}_I|\sum_{|\mathrm{i}_I|\leqslant 2^n}\epsilon_{\mathrm{i}_I}^{\mathrm{dec}}h_{n,l}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})| \\
&\leqslant \sum_{|\mathrm{i}_{I^c}|\leqslant 2^n}\mathbb{E}_{I^c}(\mathbb{E}_I|\sum_{|\mathrm{i}_I|\leqslant 2^n}\epsilon_{\mathrm{i}_I}^{\mathrm{dec}}h_{n,l}(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})|^2)^{1/2} \leqslant 2^{(\#I^c+\#I/2)n}\mathbb{E}_{I^c}(\mathbb{E}_I|h_{n,l}|^2)^{1/2} \\
&\leqslant 2^{(\#I^c+d/2)n+l+1}\log^{d/2}n\;\mathbb{P}_{I^c}(\mathbb{E}_I(h^2\wedge 2^{2nd})\geqslant 2^{2l+\#I^cn}\log^d n).
\end{aligned}
$$

Zatem, aby uzyskać (5.12), wystarczy wykazać, że

$$
\sum_{l=0}^{\infty}\sum_n 2^{l+\#I^cn}\mathbb{P}_{I^c}(\mathbb{E}_I(h^2\wedge 2^{2nd})\geqslant 2^{2l+\#I^cn}\log^d n)<\infty.
$$

To jest jednak treść Lematu 24 dla $a=2d$.

Krok 3 Nierówność (5.11) zachodzi dla dowolnego $C>0$, jeśli

$$
\begin{aligned}
A_n \subseteq \Big\{x\colon &\forall_{I\subsetneq I_d}\mathbb{E}_I(h^2\wedge 2^{2nd})\leqslant 2^{\#I^cn}\log^d n, \\
&\exists_{I\subsetneq I_d}\frac{2^{\#I^cn}}{\log^{2d}n}\leqslant \mathbb{E}_I(h^2\wedge 2^{2nd})\Big\}.
\end{aligned}
$$

Uwaga Zauważmy, że dla $I=\emptyset$, pierwszy warunek w definicji zbioru, po prawej stronie powyższej inkluzji, oznacza po prostu $h^2\leqslant 2^{nd}\log^d n$. Podobnie warunek $2^{nd}\log^{-2d}n\leqslant h^2\wedge 2^{2nd}$ jest równoważny nierówności $2^{nd}\log^{-2d}n\leqslant h^2$.

Z Lematu 21 i nierówności Czebyszewa, wystarczy wykazać, że

$$
\sum_n\frac{\mathbb{E}|\sum_{|\mathrm{i}|\leqslant 2^n}\epsilon_{\mathrm{i}}^{\mathrm{dec}}h_n(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})|^4}{2^{2nd}\log^{2d}n}<\infty.
$$

Nierówność Chinczyna dla chaosów rademacherowych (np. [12], Theorem 3.2.2) implikuje, że

$$
\begin{aligned}
L_d^{-1}\mathbb{E}|\sum_{|\mathrm{i}|\leqslant 2^n}\epsilon_{\mathrm{i}}^{\mathrm{dec}}h_n(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})|^4 &\leqslant \mathbb{E}(\sum_{|\mathrm{i}|\leqslant 2^n}h_n(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})^2)^2 \\
&=\sum_{I\subseteq I_d}\sum_{|\mathrm{i}|\leqslant 2^n}\sum_{\substack{|\mathrm{j}|\leqslant 2^n:\\ \{k\colon i_k=j_k\}=I}}\mathbb{E}h_n(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})^2h_n(\mathbf{X}_{\mathrm{j}}^{\mathbf{dec}})^2 \\
&\leqslant \sum_{I\subseteq I_d}2^{nd}2^{n(d-\#I)}\mathbb{E}[h_n(X)^2\cdot h_n(\tilde{X}(I))^2],
\end{aligned}
$$

gdzie $X=(X_1,\ldots,X_d)$ oraz $\tilde{X}(I)=((X_i)_{i\in I},(X_i^{(1)})_{i\in I^c})$, dla $(X_i)_{i\leqslant d}$ oraz $(X_i^{(1)})_{i\leqslant d}$, niezależnych o tym samy rozkładzie.

Aby zakończyć dowód tego kroku, wystarczy zatem wykazać, że dla każdego $I \subseteq I_d$,

$$\sum_n \frac{2^{-n\#I}}{\log^{2d} n} \mathbb{E}[h_n(X)^2 h_n(\tilde{X}(I))^2] < \infty.$$

a) $I = I_d$. Mamy

$$\begin{aligned}\sum_n \frac{\mathbb{E}h_n^4}{2^{nd}\log^{2d} n} &\leqslant \mathbb{E}h^4 \sum_n \frac{1}{2^{nd}\log^{2d} n}\mathbf{1}_{\{h^2 \leqslant 2^{nd}\log^d n\}} \\ &\leqslant L_d \mathbb{E}h^4 \frac{1}{h^2(\mathrm{LL}\,|h|)^d} < \infty\end{aligned}$$

z Lematu 26.

b) $I \neq I_d, \emptyset$. Oznaczmy przez $\mathbb{E}_I, \mathbb{E}_{I^c}, \tilde{\mathbb{E}}_{I^c}$ całkowanie względem odpowiednio $(X_i)_{i\in I}$, $(X_i)_{i\in I^c}$ oraz $(X_i^{(1)})_{i\in I^c}$. Niech ponadto $\tilde{h}$, $\tilde{h}_n$ oznaczają odpowiednio $h(\tilde{X}(I))$, $h_n(\tilde{X}(I))$. Wówczas

