University of Warsaw

Faculty of Mathematics, Informatics and Mechanics

Rafał Martynek

Estimates of suprema of stochastic processes with application of the chaining method

PhD dissertation

Supervisor:
dr hab. Witold Bednorz
Institute of Mathematics
University of Warsaw

September 2020

Declaration of Authorship

Author's declaration:
I hereby declare that this dissertation is my own work.

7th September, 2020

....
Rafał Martynek

Supervisor's declaration:
The dissertation is ready to be reviewed

7th September, 2020

....
dr hab. Witold Bednorz

Abstract

In the following thesis we are studying some consequences of Bednorz-Latała theorem concerning the characterization of the Bernoulli process given by the collection $(B_t)_{t\in T}$, where $B_t = \sum_{i\geq 1} \varepsilon_i t_i$, $(\varepsilon_i)_{i\geq 1}$ is a sequence of independent Bernoulli variables satisfying $\mathbb{P}(\varepsilon_i = -1) = \mathbb{P}(\varepsilon_i = 1) = 1/2$ and $T \subset \ell^2$. It states that $b(T) := \mathbf{E}\sup_{t\in T} B_t$ is comparable with

$$\inf\{\sup_{t\in T_1} |t|_1 + \gamma_2(T_2, d_2) :\ T \subset T_1 + T_2\},$$

where $\|\cdot\|_1$ is the ℓ^1 norm and γ_2 is Talagrand's γ_2-number. The main results of this thesis provide positive answers to the conjectures concerning decompositions infinitely divisible processes, empirical processes and selector processes posed in [32]. They are based on the refolmulation due to M. Talagrand of the lower bound of $b(T)$ in terms of a special functional

$$\int_T I_\mu(t)\mu(dt),$$

where μ is some probability measure on T.
This thesis includes also some generalizations of the contraction principle for Bernoulli processes and Lévy-Ottaviani type of inequality for the Bernoulli process with monotone coefficients.

AMS 2000 subject classifications: 60G15, 60G17, 60G50

Keywords: Bernoulli processes, infinitely divisible processes, process boundedness, chaining method

Rafał Martynek
September 2020

Streszczenie

Niniejsza rozprawa poświęcona jest badaniom wniosków płynących z twierdzenia Bednorza-Latały dotyczącego charakteryzacji procesu Bernoulliego rozumianego jako rodzina $(B_t)_{t\in T}$, gdzie $B_t = \sum_{i\geq 1} \varepsilon_i t_i$, $(\varepsilon_i)_{i\geq 1}$ jest ciągiem niezależnych zmiennych Bernoulliego spełniających $\mathbb{P}(\varepsilon_i = -1) = \mathbb{P}(\varepsilon_i = 1) = 1/2$, $T \subset \ell^2$. Twierdzenie to mówi, że $b(T) := \mathbf{E} \sup_{t\in T} B_t$ porównywalne jest z

$$\inf\{\sup_{t\in T_1} |t|_1 + \gamma_2(T_2, d_2) : \; T \subset T_1 + T_2\},$$

gdzie $\|\cdot\|_1$ jest normą w ℓ^1, a γ_2 jest liczbą γ_2 Talagranda. Główne wyniki rozprawy dają pozytywną odpowiedź na hipotezy dotyczące dekompozycji procesów nieskończenie podzielnych, procesów empirycznych i procesów selektorów postawionych w [32]. Wyniki te oparte są na przeformułowaniu pochodzącym od M. Talagranda dolnego oszacowania $b(T)$ w terminach specjalnego funkcjonału

$$\int_T I_\mu(t)\mu(dt),$$

gdzie μ jest pewną miarą probabilistyczną na T.

Rozprawa zajmuje się również pewnymi uogólnieniami zasady kontrakcji dla procesów Bernoulliego i nierównością typu Léviego-Ottavianiego dla procesów Bernoulliego ze współczynnikami monotonicznymi.

Klasyfikacja tematyczna: 60G15, 60G17, 60G50

Słowa kluczowe: procesy Bernoulliego, procesy nieskończenie podzielne, ograniczenia procesów, metoda łańcuchowa

Rafał Martynek
Wrzesień 2020

Contents

Chapter 1

Introduction

1.1 Upper and lower bounds of stochastic processes

1.1.1 Brief history

The study of suprema of stochastic processes understood as a family of random variables $(X_t)_{t\in T}$ indexed by some set T is one of the primal tasks in probability. Precisely, we consider $\mathbf{E}\sup_{t\in T} X_t$ and we look for upper and lower bounds of it in terms of deterministic quantities related to the geometry of the index set T. In order to guarantee that this quantity is well-defined we will assume that T is countable or define $\mathbf{E}\sup_{t\in T} X_t := \sup_{F\subset T}\mathbf{E}\sup_{t\in F} X_t$, where the supremum runs over all finite subsets F of T. It is motivated by various theoretical problems such as continuity of sample paths of stochastic processes or estimates of norms of random vectors.
The first successful result in this area concerns the centred Gaussian process $(G_t)_{t\in T}$. The fundamental observation was that the boundedness of this process is determined by the structure of the metric space (T, d_2), where $d_2(s,t) = (\mathbf{E}(G_t - G_s)^2)^{1/2}$, which is the Euclidean distance on T. It is now a classic, but essential result due to Fernique [8] and Talagrand ([28] or [32, Theorem 2.2.18]) that

$$g(T) = \mathbf{E}\sup_{t\in T} G_t$$

is comparable with so-called γ_2 number. To get the idea behind the definition of this quantity one might consider a finite set T look at the following simple observation based on the union bound. Notice that $g(T) = 2\mathbf{E}\sup_{s,t\in T}(G_t - G_s)$ and for $p \geq 1$