$$\begin{aligned}\sum_n \frac{\mathbb{E}(h_n^2 \cdot \tilde{h}_n^2)}{2^{n\#I}\log^{2d} n} &\leqslant 2\sum_n \frac{\mathbb{E}(h_n^2 \cdot \tilde{h}_n^2 \mathbf{1}_{\{|h|\leqslant|\tilde{h}|\}})}{2^{n\#I}\log^{2d} n} \\ &\leqslant 2\mathbb{E}h^2\tilde{h}^2\mathbf{1}_{\{|h|\leqslant|\tilde{h}|\}} \sum_n \frac{1}{2^{n\#I}\log^{2d} n}\mathbf{1}_{\{\mathbb{E}_{I^c}(h^2\wedge 2^{2nd})\leqslant 2^{\#In}\log^d n,\ \tilde{h}^2\leqslant 2^{2nd}\}} \\ &\leqslant 2\mathbb{E}h^2\tilde{h}^2\mathbf{1}_{\{|h|\leqslant|\tilde{h}|\}} \sum_n \frac{1}{2^{n\#I}\log^{2d} n}\mathbf{1}_{\{\mathbb{E}_{I^c}(h^2\wedge \tilde{h}^2)\leqslant 2^{\#In}\log^d n,\ \tilde{h}^2\leqslant 2^{2nd}\}} \\ &\leqslant L_d\mathbb{E}h^2\tilde{h}^2\mathbf{1}_{\{|h|\leqslant|\tilde{h}|\}}\frac{1}{(\mathbb{E}_{I^c}(h^2\wedge\tilde{h}^2))(\mathrm{LL}\,|\tilde{h}|)^d} \\ &= L_d\mathbb{E}_I\tilde{\mathbb{E}}_{I^c}\tilde{h}^2\mathbb{E}_{I^c}h^2\mathbf{1}_{\{|h|\leqslant|\tilde{h}|\}}\frac{1}{(\mathbb{E}_{I^c}(h^2\wedge\tilde{h}^2))(\mathrm{LL}\,|\tilde{h}|)^d} \\ &\leqslant L_d\mathbb{E}\frac{\tilde{h}^2}{(\mathrm{LL}\,|\tilde{h}|)^d} < \infty,\end{aligned}$$

ponownie z Lematu 26.

c) $I = \emptyset$. Chcemy udowodnić, że

$$\sum_n \frac{(\mathbb{E}h_n^2)^2}{\log^{2d} n} < \infty.$$

Niech dla $J \subsetneq I_d$

$A_n^J = \{x\colon h^2(x) \leqslant 2^{nd}\log^{dn}$ i $2^{\#J^c n}(\log n)^{-2d} \leqslant \mathbb{E}_J(h^2\wedge 2^{2nd}) \leqslant 2^{\#J^c n}(\log n)^d\}$.

Wówczas

$$\mathbb{E}h_n^2 \leqslant \sum_{J \subsetneq I_d} \mathbb{E}h^2 \mathbf{1}_{A_n^J} = \sum_{J \subsetneq I_d} \mathbb{E}h_{J,n}^2,$$

gdzie $h_{J,n}$ zostały zdefiniowane w Lemacie 27. Zatem

$$(\mathbb{E}h_n^2)^2 \leqslant 2^d \sum_{J \subsetneq J_d} (\mathbb{E}h_{J,n}^2)^2,$$

i żądana zbieżność wynika z Lematu 27.

Krok 4 Nierówność (5.11) zachodzi dla **pewnego** $C \leqslant L_d D$, jeśli

$$A_n = \Big\{x \colon \forall_{I \subsetneq I_d}\ \mathbb{E}_I(h^2 \wedge 2^{2nd}) \leqslant \frac{2^{\#I^c n}}{\log^{2d} n}\Big\}.$$

Uwaga Podobnie jak w poprzednim kroku, dla $I = \emptyset$, z warunków definiujących zbiór A_n otrzymujemy $h^2 \leqslant 2^{nd} \log^{-2d} n$.

Niniejszy krok to jedyna część dowodu, w której korzystamy z założeń o wzroście wszystkich norm $\|\cdot\|_{\mathcal{J},u}$ funkcji h. Naszym celem jest oszacowanie $\|h_n\|_{\mathcal{J}}$ i użycie Wniosku 8. Możemy założyć, że $D = 1$ (jeśli $D = 0$, to $h = 0$, w przeciwnym razie możemy przeskalować funkcję h).

Rozważmy zatem $\mathcal{J} = \{J_1, \ldots, J_k\} \in \mathcal{P}_{I_d}$ i oznaczmy jak poprzednio $X = (X_1, \ldots, X_d)$, $X_I = (X_i)_{i \in I}$. Przypomnijmy, że

$$\|h_n\|_{\mathcal{J}} = \sup\Big\{\mathbb{E}[h_n(X) \prod_{i=1}^k f_i(X_{J_i})] \colon \mathbb{E}f^2(X_{J_i}) \leqslant 1\Big\}.$$

Od tej pory, aby uprościć dość już skomplikowaną notację, pominiemy argumenty funkcji h oraz f_i i będziemy pisać po prostu h zamiast $h(X)$ oraz f_i zamiast $f_i(X_{J_i})$.

Zauważmy, że jeśli $\mathbb{E}f_i^2 \leqslant 1$, $i = 1, \ldots, k$, to dla $j = 1, \ldots, k$ oraz $J \subsetneq J_j$, z nierówności Schwarza, zastosowanej warunkowo na $X_{J_j \setminus J}$, otrzymujemy

$$\begin{aligned}
\mathbb{E}|h_n \prod_{i=1}^k f_i \mathbf{1}_{\{\mathbb{E}_J f_j^2 > a^2\}}| &\leqslant \mathbb{E}_{J_j \setminus J}\Big[(\mathbb{E}_{(J_j \setminus J)^c} \prod_{i=1}^k f_i^2)^{1/2} \mathbf{1}_{\{\mathbb{E}_J f_j^2 \geqslant a^2\}} (\mathbb{E}_{(J_j \setminus J)^c} h_n^2)^{1/2}\Big] \\
&\leqslant \mathbb{E}_{J_j \setminus J}\Big[(\mathbb{E}_J f_j^2)^{1/2} \mathbf{1}_{\{\mathbb{E}_J f_j^2 \geqslant a^2\}} (\mathbb{E}_{(J_j \setminus J)^c} h_n^2)^{1/2}\Big] \\
&\leqslant [2^{n\#(J_j \setminus J)/2} \log^{-d} n] \mathbb{E}_{J_j \setminus J}\Big[(\mathbb{E}_J f_j^2)^{1/2} \mathbf{1}_{\{\mathbb{E}_J f_j^2 \geqslant a^2\}}\Big] \\
&\leqslant [2^{n\#(J_j \setminus J)/2} \log^{-d} n] a^{-1}.
\end{aligned}$$

Zatem

$$
\begin{aligned}
\|h_n\|_{\mathcal{J}} &\leqslant \sup\{\mathbb{E}[h_n \prod_{i=1}^{k} f_i] \colon \|f_i\|_2 \leqslant 1, \|(\mathbb{E}_J f_i^2)^{1/2}\|_\infty \leqslant 2^{n\#(J_i\setminus J)/2} \log^{-d} n \text{ dla } J \subsetneq J_i\} \\
&\quad + \sum_{i=1}^{k} (2^{\#J_i} - 1) \\
&\leqslant 2^d + \sup\{\mathbb{E}[h_n \prod_{i=1}^{k} f_i] \colon \|f_i\|_2 \leqslant 1, \|(\mathbb{E}_J f_i^2)^{1/2}\|_\infty \leqslant 2^{n\#(J_i\setminus J)/2} \log^{-d} n \text{ dla } J \subsetneq J_i\}.
\end{aligned}
\tag{5.13}
$$

Rozważmy zatem dowolne f_i, $i = 1, \ldots, k$ takie, że $\|f_i\|_2 \leqslant 1$, $\|(\mathbb{E}_J f_i^2)^{1/2}\|_\infty \leqslant 2^{n\#(J_i\setminus J)/2} \log^{-d} n$ dla $J \subsetneq J_i$ (zauważmy, że ostatni warunek oznacza w szczególności, że $\|f_i\|_\infty \leqslant 2^{n\#J_i/2} \log^{-d} n \leqslant 2^{nd/2}$).