$$\mathbf{E}\sup_{s,t\in T}(G_t - G_s) \leq \left(\mathbf{E}\sup_{s,t\in T}|G_t - G_s|^p\right)^{1/p} \leq \left(\mathbf{E}\sum_{s,t\in T}|G_t - G_s|^p\right)^{1/p} \leq |T|^{2/p}\sup_{s,t\in T}(\mathbf{E}|G_t - G_s|^p)^{1/p}, \tag{1.1}$$

where $|\cdot|$ denotes the cardinality of T. So, if we take $p = 2^n$ and $|T| = 2^{2^n}$ and denote $(\mathbf{E}|G_t - G_s|^p)^{1/p} = \|G_t - G_s\|_p$ we get that

$$g(T) \leq 4\sup_{s,t\in T}\|G_t - G_s\|_{2^n} \leq L2^{n/2}\sup_{s,t\in T} d_2(s,t) = 2^{n/2}\Delta_2(T),$$

where we used the fact (see e.g. [35, (2.11)]) $\|G_t\|_p \leq L\sqrt{p}\|G_t\|_2$ and denoted by $\Delta_2(\cdot)$ the diameter of the set with respect to distance d_2. Here and in the rest of this dissertation we denote by L a universal constant which might be different at each occurance. To define γ numbers we will need the definition of admissible sequence of partititions of T.

Definition 1. Let $N_n = 2^{2^n}$ for $n \geq 1$ and $N_0 = 1$ and consider T with some distance d. We will call nested sequence of partitions $(\mathcal{A}_n)_{n\geq 0}$ of set T admissible if it holds that $|\mathcal{A}_0| = 1$ together with $|\mathcal{A}_n| \leq N_n$ for $n \geq 1$, where by $|\cdot|$ we denote the cardinality.

By $A_n(t)$ we will denote the unique element of partition $\mathcal{A}_n$ that contains $t \in T$ and by $\Delta(\cdot)$ the diameter of set in distance d. Given $\alpha > 0$ define

$$\gamma_\alpha(T,d) = \inf \sup_{t\in T} \sum_{n\geq 0} 2^{n/\alpha}\Delta(A_n(t)),$$

where the infimum is taken over all admissible sequences. The Fernique-Talagrand bound (for historical reasons also referred as the Majorising Measure Theorem) states that there is a universal constant L such that

$$\frac{1}{L}\gamma_2(T,d_2) \leq g(T) \leq L\gamma_2(T,d_2), \tag{1.2}$$

where d_2 is an Euclidean distance on T. The definition of admissible sequence of partitions as well as of γ_2 number might look confusing at the first sight. In order to better understand the above inequality one should look closer into the idea of generic chaining (see [32, Chapter 2] or [35, Chapter 8]). The upper bound should be treated as a strengthening of a Dudley's entropy bound ([35, Theorem 8.1.3], [32, Proposition 2.2.10]). It follows by combining the union bound and the increment condition (see [32, (1.4)], [35, (2.10)]) given by

$$\forall u > 0, \quad \mathbf{P}(|G_t - G_s| \geq u) \leq 2\exp\left(-\frac{u^2}{d_2(s,t)^2}\right).$$

Equivalently, we can write already mentioned fact the $\|G_t - G_s\|_p$ is comparable with $\sqrt{p}\|G_t - G_s\|_2$. The above condition combined with the chaining method motivates the definition of γ_2 number. Method of obtaining upper bounds will be explained in Section 1.2. The lower bound is a result of the essential partitioning scheme which is a multiscale combination of Sudakov's minoration (see e.g. [35, Theorem 7.4.1] or [32, p. 2.4.2]) with concetration inequality ([32, Lemma 2.4.7]). The minoration is in some sense reverse of (1.1) i.e. consider $|T| = 2^{2^n}$ and $A > 0$. It states that

$$\text{if } \forall s,t \in T, \|G_t - G_s\|_{2^n} \geq A, \text{then } g(T) \geq L^{-1}A.$$

We will outline the generalized form of the partitioning scheme in Section 1.1.2. For the most insightful account of developments leading to the modern formulation of the Majorising Measure Theorem given in (1.2) see [32, Section 2.8]. The natural extension of the Gaussian case follows from the fact that separable

Gaussian process has a canonical representation given by the Karhunen-Loève representation ([19, Corollary 5.3.4])

$$G_t := \sum_{i\geq 1} t_i g_i,$$

where g_i's are independent standard normal variables. The problem then is to consider the processes generated by random variables other than g_i's, for example exponential or Weibull variables. This leads to the definition of canonical stochastic processes. For $T \subset \ell^2$ and $t \in T$ define independent symmetric random variables Y_i satisfying for $u > 0$ and $p \geq 1$, $\mathbf{P}(|Y_i| \geq u) \leq 2\exp(-u^p/2)$ and

$$X_t := \sum_{i\geq 1} t_i Y_i.$$

The increment condition can be then deduced from the following (see [33, Exercise 7.2.2]) for $p \leq 2$ and q such that $1/p + 1/q = 1$,

$$\mathbf{P}(|X_t| \geq u) \leq 2\exp\left(-\frac{1}{L}\min\left(\frac{u^2}{\sum_{i\geq 1} t_i^2}, \frac{u^p}{(\sum_{i\geq 1}|t_i|^q)^{p/q}}\right)\right),$$

while for $p > 2$

$$\mathbf{P}(|X_t| \geq u) \leq 2\exp\left(-\frac{1}{L}\max\left(\frac{u^2}{\sum_{i\geq 1} t_i^2}, \frac{u^p}{(\sum_{i\geq 1}|t_i|^q)^{p/q}}\right)\right).$$

Let us denote by d_∞ and d_p the distances induced by ℓ^∞ and ℓ^p norms respectively. As will become clear from Theorem 3 (see also [33, Chapter 7.2]) this implies that for $p = 1$,

$$\mathbf{E}\sup_{t\in T} X_t \leq L(\gamma_2(T, d_2) + \gamma_1(T, d_\infty)),$$

for $p \leq 2$,

$$\mathbf{E}\sup_{t\in T} X_t \leq L(\gamma_2(T, d_2) + \gamma_p(T, d_q))$$

and for $p > 2$, we get $\mathbf{E}\sup_{t\in T} X_t \leq L\gamma_2(T, d_2)$ and $\mathbf{E}\sup_{t\in T} X_t \leq L\gamma_p(T, d_q)$, which can be interpolated to get

$$\mathbf{E}\sup_{t\in T} X_t \leq L\inf\{(\gamma_2(T_1, d_2) + \gamma_p(T_2, d_q)) : \ T \subset T_1 + T_2\}.$$

The above inequalities can be reversed (see [33, Chapter 7.2], [32, Chapter 10.2]) which should be treated as a generalization of the Majorising Measure Theorem. We give the account of this result in Section 2.2.

A process of the crucial importance for our study is the Bernoulli process which can be seen as a limiting case of the previous example when $p = \infty$. First, by $(\varepsilon_i)_{i\geq 1}$ we denote a sequence of random signs (Bernoulli sequence) i.e. $\mathbb{P}(\varepsilon_i = \pm 1) = 1/2$. By the Bernoulli process we mean a collection of random variables

$(B_t)_{t\in T}$, where

$$B_t := \sum_{i\geq 1} t_i\varepsilon_i$$

and T is again a subset of ℓ^2. We consider a quantity

$$b(T) := \mathbf{E}\sup_{t\in T} B_t.$$

We can bound $b(T)$ from above as follows. On the one hand, we have the obvious upper bound given by $|B_t| \leq \|t\|_1 = \sum_{i\geq 1}|t_i|$ so that $b(T) \leq \sup_{t\in T}\|t\|_1$. On the other hand, we can observe that if we consider a canonical Gaussian process $G_t = \sum_{i\geq 1} t_i g_i$, where g_i's are independent standard normal variables independent also of ε_i's then by Jensen's inequality we have the following bound

$$g(T) = \mathbf{E}\sup_{t\in T} G_t = \mathbf{E}\sup_{t\in T}\sum_{i\geq 1} t_i\varepsilon_i|g_i| \geq \mathbf{E}\sup_{t\in T}\sum_{i\geq 1} t_i\varepsilon_i\mathbf{E}|g_i| = \sqrt{\frac{2}{\pi}}b(T).$$

With (1.2) in hand we see that $b(T)$ can be bounded from above by $\sup_{t\in T}\|t\|_1$ and $\gamma_2(T,d_2)$. Formally, since obviously if $T\subset T_1+T_2 = \{t^1+t^2 : t^1\in T_1, t^2\in T_2\}$ then $b(T)\leq b(T_1)+b(T_2)$ we can formulate the upper bound as

$$\begin{aligned} b(T) &\leq L\inf\{\sup_{t\in T_1}\|t\|_1 + g(T_2) : T\subset T_1+T_2\} \\ &\leq L\inf\{\sup_{t\in T_1}\|t\|_1 + \gamma_2(T_2) : T\subset T_1+T_2\}. \end{aligned}$$

The milestone result due to Bednorz and Latała states that the above bound can be reversed. It was conjectured by Fernique and the problem was open for about 25 years as a Bernoulli conjecture until the celebrated paper of mentioned authors (see [2]). The central theme of this dissertation will be a study of random series of functions which are conditionally Bernoulli processes and the main results obtained are consequences of this theorem.

Theorem 1. *Let $T\subset \ell^2$. There exists a universal constant L such that*

$$\inf\{\gamma_2(T_1,d_2) + \sup_{t\in T_2}\|t\|_1; T\subset T_1+T_2\} \leq Lb(T).$$

1.1.2 General partitioning scheme

In order to prove lower bounds such as in (1.2) or Theorem 1 we have to build appropriate sequence of partitions of the index set T. The crucial concept is to relate the geometry of T with the structure of the stochastic process by considering distances on T induced by this process together with Sudakov's minoration and concentration inequality. The main step when going beyond the Gaussian case is to consider the whole family of distances. We will now provide a general setting after [32, Chapter 10] and we will consider a specific examples in Chapter 4 and 5.

Consider a family of maps $(\varphi_j)_{j\in\mathbb{Z}}$ such that

$$\varphi_j : T\times T \to \mathbb{R}\cup\{\infty\}\ ,\ \varphi_{j+1}\geq\varphi_j\geq 0\ ,\ \varphi_j(s,t)=\varphi_j(t,s).$$

Define

$$B_j(t,c)=\{s\in T:\varphi_j(s,t)\leq c\}.$$

The idea which meets the geometry of the set T with the stochastic process are functionals which satisfy so-called growth condition (see [32, Section 2.3] for the introduction in the Gaussian case). Functionals we want to consider are non-decreasing maps from subsets of T to $\mathbb{R}^+$. The examples to bear in mind are $\gamma_2(T,d_2)$ or $\mathbf{E}\sup_{t\in T}X_t$. Let $F_{n,j}$ denote the functional and assume that

$$F_{n+1,j}\leq F_{n,j}\ ,\ F_{n,j+1}\leq F_{n,j}.$$

Definition 2. We say that functionals $F_{n,j}$ satisfy the growth condition for $r=2^{\kappa-3}$, $\kappa\in\mathbb{Z}$, if the following occurs. Consider any $j\in\mathbb{Z}$, any $n\geq 1$ and $m=N_n$. Consider any sets $(H_l)_{1\leq l\leq m}$ that are separated in the following sense: there exist points $u,t_1,\ldots,t_m$ in T for which $H_l\subset B_{j+2}(t_l,2^{n+\kappa})$ and

$$\forall l,l'\leq m,\ \ l\neq l',\ \ \varphi_{j+1}(t_l,t_{l'})\geq 2^{n+1}, \tag{1.3}$$

$$\forall l\leq m,\ \ t_l\in B_j(u,2^n). \tag{1.4}$$

Then,

$$F_{n,j}(\bigcup_{l\leq m}H_l)\geq 2^n r^{-j-1}+\min_{l\leq m}F_{n+1,j+1}(H_l). \tag{1.5}$$