Z założenia (5.2) dostajemy,

$$
|\mathbb{E}[h \prod_{i=1}^{k} f_i]| \leqslant \|h\|_{\mathcal{J}, 2^{nd/2}} \leqslant K \log^{(d-\deg \mathcal{J})/2} n
$$

dla dowolnego $K > 1$ i odpowiednio dużych n (korzystamy tu z założenia $D = 1$).

Ponadto,

$$
\begin{aligned}
\mathbb{E}|h\mathbf{1}_{\{|h|^2 \geqslant 2^{nd} \log^d n\}} \prod_{i=1}^{k} f_i| &\leqslant \mathbb{E}|h|\mathbf{1}_{\{|h|^2 \geqslant 2^{nd} \log^d n\}} \prod_{i=1}^{k} \|f_i\|_\infty \\
&\leqslant [2^{nd/2} \log^{-d \deg \mathcal{J}} n] \mathbb{E}|h|\mathbf{1}_{\{|h|^2 \geqslant 2^{nd} \log^d n\}} \\
&\leqslant [2^{nd/2} \log^{-d} n] \cdot \frac{K \log^{d-1} n}{2^{nd/2} \log^{d/2} n} \leqslant 1
\end{aligned}
\tag{5.14}
$$

dla dużych n, przy czym szacując $\mathbb{E}|h|\mathbf{1}_{\{|h|^2 \geqslant 2^{nd} \log^d n\}}$, skorzystaliśmy z Lematu 25.

Oznaczmy teraz $\tilde{h}_n = h\mathbf{1}_{\{h^2 \leqslant 2^{nd} \log^d n\}}$ i rozważmy dowolne $I \subsetneq I_d$.

$$\mathbb{E}|\tilde{h}_n \prod_{i=1}^k f_i \mathbf{1}_{\{\mathbb{E}_I(h^2 \wedge 2^{2nd}) \geqslant 2^{n\# I^c} \log^d n\}}|$$

$$\begin{aligned}
&\leqslant \mathbb{E}_{I^c}\Big[(\mathbb{E}_I \tilde{h}_n^2)^{1/2} \mathbf{1}_{\{\mathbb{E}_I(h^2 \wedge 2^{2nd}) \geqslant 2^{n\# I^c} \log^d n\}} \prod_{i=1}^k (\mathbb{E}_{J_i \cap I} f_i^2)^{1/2}\Big] \\
&\leqslant \prod_{i\colon J_i \cap I^c \neq \emptyset} [2^{n\#(J_i \cap I^c)/2} \log^{-d} n] \mathbb{E}_{I^c}[(\mathbb{E}_I \tilde{h}_n^2)^{1/2} \mathbf{1}_{\{\mathbb{E}_I(h^2 \wedge 2^{2nd}) \geqslant 2^{n\# I^c} \log^d n\}}] \\
&\leqslant [2^{n\# I^c/2} \log^{-d} n] (\mathbb{E}\tilde{h}_n^2)^{1/2} [\mathbb{P}_{I^c}(\mathbb{E}_I(h^2 \wedge 2^{2nd}) \geqslant 2^{n\# I^c} \log^d n)]^{1/2} \\
&\leqslant K[2^{n\# I^c/2} \log^{-d} n][\log^{(d-1)/2} n]\Big(\frac{\mathbb{E}(h^2 \wedge 2^{2nd})}{2^{n\# I^c} \log^d n}\Big)^{1/2} \\
&\leqslant \tilde{K}[2^{n\# I^c/2} \log^{-d} n][\log^{(d-1)/2} n] \frac{\log^{(d-1)/2} n}{2^{n\# I^c/2} \log^{d/2} n} \leqslant 1
\end{aligned} \tag{5.15}$$

dla dużych n.

Z ostatnich trzech nierówności otrzymujemy dla dużych n,

$$\begin{aligned}
|\mathbb{E}[h_n \prod_{i=1}^k f_i]| \leqslant& |\mathbb{E}h \prod_{i=1}^k f_i| + |\mathbb{E}h \mathbf{1}_{A_n^c} \prod_{i=1}^k f_i| \leqslant K \log^{(d-\deg \mathcal{J})/2} n + \mathbb{E}|h \prod_{i=1}^k f_i \mathbf{1}_{\{h^2 \geqslant 2^{nd} \log^d n\}}| \\
&+ \sum_{I \subsetneq I_d} \mathbb{E}|\tilde{h}_n \prod_{i=1}^k f_i \mathbf{1}_{\{\mathbb{E}_I(h^2 \wedge 2^{2nd}) \geqslant 2^{n\# I^c} \log^{-2d} n\}}| \\
\leqslant& K \log^{(d-\deg \mathcal{J})/2} n + 1 \\
&+ \sum_{I \subsetneq I_d} \mathbb{E}|\tilde{h}_n \prod_{i=1}^k f_i \mathbf{1}_{\{\mathbb{E}_I h^2 \wedge 2^{2nd} \geqslant 2^{n\# I^c} \log^d n\}}| \\
&+ \sum_{I \subsetneq I_d} \mathbb{E}|\tilde{h}_n \prod_{i=1}^k f_i \mathbf{1}_{\{2^{n\# I^c/2} \log^{-2d} n \leqslant \mathbb{E}_I h^2 \wedge 2^{2nd} \leqslant 2^{n\# I^c} \log^d n\}}| \\
\leqslant& K \log^{(d-\deg \mathcal{J})/2} n + 1 + 2^d - 1 + \sum_{I \subsetneq I_d} \gamma_{I,n},
\end{aligned}$$

gdzie

$$\gamma_{I,n}^2 = \mathbb{E}\tilde{h}_n^2 \mathbf{1}_{\{2^{n\# I^c/2} \log^{-2d} n \leqslant \mathbb{E}_I h^2 \wedge 2^{2nd} \leqslant 2^{n\# I^c} \log^d n\}}.$$

Ostatecznie, biorąc pod uwagę (5.13), uzyskaliśmy zatem oszacowanie

$$\|h_n\|_{\mathcal{J}} \leqslant K \log^{(d-\deg \mathcal{J})/2} n + \sum_{I \subsetneq I_d} \gamma_{I,n}$$

dla dowolnego $K > 2^d + 2^d + 1 = 2^{d+1} + 1$ i dostatecznie dużych n.