To get the intuition behind this definition it is good to keep in mind the Sudakov's minoration (see [32, Lemma 2.4.2]) for Gaussian processes which states that if we have a well-controlled number of points which are separated i.e. $\forall l,l'\leq m,\ l\neq l'$ for some number m (think of 2^n) we have that $d_2(t_l,t_{l'})\geq a$ then

$$\mathbf{E}\sup_{l\leq m}G_{t_l}\geq\frac{a}{L}\sqrt{\log m}.$$

This can be generalized for subsets of well-separated balls (see [32, Proposition 2.4.9]). If points t_l and numbers m and a are as previously and if we consider subsets $H_l\subset B(t_l,\sigma)$ for $\sigma>0$ sufficiently smaller than a by denoting $H=\cup_{l\leq m}H_l$ we have

$$\mathbf{E}\sup_{t\in H}G_t\geq\frac{a}{L}\sqrt{\log m}+\min_{l\leq m}\mathbf{E}\sup_{t\in H_l}G_t,$$

which should be compared with (1.5). Indeed, let us look at the example, where $\varphi_j(s,t)=r^{2j}d(s,t)^2$ for a distance d on T. Notice that by denoting $B(t,r)$ the ball in distance d centred at t of radius r we have

$$B_j(t,c)=B(t,r^{-j}\sqrt{c}),$$

so the condition (1.3) means that

$$\forall l, l' \leq m, \;\; l \neq l', \;\; d(t_l, t_{l'}) \geq 2^{(n+1)/2} r^{-j-1} := a. \tag{1.6}$$

It can then be deduced (see discussion after [32, Definition 10.1.1]) that (1.6) reads

$$F_{n,j}(\bigcup_{l\leq m} H_l) \geq 2^{(n-1)/2} a + \min_{l \leq m} F_{n+1,j+1}(H_l).$$

This suggest that (1.5) is uniform in j with the term r^{-j-1} being the normalization factor. Also, it recovers the growth condition in the Gaussian case [32, Definition 2.7.1] and the version of Sudakov's minoration explained above for

$$F_{n,j}(H) = F(H) = g(H).$$

Let us also record an important fact that the functional given by $\gamma_2(T, d_2)$ satisfies the growth condition (see [32, Theorem 2.3.15]).
The general partitioning scheme is the following.

Theorem 2. *Let $F_{n,j}$ be functionals as above and assume that they satisfy the growth condition of Definition 2. Suppose that for some $j_0 \in \mathbb{Z}$ we have*

$$\forall s, t \in T, \; \varphi_{j_0}(s,t) \leq 1. \tag{1.7}$$

Then, there exists an admissible sequence of partitions $(\mathcal{A}_n)_{n\geq 0}$ and for each $A \in \mathcal{A}_n$ an integer $j(A)$ and a point $t_{n,A} \in T$ such that

$$A \in \mathcal{A}_n, C \in \mathcal{A}_{n-1}, A \subset C \implies j_{n-1}(C) \leq j_n(A) \leq j_{n-1}(C) + 1, \tag{1.8}$$

$$\forall t \in T, \; \sum_{n\geq 0} 2^n r^{-j_n(A_n(t))} \leq L(rF_{0,j_0}(T) + r^{-j_0}), \tag{1.9}$$

$$\forall n \geq 0, \; \forall A \in \mathcal{A}_n, \;\; A \subset B_{j_n(A)}(t_{n,A}, 2^n). \tag{1.10}$$

Looking back at the example $\varphi_j(s,t) = r^{2j} d(s,t)^2$, we see that (1.7) means that $\Delta(T) \leq r^{-j_0}$ while by (1.10) we have $\Delta(A) \leq 2r^{-j_n(A)} 2^{n/2}$, so by (1.9) we have

$$\forall t \in T, \; \sum_{n \geq 0} 2^{n/2} \Delta(A_n(t)) \leq L(rF_{0,j_0}(T) + r^{-j_0}).$$

Now, if we take j_0 to be the largest integer for which $\Delta(T) \leq r^{-j_0}$ the we conclude that

$$\gamma_2(T,d) \leq L(rF_{0,j_0} + \Delta(T))$$

which recovers the lower bound in Majorising Measure Theorem [33, Theorem 2.8.1] by defininig $F_{0,j_0} = g(T)$ and observing that $\Delta(T) \leq \sqrt{2\pi} g(T)$.
The reason why we only give a brief summary of the powerful tool given in Theorem 2, the proof of which is of separate interest, is the fact that in the setting we want to work in we will not be able to apply the general partitioning scheme. However, it will be helpful in refolmulating the lower bound of the Bernoulli process. We will be explain this in much more detail in Chapter 4.

1.1.3 Random series of functions

The main results of this dissertation concern lower bounds for infinitely divisible processes, empirical processes and selector processes. Each of these processes can be treated as conditionally Bernoulli process. The study of such processes is in some sense a natural next step of investigation after characterization of Bernoulli processes. Historically though it was the study of random Fourier sequence conveyed by X. Fernique (see [32, Chapter 3.3] and [2, Section 8]) which inspired the questions on Bernoulli process and the fact that symmetric random Fourier series can be interpreted as conditionally Bernoulli series played the main role in the Marcus-Pisier theorem (see [32, Theorem 3.2.1]).
A leitmotiv of the theory of suprema of stochastic processes developed and gathered by M. Talagrand in [32] is that the stochastic process which is conditionally a Bernoulli process can be decomposed into the positive part and the part which captures all possible cancellations between the terms and which can be explained by the chaining. This phenomenon was already present in the Bernoulli case where the bound $\sup_{t\in T_2}\|t\|_1$ should be perceived as the part of the process which owes nothing to cancellation while $\gamma_2(T_1, d_2)$ explains the part of the process which gathers all the cancellations.
It will be convenient to work in the general setting of random series of functions to formulate the main question. Let us consider an index set T, and a random sequence $(Z_i)_{i\geq 1}$ of functions on T. This sequence does not have to be independent (as we will see in Chapter 4). Consider an independent Bernoulli sequence $(\varepsilon_i)_{i\geq 1}$, which is independent of the sequence $(Z_i)_{i\geq 1}$. Define $X_t = \sum_{i\geq 1}\varepsilon_i Z_i(t)$. Our goal is to provide under certain conditions a lower bound of the quantity

$$S := \mathbf{E}\sup_{t\in T} X_t.$$

As we will outline in Sections 1.2 and 1.3 the question of upper bounds is significantly easier. Let us skip the usual measurability of S discussion now for the sake of exposure and come back to it in Chapter 2. The crucial property of the series X_t is that conditioned on the Z_i we are considering a Bernoulli process.
As introduced before, there are three important examples of processes which can be expressed as such a series. The first class are infinitely divisible processes which have series represantation due to Rosiński [23]. This case will be treated in detail in Chapter 4 and the main tools presented there wil be used for another two classes of processes which are empirical processes and selector processes.
The contents of the next sections follow closely the state of art gathered in [33] and [32]. They aim to present the upper bounds which are already known and which we want to reverse as well as very useful symetrization inequalities in the spirit of Giné-Zinn [10].

1.2 Tools for upper bounds

The following is a known result stating that if the increments of a centred process satisfy Bernstein's inequality, then we can control the supremum of this process from above. We restate it here after [33] and [32] together with the proof for completeness.

Theorem 3. *Consider a set T provided with two distances d_1 and d_2. Consider a centered process $(X_t)_{t\in T}$ which satisfies $\forall s,t \in T$, $\forall u > 0$*

$$\mathbf{P}\left(|X_s - X_t| \geq u\right) \leq 2\exp\left(-\min\left(\frac{u^2}{d_2(s,t)^2}, \frac{u}{d_1(s,t)}\right)\right). \tag{1.11}$$

Then

$$\mathbf{E}\sup_{s,t\in T}|X_s - X_t| \leq L(\gamma_1(T,d_1) + \gamma_2(T,d_2)).$$

First, we need the following result which belongs independently to Latała and Mendelson (see [32, Excercise 2.2.25], [33, Theorem 2.6.10] or [17]). Consider a distance $\delta_n(s,t) = \|X_s - X_t\|_{2^n}$ for $n \geq 0$, where by $\|X\|_p$ we denote the p-th norm of random variable X i.e. $(\mathbf{E}|X|^p)^{1/p}$. Let $D_n(\cdot)$ be a diameter in the distance δ_n. This distance will be important in Chapter 2 and another form of the next result will be provided in Section 2.2, we therefore omit the proof.

Theorem 4. *Consider an admissible sequence of partitions $(\mathcal{A}_n)_{n\geq 0}$ of partitions of T. Then,*

$$\mathbf{E}\sup_{s,t\in T}|X_s - X_t| \leq L\sup_{t\in T}\sum_{n\geq 0} D_n(A_n(t)). \tag{1.12}$$

We see that what we need to show is the following.

Theorem 5. *Suppose that (1.11) is satisfied. Then, there exists an admissible sequence of partitions $(\mathcal{A}_n)_{n\geq 0}$ such that*

$$\sup_{t\in T} D_n(A_n(t)) \leq L(\gamma_1(T,d_1) + \gamma_2(T,d_2)). \tag{1.13}$$

Proof. Denote by $\Delta_j(A)$ the diameter of the set A in the distance d_j, $j = 1,2$. Consider an admissible sequence $(\mathcal{B}_n)_{n\geq 0}$ of partitions of T such that

$$\forall t \in T, \sum_{n\geq 0} 2^n \Delta_1(B_n(t)) \leq 2\gamma_1(T,d_1) \tag{1.14}$$

and an admissible sequence $(\mathcal{C}_n)_{n\geq 0}$ such that

$$\forall t \in T, \sum_{n\geq 0} 2^{\frac{n}{2}} \Delta_2(C_n(t)) \leq 2\gamma_2(T,d_2), \tag{1.15}$$

where we denote by $B_n(t)$ and $C_n(t)$ the unique element of $\mathcal{B}_n$ and $\mathcal{C}_n$ respectively which contains t. We define the partition $\mathcal{A}_n$ by $\mathcal{A}_0 = \{T\}$ and for $n \geq 1$, $\mathcal{A}_n = \{B \cap C : B \in \mathcal{B}_{n-1},\ C \in \mathcal{C}_{n-1}\}$. Obviously, it is nested because both $(\mathcal{B}_n)$ and $(\mathcal{C}_n)$ are nested. Moreover, for $n \geq 1$, $|\mathcal{A}_n| \leq N_{n-1} \cdot N_{n-1} = N_n$, so $\mathcal{A}_n$ is admissible. Now, notice that for non-negative random variable Y satisfying for some $A, B > 0$ and all $u > 0$

$$\mathbf{P}(Y \geq u) \leq 2\exp\left(-\min\left(\frac{u^2}{A^2}, \frac{u}{B}\right)\right)$$

we have that $\|Y\|_p \leq L(A\sqrt{p} + Bp)$. Indeed we have

$$\mathbf{E}Y^p = \int_0^\infty pt^{p-1}\mathbf{P}(Y \geq t)dt \leq \int_0^\infty pt^{p-1}(\exp(-\frac{u^2}{A^2}) + \exp(-\frac{u}{B}))dt$$