Zauważmy, że z Lematu 27 wynika, że dla każdego $I \subsetneq I_d$

$$\sum_n \frac{\gamma_{I,n}^4}{\log^{2d} n} < \infty. \tag{5.16}$$

Zatem, korzystając z Wniosku 8, otrzymujemy

$$\begin{aligned}
&\sum_{n=1}^{\infty} \mathbb{P}\Big(\Big|\sum_{|\mathrm{i}|\leqslant 2^n} \pi_d h_n(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| \geqslant C 2^{nd/2} \log^{d/2} n\Big) \\
&\leqslant K_d \sum_{\mathcal{J}\in\mathcal{P}_{I_d}} \sum_{I\subsetneq I_d} \sum_n \exp\Big[-\frac{1}{K_d}\Big(\frac{C2^{nd/2}\log^{d/2} n}{2^{nd/2}(\gamma_{I,n} + K\log^{(d-\deg\mathcal{J})/2} n)}\Big)^{2/\deg\mathcal{J}}\Big] \\
&\quad + K_d \sum_{I\subsetneq I_d} \sum_n \exp\Big[-\frac{1}{K_d}\Big(\frac{C2^{nd/2}\log^{d/2} n}{2^{n\#I/2} 2^{n\#I^c/2}\log^{-2d} n}\Big)^{2/(d+\#I^c)}\Big] \\
&\leqslant \tilde{K}_d \sum_{I\subsetneq I_d} \sum_n \exp\Big[-\frac{1}{\tilde{K}_d}\Big(\frac{C\log^{d/2} n}{\gamma_{I,n}}\Big)^{2/\deg\mathcal{J}}\Big] \\
&\quad + \tilde{K}_d \sum_{\mathcal{J}\in\mathcal{P}_{I_d}} \sum_n \exp\Big[-\frac{1}{\tilde{K}_d}\Big(\frac{C\log^{d/2} n}{K\log^{(d-\deg\mathcal{J})/2} n)}\Big)^{2/\deg\mathcal{J}}\Big] \\
&\quad + \tilde{K}_d \sum_{I\subsetneq I_d} \sum_n \exp\Big[-\frac{1}{\tilde{K}_d}\Big(C\log^{5d/2} n\Big)^{1/d}\Big].
\end{aligned}$$

Ostatni szereg po prawej stronie powyższej nierówności jest zbieżny dla dowolnego $C > 0$. Drugi z szeregów jest zbieżny dla $C \leqslant L_d$ (przypomnijmy, że za K możemy wziąć dowolną liczbę większą od $2^{d+1} + 1$). Zbieżność pierwszego szeregu wynika z faktu, że dla $x > 0$,

$$e^{-x} \leqslant K_d \frac{1}{x^{2\deg\mathcal{J}}}$$

oraz warunku (5.16). Kończy to dowód Kroku 4.

Aby udowodnić dostateczność warunku (5.2), z Wniosku 10 wystarczy wykazać zbieżność szeregu

$$\sum_n \mathbb{P}\Big(\Big|\sum_{|\mathrm{i}|\leqslant 2^n} h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| \geqslant C2^{nd/2}\log^{d/2} n\Big) \tag{5.17}$$

dla pewnego $C < L_d D$. W tym celu, dla każdego n rozbijamy Σ^d na cztery rozłączne zbiory A_n^i, $i = 1, \ldots, 4$, tak aby zbiór A_n^i spełniał warunki z kroku i powyżej (jest to możliwe, gdyż zbiory z kroków 1-4 sumują się do Σ^d). Wówczas, dla $C = L_d D$, z nierówności trójkąta i nierówności (5.11), udowodnionej dla każdego z ciągów A_n^i, zbieżny jest szereg

$$\sum_n \mathbb{P}\Big(\Big|\sum_{|\mathrm{i}|\leqslant 2^n} \pi_d h(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})\Big| \geqslant C2^{nd/2}\log^{d/2} n\Big)$$

To kończy dowód, gdyż z kanoniczności mamy $\pi_d h = h$. □

Uwaga Zauważmy, że dowodząc dostateczności w Twierdzeniu 22, nie używaliśmy założenia o symetryczności funkcji h.

5.2 Prawo iterowanego logarytmu dla zwykłych U-statystyk

Korzystając z powyższego twierdzenia, możemy udowodnić główny wynik tej części pracy.

TWIERDZENIE 23. *Dla symetrycznej funkcji $h\colon \Sigma^d \to \mathbb{R}$, prawo iterowanego logarytmu*

$$\limsup_{n\to\infty} \frac{1}{(n\log\log n)^{d/2}} \Big| \sum_{\mathrm{i}\in I_n^d} h(\mathbf{X}_\mathrm{i}) \Big| \leqslant C < \infty \text{ p.n.} \tag{5.18}$$

zachodzi wtedy i tylko wtedy, gdy h jest kanoniczna (względem rozkładu X_1) oraz dla wszystkich $\mathcal{J} \in \mathcal{P}_{I_d}$

$$\limsup_{u\to\infty} \frac{1}{(\log\log u)^{(d-\deg\mathcal{J})/2}} \|h\|_{\mathcal{J},u} \leqslant D < \infty. \tag{5.19}$$

Dokładniej, (5.18) implikuje (5.19) z $D = L_d C$ i odwrotnie, (5.19) wraz z kanonicznością implikuje (5.18) z $C = L_d D$.

DOWÓD. Dostateczność wynika z Wniosku 11 oraz Twierdzenia 22. Aby udowodnić konieczność, zauważmy że z Lematu 28 oraz Wniosku 12, h spełnia zrandomizowane, uniezależnione prawo iterowanego logarytmu (4.12), a zatem, z Twierdzenia 22, warunki wzrostu dla $\|h\|_{\mathcal{J},u}$ także są spełnione (jak łatwo zauważyć, normy $\|\cdot\|_{\mathcal{J},u}$ jądra $h(X_1,\ldots,X_d)$ oraz $\varepsilon_1\cdots\varepsilon_d h(X_1,\ldots,X_d)$ są równe). Zatem, aby zakończyć dowód, pozostaje wykazać kanoniczność funkcji h. Warunek na całkowalność (5.3) implikuje, że $\mathbb{E}|\pi_d h|^p < \infty$ dla wszystkich $p < 2$, zatem z prawa wielkich liczb typu Marcinkiewicza dla kanonicznych U-statystyk (Giné i Zinn, [22]) wynika, że

$$\frac{1}{n^{d/p}} \sum_{\mathrm{i}\in I_n^d} \pi_d h(\mathbf{X}_\mathrm{i}) \to 0 \text{ p.n.}$$

dla $n \to \infty$. Ponadto, z założenia o zachodzeniu prawa iterowanego logarytmu, mamy także

$$\frac{1}{n^{d/p}} \sum_{\mathrm{i}\in I_n^d} h(\mathbf{X}_\mathrm{i}) \to 0 \text{ p.n.}$$