It then remains to observe that $u^{p-1}\exp(-u^2) \leq Lp^{(p-1)/2}\exp(-u^2/2)$ and apply it for $u = t/A$ for the first integral and that $u^{p-1}\exp(-u) \leq Lp^{p-1}\exp(-u/2)$ and apply it for $u = t/B$ for the second integral. Having established this we see that $D_n(A_n(t)) \leq L(2^n\Delta_1(B_{n-1})) + 2^{n/2}\Delta_2(C_{n-1})$, so (1.13) follows by summing (1.14) and (1.15). □

Proof of Theorem 3. It now follows by combining (1.12) and Theorem 5. □

We will use the following form of Bernstein's inequality.

Theorem 6. *Let $(W_i)_{i\geq 1}$ be a sequence of independent random variables with $\mathbf{E}W_i = 0$ and for some number U, $|W_i| \leq U$ for each i. Then, for $v > 0$*

$$\mathbf{P}\left(\left|\sum_{i\geq 1} W_i\right| \geq v\right) \leq 2\exp\left(-\min\left(\frac{v^2}{4\sum_{i\geq 1}\mathbf{E}W_i^2}, \frac{v}{2U}\right)\right). \qquad (1.16)$$

Proof. We will argue using Chebychev's inequality, so we will need a bound for $\mathbf{E}\exp(\lambda|W_i|)$, $\lambda \in \mathbb{R}$. To this end, observe that for $|x| \leq 1$ we have

$$|e^x - 1 - x| \leq x^2\sum_{k\geq 2}\frac{1}{k!} = x^2(e-2) \leq x^2,$$

so since $\mathbf{E}W_i = 0$ for $U|\lambda| \leq 1$ we have

$$|\mathbf{E}\exp(\lambda W_i) - 1| \leq \lambda^2\mathbf{E}W_i^2.$$

So, $\mathbf{E}\exp(\lambda W_i) \leq 1 + \lambda^2\mathbf{E}W_i^2 \leq \exp(\lambda^2\mathbf{E}W_i^2)$ and

$$\mathbf{E}\exp\left(\lambda\sum_{i\geq 1} W_i\right) = \prod_{i\geq 1}\mathbf{E}\exp(\lambda W_i) \leq \exp\left(\lambda^2\sum_{i\geq 1}\mathbf{E}W_i^2\right).$$

Therefore, for $0 \leq \lambda \leq 1/U$

$$\begin{aligned}\mathbf{P}\left(\sum_{i\geq 1} W_i \geq v\right) &\leq \exp(-\lambda u)\exp\left(\lambda\sum_{i\geq 1}\mathbf{E}W_i\right)\\ &\leq \exp\left(\lambda^2\sum_{i\geq 1} W_i^2 - \lambda v\right).\end{aligned}$$

Finally, we optimize λ. Namely, if $Uv \le 2\sum_{i\ge1}\mathbf{E}W_i^2$ put $\lambda = v/(2\sum_{i\ge1}\mathbf{E}W_i^2)$ to get the bound $\exp(-v^2/(4\sum_{i\ge1}\mathbf{E}W_i^2))$. If $Uv > 2\sum_{i\ge1}\mathbf{E}W_i^2$, put $\lambda = 1/U$ and observe that

$$\frac{1}{U^2}\sum_{i\ge1}\mathbf{E}W_i^2 - \frac{v}{U} \le \frac{Uv}{2U^2} - \frac{v}{U} = -\frac{v}{2U}.$$

Therefore, we get the bound

$$\mathbf{P}\left(\sum_{i\ge1} W_i \ge v\right) \le \exp\left(-\min\left(\frac{v^2}{4\sum_{i\ge1}\mathbf{E}W_i^2}, \frac{v}{2U}\right)\right).$$

We get the same bound for $\mathbf{P}(\sum_{i\ge1} W_i \le -v)$ by changing W_i into $-W_i$. □

1.3 Upper bounds for empirical and selector processes

Let us now present two out of three main processes we want to study. We postpone introducing infinitely divisible processes until Chapter 4.

First example is constituted by empirical processes. We will follow a standard notation used in the study of such processes. Let (Ω, μ) be a probability space and $\mathcal{F}$ be a bounded, countable subset of $L^2(\mu)$ (it is the only instance when we do not denote the index set by T). Since $\mathcal{F}$ is countable we do not have to distinguish between actual functions on Ω and classes of functions in $L^2(\mu)$. Denote $\mu(f) = \int f d\mu$. Consider independent random variables $(X_i)_{i\le N}$ all distributed like μ. In this case the object of our study is

$$S_N(\mathcal{F}) = \mathbf{E}\sup_{f\in\mathcal{F}}\left|\sum_{i\le N}(f(X_i) - \mu(f))\right|. \tag{1.17}$$