Zauważmy, że z rozkładu Hoeffdinga (Lemat 22),

$$
\begin{aligned}
&\sum_{\mathrm{i}\in I_n^d}(h(\mathbf{X}_{\mathrm{i}})-\pi_d h(\mathbf{X}_{\mathrm{i}}))\\
&\qquad=\sum_{k=0}^{d-1}\binom{d}{k}\cdot\frac{(n-k)!}{n!}\cdot\frac{n!}{(n-d)!}\sum_{\substack{i_1,\ldots,i_k\leqslant n\\ i_j\neq i_l \text{ dla j}\neq\text{l}}}\pi_k h(X_{i_1},\ldots,X_{i_k})\\
&\qquad=(n-d+1)\sum_{\substack{i_1,\ldots,i_{d-1}\leqslant n\\ i_j\neq i_l \text{ dla j}\neq\text{l}}} g(X_{i_1},\ldots,X_{i_{d-1}}),
\end{aligned}
\tag{5.20}
$$

gdzie

$$
g(x_1,\ldots,x_{d-1})=\frac{1}{(d-1)!}\sum_{\sigma}\tilde{g}(x_{\sigma(1)},\ldots,x_{\sigma(d-1)}),
$$

przy czym suma rozciąga się na wszystkie permutacje σ zbioru I_{d-1} oraz

$$
\tilde{g}(x_1,\ldots,x_{d-1})=\sum_{k=0}^{d-1}\binom{d}{k}\pi_k h(x_1,\ldots,x_k).
$$

Otrzymujemy więc

$$
\frac{n-d+1}{n^{d/p}}\left|\sum_{\substack{i_1,\ldots,i_{d-1}\leqslant n\\ i_j\neq i_l \text{ dla j}\neq\text{l}}} g(X_{i_1},\ldots,X_{i_{d-1}})\right|\to 0 \text{ p.n.}
$$

Stąd ciąg

$$
\frac{1}{n^{d/p-1}}\left|\sum_{\substack{i_1,\ldots,i_{d-1}\leqslant n\\ i_j\neq i_l \text{ dla j}\neq\text{l}}} g(X_{i_1},\ldots,X_{i_{d-1}})\right|
$$

jest stochastycznie ograniczony. Kładąc $p=2d/(d+1)$ otrzymujemy normowanie $n^{-(d-1)/2}$, takie jak w centralnym twierdzeniu graniczny dla U-statystyk rzędu $d-1$ (np. [12], Theorem 4.2.4). Z wyników Giné i Zinna ([23], Theorem 1 lub [12], Theorem 4.2.6) wynika, że g jest kanoniczna oraz $\mathbb{E}g^2<\infty$, a zatem g spełnia centralne twierdzenie graniczne dla kanonicznych U-statystyk (np. [44] lub [12], Theorem 4.2.4), z którego dostajemy

$$
g(X_1,\ldots,X_{d-1})=0 \text{ p.n.}
$$

Stosując (5.20) dla $n=d$, otrzymujemy $h=\pi_d h$, co dowodzi kanoniczności h. $\square$

5.3 Zbiór graniczny

Dla $\mathbb{E}h^2<\infty$, zbiór graniczny (zbiór punktów skupienia) w prawie iterowanego logarytmu (4.2) jest z prawdopodobieństwem 1 równy

$$
2^{d/2}\{\mathbb{E}h(X_1,\ldots,X_d)f(X_1)\cdots f(X_d)\colon \mathbb{E}f^2(X_1)\leqslant 1\},
$$

co zostało udowodnione przez Arconesa i Giné [5]. W ogólności nie jest to prawda. Dla $d = 2$ wiadomo [24], że zbiór ten jest odcinkiem. Jego końce znane są jedynie w pewnych specjalnych przypadkach [32], w których granica górna z prawa iterowanego logarytmu okazuje się być dość skomplikowaną funkcją „deterministycznych" granic górnych występujących w warunkach koniecznych i dostatecznych. Naturalna wydaje się hipoteza, że także w ogólnej sytuacji, granica górna w prawie iterowanego logarytmu jest funkcją tych wielkości.

Rozważania na temat prawa iterowanego logarytmu dla rzeczywistych U-statystyk chcielibyśmy zakończyć następującym twierdzeniem.

Twierdzenie 24. *Zbiór graniczny w prawie iterowanego logarytmu (4.2), tj. zbiór punktów skupienia ciągu* $((n\operatorname{loglog} n)^{-d/2}\sum_{|\mathrm{i}|\leqslant n} h(\mathbf{X}_{\mathrm{i}}))$, *jest deterministycznym przedziałem.*

Dowód. Jeśli wykażemy, że zbiór punktów skupienia jest przedziałem, fakt, że jest z prawdopodobieństwem 1 stały, będzie wynikał z prawa 0-1 Hewitta-Savage'a.

Aby udowodnić, że zbiór graniczny jest przedziałem, wystarczy wykazać, że

$$\limsup_{n\to\infty}\left|\frac{\sum_{\mathrm{i}\in I_n^d} h(\mathbf{X}_{\mathrm{i}})}{n^{d/2}\operatorname{loglog}^{d/2} n} - \frac{\sum_{\mathrm{i}\in I_{n-1}^d} h(\mathbf{X}_{\mathrm{i}})}{(n-1)^{d/2}\operatorname{loglog}^{d/2}(n-1)}\right| = 0 \text{ p.n.} \tag{5.21}$$

Zauważmy, że

$$\begin{aligned} 0 \leqslant \frac{1}{(n-1)^{d/2}\operatorname{loglog}^{d/2}(n-1)} - \frac{1}{n^{d/2}\operatorname{loglog}^{d/2} n} &\leqslant \frac{n^{d/2}-(n-2)^{d/2}}{n^{d/2}(n-2)^{d/2}\operatorname{loglog}^{d/2} n} \\ &\leqslant \frac{d}{2}\frac{2}{n(n-2)^{d/2}\operatorname{loglog}^{d/2} n}. \end{aligned}$$

Z powyższej nierówności i prawa iterowanego logarytmu, wynika, że

$$\lim_{n\to\infty}\Big|\sum_{\mathrm{i}\in I_n^d} h(\mathbf{X}_{\mathrm{i}})\Big|\Big(\frac{1}{(n-1)^{d/2}\operatorname{loglog}^{d/2}(n-1)} - \frac{1}{n^{d/2}\operatorname{loglog}^{d/2} n}\Big) = 0 \quad \text{p.n.}$$