The quantity $\sup_{f\in\mathcal{F}}|\sum_{i\le N}(f(X_i)-\mu(f))|$ is called discrepancy bound because since

$$\left|\sum_{i\le N}(f(X_i) - \mu(f))\right| = N\left|\frac{1}{N}\sum_{i\le N} f(X_i) - \mu(f)\right|$$

it measures the discrepancy between the true measure $\mu(f)$ and the 'empirical measure' $N^{-1}\sum_{i\le N} f(X_i)$. First, observe that

$$S_N(\mathcal{F}) \le 2\mathbf{E}\sup_{f\in\mathcal{F}}|f(X_i)|, \tag{1.18}$$

because

$$S_N(\mathcal{F}) \le \mathbf{E}\sup_{f\in\mathcal{F}}\sum_{i\le N}|f(X_i) - \mu(f)| \le \mathbf{E}\sup_{f\in\mathcal{F}}\sum_{i\le N}|f(X_i)| + N\sup_{f\in\mathcal{F}}\mu(f)$$

and by Jensen's inequality $N\sup_{f\in\mathcal{F}}\mu(f) \le \mathbf{E}\sup_{f\in\mathcal{F}}\sum_{i\le N}|f(X_i)|$. Moreover, we can now apply Bernstein's inequality (1.16) for $W_i = f(X_i) - \mu(f)$ if $i \le N$

and $W_i = 0$ if $i > 0$. Notice that $\mathbf{E}W_i = 0$, $|W_i| \le 2\sup|f| = 2\|f\|_\infty$ and $\mathbf{E}W_i^2 \le \mathbf{E}f^2(X_i) = \|f\|_2^2$. We get that

$$\mathbf{P}\left(\left|\sum_{i\le N}(f(X_i)-\mu(f))\right| \ge v\right) \le 2\exp\left(-\frac{1}{4}\min\left(\frac{v^2}{N\|f\|_2^2}, \frac{v}{\|f\|_\infty}\right)\right).$$

Finally, define $Z_f = \sum_{i\le N}(f(X_i)-\mu(f))$ and assume $0 \in \mathcal{F}$. Then, by Theorem 3

$$\begin{aligned}\mathbf{E}\sup_{f\in\mathcal{F}}|Z_f| \le \mathbf{E}\sup_{f,f'\in\mathcal{F}}|Z_f - Z_{f'}| &\le L(\gamma_2(\mathcal{F}, 2\sqrt{N}d_2) + \gamma_1(\mathcal{F}, 4d_\infty)) \\ &= L(2\sqrt{N}\gamma_2(\mathcal{F}, d_2) + 4\gamma_1(\mathcal{F}, d_\infty)), \quad (1.19)\end{aligned}$$

where d_2 and d_∞ are distances induced by $L^2(\mu)$ and $L^\infty(\mu)$ norms repsectively.
The second example is a selector process. For a fixed number $0 < \delta < 1$ we define a sequence of random variables $(\delta_i)_{i\le M}$ which is 1 with probability δ and 0 with probability $1-\delta$. These variables are called selectors becasue they allow to select a random subset of $\{1, 2, \ldots, M\}$ of cardinality equal about δM. By the selector process we understand the family of random variables $\sum_{i\le M} t_i(\delta_i - \delta)$, where t varies over the set T of sequences. Similarly as we have seen before we want to study the quantity

$$\delta(T) = \mathbf{E}\sup_{t\in T}\left|\sum_{i\le M} t_i(\delta_i - \delta)\right|. \quad (1.20)$$

First type of the control we have over $\delta(T)$ is again

$$\delta(T) \le \mathbf{E}\sup_{t\in T}\sum_{i\le M}|t_i|\delta_i. \quad (1.21)$$

To apply Bernstein's inequality define $W_i = t_i(\delta_i - \delta)$ for $i \le M$ and $W_i = 0$ for $i > M$. Obviously, $\mathbf{E}W_i = 0$. Moreover, $|W_i| = |t_i||\delta_i - \delta| \le 2\max_{i\le M}|t_i|$ and $\mathbf{E}W_i^2 = \delta(1-\delta)t_i^2 \le \delta t_i^2$, so (1.16) implies

$$\mathbf{P}\left(\left|\sum_{i\le M} t_i(\delta_i - \delta)\right| \ge v\right) \le 2\exp\left(-\frac{1}{4}\min\left(\frac{v^2}{\sum_{i\ge 1} t_i^2}, \frac{v}{\max_{i\le M}|t_i|}\right)\right).$$

As in the case of empirical processes Theorem 3 implies that

$$\delta(T) \le L(\sqrt{\delta}\gamma_2(T, d_2) + \gamma_1(T, d_\infty)), \quad (1.22)$$

where $d_2^2(s,t) = \sum_{i\le M}|t_i - s_i|^2$ and $d_\infty(s,t) = \max_{i\le M}|t_i - s_i|$.
It is good to see both of the presented cases as examples of a general case i.e. bounding

$$S = \mathbf{E}\sup_{t\in T}\sum_{i\le N}\varepsilon_i Z_i(t).$$

Let us assume that $Z_i = 0$ for $i > N$ and that Z_i are independent. They however do not have to be identically distributed. Let λ_i be the law of Z_i so that it is a probability measure on the space $\mathcal{F} = \mathbb{R}^T$ of functions on T. Let $\nu = \sum_{i\le N} \lambda_i$. We can therefore see the elements of T as functions on the measured space $(\mathcal{F}, \nu)$ and the two distances we want to use is d_2 and d_∞ induced by $L^2(\nu)$ and $L^\infty(\nu)$ norm respectively. What might be confusing is the fact that we use functions $Z_i(t)$ rather than $t(Z_i)$. The fact is we can use them interchangeably as will become clear in Chapter 5 (see also discussion after [32, Theorem 11.2.10]). After the two examples it should not be surprising that we can bound S not only by $\mathbf{E}\sup_{t\in T}\sum_{i\le N}|Z_i(t)|$ but also as follows

$$S \le L(\gamma_2(T, d_2) + \gamma_1(T, d_\infty)). \tag{1.23}$$

Reversing the bounds (1.18) and (1.19) in the case of empirical processes was a subject of [32, Research problem 9.1.3], while reversing the bounds (1.21) and (1.22) in the case of selector processes was a subject of [32, Conjecture 12.3.3] under the name of Generalized Bernoulli Conjecture. We will answer both of them positively in Chapter 5. The results will be formulated in the form of decomposition theorem similarly as for the Bernoulli process in Bednorz-Latała Theorem (Theorem 1). We will formulate the same upper bounds and decomposition for infinitely divisible processes in Chapter 4. At this stage it might not be clear that both of these results will folllow from the same approach. It might be even more unclear where do Bernoulli variables pop out in the whole story since they are not present in the definition of $S_N(\mathcal{F})$ nor $\delta(T)$. We will clarify this in the next section.