Zatem, aby wykazać (5.21), wystarczy udowodnić, że

$$\limsup_{n\to\infty}\frac{1}{n^{d/2}\operatorname{loglog}^{d/2} n}\Big|\sum_{\mathrm{i}\in I_n^d,\ i_d=n} h(\mathbf{X}_{\mathrm{i}})\Big| = 0 \text{ p.n.}$$

Powyższa równość wynika ze zbieżności

$$\sum_n \sum_{2^{n-1}<k\leqslant 2^n} \mathbb{P}\Big(\Big|\sum_{\mathrm{i}\in I_k^d,\ i_d=k} h(\mathbf{X}_{\mathrm{i}})\Big| > \delta 2^{nd/2}\log^{d/2} n\Big) < \infty \tag{5.22}$$

dla wszystkich $\delta > 0$. Niech $\bar{\pi}_{d-1}$ oznacza projekcję Hoeffdinga jedynie względem pierwszych $d-1$ zmiennych. Podobnie jak w dowodzie prawa iterowanego logarytmu, kanoniczność h daje $\bar{\pi}_{d-1}h = h$, a zatem wystarczy udowodnić, że

$$\sum_n \sum_{2^{n-1}<k\leqslant 2^n} \mathbb{P}\Big(\Big| \sum_{\mathrm{i}\in I_k^d,\, i_d=k} \bar{\pi}_{d-1}h(\mathbf{X}_\mathrm{i})\Big| > \delta 2^{nd/2}\log^{d/2} n\Big) < \infty.$$

Będziemy teraz postępować jak w dowodzie Twierdzenia 22, tzn. udowodnimy zbieżność powyższego szeregu z funkcją h zastąpioną przez $h_n = h\mathbf{1}_{A_n}$ dla odpowiednich zbiorów A_n.

Krok 1

$$A_n \subseteq \{x \in \Sigma^d \colon h^2(x) \geqslant 2^{nd}\log^d n\}.$$

Ponieważ dla $2^{n-1} < k \leqslant 2^n$, $\#\{\mathrm{i} \in I_k^d \colon i_d = k\} \leqslant 2^{n(d-1)}$, możemy użyć nierówności Czebyszewa, dokładnie jak w pierwszym kroku dowodu Twierdzenia 22.

Krok 2

$$A_n \subseteq \{x \colon h^2(x) \leqslant 2^{nd}\log^d n,\ \exists_{I\subseteq I_{d-1}, I\neq\emptyset}\ \mathbb{E}_I(h^2 \wedge 2^{2nd}) \geqslant 2^{\#I^c n}\log^d n\}.$$

Zauważmy, że z twierdzenia o uniezależnianiu i Lematu 21, zastosowanego warunkowo na X_k, mamy

$$\begin{aligned}\mathbb{E}\Big|\sum_{\mathrm{i}\in I_k^d,\, i_d=k} \bar{\pi}_{d-1}h_n(\mathbf{X}_\mathrm{i})\Big| &\leqslant L_d\mathbb{E}\Big|\sum_{\mathrm{i}\in I_k^d,\, i_d=k} \bar{\pi}_{d-1}h_n(\mathbf{X}_\mathrm{i}^{\mathrm{dec}})\Big| \\ &\leqslant 2^{d-1}L_d\mathbb{E}\Big|\sum_{\mathrm{i}\in I_k^d,\, i_d=k} \epsilon_\mathrm{i}^{\mathrm{dec}} h_n(\mathbf{X}_\mathrm{i}^{\mathrm{dec}})\Big|.\end{aligned}$$

Zatem, jeśli zdefiniujemy zbiory $A_n(I)$ oraz $A_{n,l}(I)$ (dla $I \subseteq I_{d-1}, I \neq \emptyset$) jak w kroku 2 dowodu Twierdzenia 22, wystarczy wykazać

$$\sum_n \sum_{2^{n-1}<k\leqslant 2^n} \frac{1}{2^{nd/2}\log^{d/2} n} \sum_{l=0}^{\infty} \mathbb{E}\Big|\sum_{\mathrm{i}\in I_k^d,\, i_d=k} \epsilon_\mathrm{i}^{\mathrm{dec}} h_{n,l}(\mathbf{X}_\mathrm{i})\Big| < \infty,$$

gdzie dla ustalonego I, funkcje $h_{n,l}$ są zdefiniowane jak w dowodzie Twierdzenia 22. Ale dla każdych $2^{n-1} < k \leqslant 2^n, l$ przez argument podobny do przedstawionego w tamtym dowodzie

$$\begin{aligned}\mathbb{E}\Big|\sum_{\mathrm{i}\in I_k^d,\, i_d=k} \epsilon_\mathrm{i}^{\mathrm{dec}} h_{n,l}(\mathbf{X}_\mathrm{i})\Big| \leqslant & [2^{(\#I^c+d/2-1)n+l+1}\log^{d/2} n] \times \\ & \times \mathbb{P}_{I^c}(\mathbb{E}_I(h^2 \wedge 2^{2nd}) \geqslant 2^{2l+\#I^c n}\log^d n).\end{aligned}$$

Stąd

$$\sum_{2^{n-1}<k\leqslant 2^n} \mathbb{E}|\sum_{\mathrm{i}\in I_k^d,\ i_d=k} \epsilon_{\mathrm{i}}^{\mathrm{dec}} h_{n,l}(\mathbf{X}_{\mathrm{i}})| \leqslant [2^{(\#I^c+d/2)n+l+1}\log^{d/2} n]\times$$

$$\times\, \mathbb{P}_{I^c}(\mathbb{E}_I(h^2\wedge 2^{2nd}) \geqslant 2^{2l+\#I^c n}\log^d n)$$

i możemy zakończyć jak w kroku 2 dowodu twierdzenia 22.

Krok 3

$$A_n \subseteq \{x\colon h^2(x) \leqslant 2^{nd}\log^d n,\ \forall_{I\subseteq I_{d-1}, I\neq\emptyset}\ \mathbb{E}_I(h^2\wedge 2^{2nd}) \leqslant 2^{\#I^c n}\log^d n\}.$$

Używając podobnych argumentów o uniezależnianiu i porównywaniu momentów, jak powyżej oraz nierówności Chinczyna dla chaosów rademacherowych, otrzymujemy

$$\begin{aligned}\mathbb{E}|\sum_{\mathrm{i}\in I_k^d,\ i_d=k} \bar{\pi}_{d-1}h_n(\mathbf{X}_{\mathrm{i}})|^4 &\leqslant L_d\mathbb{E}|\sum_{\mathrm{i}\in I_k^d,\ i_d=k} \bar{\pi}_{d-1}h_n(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})|^4\\ &\leqslant 2^{4(d-1)}L_d\mathbb{E}|\sum_{\mathrm{i}\in I_k^d,\ i_d=k} \epsilon_{\mathrm{i}}^{\mathrm{dec}} h_n(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}})|^4\\ &\leqslant \tilde{L}_d\mathbb{E}(\sum_{|\mathrm{i}|\leqslant k,\ i_d=k} h_n^2(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}))^2,\end{aligned}$$

przy czym w ostatniej nierówności dodaliśmy składniki związane z „przekątną” jedynie aby uczynić ten krok bardziej podobnym do kroku 3 z dowodu Twierdzenia 22. Wystarczy zatem udowodnić, że

$$\sum_n \frac{2^n\mathbb{E}(\sum_{|\mathrm{i}|\leqslant 2^n,\ i_d=2^n} h_n^2(\mathbf{X}_{\mathrm{i}}^{\mathrm{dec}}))^2}{2^{2nd}\log^{2d} n} < \infty.$$

To jednak może znów być wykazane jak w Twierdzeniu 22 (krok 3), przez porównanie jedynie z punktami a), b). Przypadek c) (który jako jedyny używał oszacowań z dołu w definicji zbioru A_n i czynił niezbędnym krok 4 dowodu) nie pojawia się tutaj, gdyż indeks i_d jest w naszej sytuacji ustalony. Dowód Twierdzenia może być zatem zakończony jak dla Twierdzenia 22, przez rozbicie Σ^d na 3 rozłączne części (dla każdego n), odpowiadające powyższym krokom 1–3.

□

5.4 Uwagi nt. prawa iterowanego logarytmu dla U-statystyk o wartościach w przestrzeni Hilberta

Na zakończenie podamy bez dowodu, otrzymaną niedawno wspólnie z Rafałem Latałą charakteryzację prawa iterowanego logarytmu dla U-statystyk o wartościach w przestrzeni Hilberta. Aby sformułować warunki konieczne i dostateczne, wprowadźmy odpowiedniki norm $\|\cdot\|_{\mathcal{J},u}$.

DEFINICJA 18. *Niech H będzie ośrodkową przestrzenią Hilberta, zaś $h\colon \Sigma^d \to H$ funkcją kanoniczną. Dla dowolnego zbioru $K \subseteq I_d$, rozbicia $\mathcal{J} = \{J_1, \ldots, J_k\} \in \mathcal{P}_{I_d\setminus K}$ oraz $u > 0$, zdefiniujmy*

$$\|h\|_{K,\mathcal{J},u} = \sup\{\mathbb{E}\langle h(X), g(X_K)\rangle \prod_{i=1}^{k} f_i(X_{J_i})\colon g\colon \Sigma^K \to H,$$
$$f_i\colon \Sigma^{J_i} \to \mathbb{R}, \|g\|_2, \|f_i\|_2 \leqslant 1, \|g\|_\infty, \|f_i\|_\infty \leqslant u\},$$

przy czym dla $K = \emptyset$ powyższa definicja interpretowana jest analogicznie jak Definicja 12.

Wówczas zachodzi następujące twierdzenie.

TWIERDZENIE 25 (Adamczak, Latała, [4]). *Dla dowolnej funkcji $h\colon \Sigma^d \to \mathbb{R}$, prawo iterowanego logarytmu*

$$\limsup_{n\to\infty} \frac{1}{(n\mathrm{LL}\, n)^{d/2}} \Big| \sum_{\mathbf{i}\in I_n^d} h(\mathbf{X}_{\mathbf{i}}) \Big| \leqslant C < \infty \text{ p.n.} \tag{5.23}$$

zachodzi wtedy i tylko wtedy, gdy

$$\mathbb{E}\frac{|h|^2}{(\mathrm{LL}\,|h|)^d} < \infty,$$

h jest kanoniczna względem rozkładu X_1 oraz dla każdego zbioru $K \subseteq I_d$ oraz rozbicia $\mathcal{J} \in \mathcal{P}_{I_d}$

$$\limsup_{u\to\infty} \frac{\|h\|_{K,\mathcal{J},u}}{(\mathrm{loglog}\, u)^{(d-\deg\mathcal{J})/2}} \leqslant D < \infty. \tag{5.24}$$

Dokładniej, (5.23) implikuje (5.24) z $D = L_d C$ oraz (5.24) wraz z pozostałymi warunkami, implikuje (5.23) z $C = L_d D$.

Bibliografia

[1] Adamczak R. (2005). Logarithmic Sobolev inequalities and concentration of measure for convex functions and polynomial chaoses. *Bull. Pol. Acad. Sci. Math.* **53**, no. 2, 221-238. MR2163396.

[2] Adamczak, R. (2006). Moment Inequalities for U-statistics. Praca ukaże się w *Ann. Probab.* **34**

[3] Adamczak, R., Latała, R. (2006) LIL for canonical U-statistics. Preprint.

[4] Adamczak, R., Latała, R. (2006) LIL for canonical U-statistics in Hilbert spaces. W przygotowaniu.

[5] Arcones, M., Giné, E. (1995). On the law of the iterated logarithm for canonical U-statistics and processes. *Stochastic Process. Appl.* **58** 217–245. MR1348376.

[6] Arcones, M., Giné, E. (1993). On decoupling, series expansion and tail behaviour of chaos processes. *J. Theoret. Probab.* **6** , 101-122. MR1201060.

[7] Borell, C. (1984). On a Taylor series of a Wiener polynomial. In *Seminar Notes on Multiple Stochastic Integration, polynomial chaos and their integration.* Case Western Reserve Univ., Cleveland.

[8] Bousquet O., Boucheron S., Lugosi G., Massart P.(2005). Moment inequalities for functions of independent random variables. *Ann. Probab.* **33**, no. 2, 514-560. MR2123200.

[9] Barthe, F., Roberto, C. (2004). Sobolev inequalities for probability measures on the real line. *Studia Math.* **159** (2003), no. 3, 481–497. MR2052235.

[10] Clémençon S., Lugosi G., Vayatis N. (2006). Ranking and empirical minimization of U-statistics. Preprint.

[11] Dehling, H., Denker, M., Philipp, W. (1986). *Probab. Theory Relat. Fields* **72**, no. 1, 111–131. MR0835162.

[12] de la Peña, V.H. and Giné, E. (1999). *Decoupling. From Dependence to Independence*. Springer, New York. MR1666908.

[13] de la Peña, V.H. and Montgomery-Smith, S. (1994). Bounds for the tail probabilities of U-statistics and quadratic forms. *Bull. Amer. Math. Soc.* **31** 223–227. MR1261237.

[14] de la Peña, V. H., Montgomery-Smith, S. J. (1995) Decoupling inequalities for the tail probabilities of multivariate U-statistics. *Ann. Probab.* **23**, no. 2, 806–816. MR1334173.

[15] Gikhman, I., Skorohod, A.V. (1975, 2004) Theory of stochastic processes II. Springer, Berlin, Heidelberg, New York 1975, 2004.

[16] Gentil, I., Guillin, A., Miclo, L. (2005) Modified logarithmic Sobolev inequalities and transportation inequalities. *Probability Theory and Related Fields*, **133(3)**, 409-436. MR2198019.

[17] Gentil, I., Guillin, A., Miclo L. (2006) Modified logarithmic Sobolev inequalities in null curvature. Praca ukaże się w Revista Matematica Iberoamericana. Dostepna pod adresem http://www.arxiv.org/abs/math.PR/0503585.

[18] Giné E., Latała, R., Zinn, J. (2000). Exponential and moment inequalities for U-statistics. *High Dimensional Probability II*, 13–38. Progr. Probab., 47, Birkhauser, Boston. MR1857312.

[19] Giné E., Mason D. (2006). On local U-statistic processes and the estimation of densities of functions of several variables. Praca ukaże się w *Ann. Statist.*

[20] Giné E., Mason D. (2006). Laws of the iterated logarithm for the local U-statistic process. Praca ukaże się w *J. Theoret. Probab.*

[21] Giné, E., Zhang, C.-H. (1996). On integrability in the LIL for degenerate U-statistics. *J. Theoret. Probab.* **9**, 385–412. MR1385404.

[22] Giné, E., Zinn, J. (1992). Marcinkiewicz type laws of large numbers and convergence of moments for U-statistics. *Probability in Banach spaces*, **8**, 273–291, Progr. Probab., 30, Birkhauser, Boston. MR1227625

[23] Giné, E., Zinn, J. (1994). A remark on convergence in distribution of U-statistics. *Ann. Probab.* **22** 117–125. MR1258868.

[24] Giné E., Kwapień, S., Latała, R., Zinn, J. (2001). The LIL for canonical U-statistics of order 2, *Ann. Probab.* **29** 520–557. MR1825163.

[25] Halmos, P.R. (1946). The theory of unbiased estimation. *Ann. Math. Statistics* **17** 34–43. MR0015746.

[26] Hitczenko, P. (1993). Domination inequality for martingale transforms of a Rademacher sequence. *Israel J. Math.* **84** 161–178. MR1244666.

[27] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statistics* **19** 293–325. MR0026294.

[28] Hoeffding, W. (1961). The strong law of large numbers for U-statistics. *Inst. Statist. Mimeo Ser.* **302**, University of North Carolina. Chapel Hill, North Carolina.

[29] Houdré, C., Reynaud-Bouret, P. (2003). Exponential Inequalities, with Constants, for U-statistics of Order Two. *Stochastic inequalities and applications*, 55–69. Progr. Probab., 56, Birkhauser, Basel. MR2073426.

[30] Kwapień, S. (1975). A theorem on the Rademacher series with vector valued coefficients. Probability in Banach Spaces. Oberwolfach 1975. Lecture Notes in Mathematics, vol. 526. Springer, Berlin Heildeberg 1976, pp. 157-158. MR0451333.

[31] Gluskin, E.D., Kwapień, S. (1995). Tail and moment estimates for sums of independent random variables with logarithmically concave tails. *Studia Math.* **114** , no. 3, 303–309. MR1338834.

[32] Kwapień, S., Latała, R., Oleszkiewicz, K. and Zinn, J. (2003). On the limit set in the law of the iterated logarithm for U-statistics of order two. *High dimensional probability*, III (Sandjberg), 111–126. Progr. Probab., 55, Birkhauser, Basel. MR2033884.

[33] Kwapień S. Woyczynski W. (1992) *Random Series and Stochastic Integrals. Single and Multiple.* Birkhäuser, Boston. MR1167198.

[34] Latała, R. (1999). Tail and moment estimates for some types of chaos. *Studia Math.* **135**, 39-53. MR1686370.

[35] Latała, R. (2006). Estimation of moments and tails of Gaussian chaoses. Praca ukaże się w *Ann. Probab.* **34**.

[36] Latała, R., Zinn, J. (2000). Necessary and sufficient conditions for the strong law of large numbers for U-statistics. *Ann. Probab.* **28** 1908–1924. MR1813848.

[37] Latała, R., Oleszkiewicz, K. (2000). Between Sobolev and Poincaré. *Geometric aspects of functional analysis*, 147–168, Lecture Notes in Math., 1745, Springer, Berlin, 2000. MR1796718.

[38] Ledoux,M. (2001). The concentration of measure phenomenon. Mathematical Surveys and Monographs, **89**. American Mathematical Society, Providence, RI. MR1849347.

[39] Ledoux, M., Talagrand, M. (1991). *Probability in Banach spaces. Isoperimetry and processes.* Ergebnisse der Mathematik und ihrer Grenzgebiete (3) 23. Springer-Verlag, Berlin. MR1102015.

[40] Łochowski R. (2006). Moment and Tail Estimates for Multidimensional Chaoses Generated by Symmetric Random Variables with Logarithmically Concave Tails. Banach Center Publ. **72**, 161-176.

[41] Montgomery-Smith, S. (1993). Comparison of sums of independent identically distributed random vectors. *Probab. Math. Statist.* **14** 281–285. MR1321767.

[42] Paley, R.E.A.C., Zygmund, A. (1932). A note on analytic functions in the unit circle. PRoceedings Cambridge Phil. Soc. **28**, 266–272.

[43] Reynaud-Bouret, P. (2003). Adaptive estimation of the intensity of inhomogeneous Poisson processes via concentration inequalities. *Probab. Theory Related Fields* **126**, no. 1, 103–153. MR1981635.

[44] Rubin, H., Vitale, R.A. (1980). Asymptotic distribution of symmetric statistics. *Ann. Statist.* **8** 165–170. MR0557561.

[45] Samson, P.M. (2000). Concentration of measure inequalities for Markov chains and Φ-mixing processes. *Ann. Probab.* **28**, 416-461. MR1756011.

[46] Serfling, R.J. (1971). The law of the iterated logarithm for U-statistics and related von Mises functionals. *Ann. Math. Statist.* **42** 1794.

[47] Talagrand, M. (1996). New concentration inequalities in product spaces, *Inventionnes Math* **126**, 505-563. MR1419006.

[48] Talagrand, M. (1996). A New Look at Independence, *Ann. Probab.* **24**, 1-34. MR1387624.