1.4 Symmetrization inequalities

We start this section with a crucial property of a Bernoulli process called a comparison principle. It will be studied in much more detail in Chapter 2. In its' simplest form it states what follows. For the proof see e.g. [32, Theorem 5.3.6].

Theorem 7. *Let θ_i's be contractions from $\mathbb{R}$ to $\mathbb{R}$ i.e. for each $i \ge 1$, $|\theta_i(s) - \theta_i(t)| \le |s-t|$ and $\theta_i(0) = 0$. Then for each finite subset T of ℓ^2 we have*

$$\mathbf{E}\sup_{t\in T}\sum_{i\ge 1}\theta_i(t_i)\varepsilon_i \le \mathbf{E}\sup_{t\in T}\sum_{i\ge 1} t_i\varepsilon_i. \tag{1.24}$$

We will need a following consequence of the comparison principle.

Corollary 1. *For each finite subset T of ℓ^2 we have that*

$$\mathbf{E}\sup_{t\in T}\left|\sum_{i\ge 1}|t_i|\varepsilon_i\right| \le 2\mathbf{E}\sup_{t\in T}\left|\sum_{i\ge 1} t_i\varepsilon_i\right|.$$

Proof. Denote by $x^+ = \max\{x, 0\}$, so that $|x| = x^+ + (-x)^+$. Hence, by symmetry

$$\mathbf{E}\sup_{t\in T}\left|\sum_{i\geq 1}|t_i|\varepsilon_i\right| \leq 2\mathbf{E}\sup_{t\in T}\left(\sum_{i\geq 1}|t_i|\varepsilon_i\right)^+ = 2\mathbf{E}\sup_{t\in T'}\sum_{i\geq 1}|t_i|\varepsilon_i,$$

where $T' = T\cup\{0\}$. By (1.24) applied for T' and the contraction given by $|\cdot|$ we get

$$\mathbf{E}\sup_{t\in T'}\sum_{i\geq 1}|t_i|\varepsilon_i \leq \mathbf{E}\sup_{t\in T'}\sum_{i\geq 1}t_i\varepsilon_i.$$

The result follows since $\mathbf{E}\sup_{t\in T'}\sum_{i\geq 1}t_i\varepsilon_i \leq |\mathbf{E}\sup_{t\in T}\sum_{i\geq 1}t_i\varepsilon_i|$. □

In order to see that S is the correct quantity to bound we need the following inequalities (cf.[33, Lemma 9.1.11]).

Theorem 8. *Let $S_N(\mathcal{F})$ be as in (1.17) and assume for each $f\in\mathcal{F}$, $\mu(f)=0$. Then,*

$$\mathbf{E}\sup_{f\in\mathcal{F}}\left|\sum_{i\leq N}\varepsilon_i f(X_i)\right| \leq 2S_N(\mathcal{F}). \tag{1.25}$$

For $\delta(T)$ as in (1.20) we have

$$\mathbf{E}\sup_{t\in T}\left|\sum_{i\leq M}\varepsilon_i t_i\delta_i\right| \leq 3\delta(T) \tag{1.26}$$

Proof. We start with (1.25). We will condition on the sequence $(\varepsilon_i)_{i\leq N}$. For this define $I = \{i\leq N:\ \varepsilon_i = 1\}$ and $J = \{i\leq N:\ \varepsilon_i = -1\}$, so that

$$\mathbf{E}\sup_{f\in\mathcal{F}}\left|\sum_{i\leq N}\varepsilon_i f(X_i)\right| \leq \mathbf{E}\sup_{f\in\mathcal{F}}\left|\sum_{i\in I}f(X_i)\right| + \mathbf{E}\sup_{f\in\mathcal{F}}\left|\sum_{i\in J}f(X_i)\right|.$$

Since $\mathbf{E}f(X_i) = 0$ for each $f\in\mathcal{F}$ and $i\leq N$ we can denote by $\mathbf{E}^I$ the expactation with respect to X_i's for $i\notin I$ and by Jensen's inequality deduce that

$$\sup_{f\in F}\left|\sum_{i\in I}f(X_i)\right| = \sup_{f\in F}\left|\mathbf{E}^I\sum_{i\leq N}f(X_i)\right| \leq \mathbf{E}^I\sup_{f\in F}\left|\sum_{i\leq N}f(X_i)\right|.$$

We get the same inequality for J and get (1.25) by taking expactions of sides and summating.
Now, we move to the proof of (1.26). First, observe that by triangle inequality

$$\mathbf{E}\sup_{t\in T}\left|\sum_{i\leq M}\varepsilon_i t_i\delta_i\right| \leq \mathbf{E}\sup_{t\in T}\left|\sum_{i\leq M}\varepsilon_i t_i(\delta_i-\delta)\right| + \delta\mathbf{E}\left|\sum_{i\leq M}\varepsilon_i t_i\right|. \tag{1.27}$$

Consider an independent copy of the sequence $(\delta_i)_{i\le M}$ given by $(\delta_i')_{i\le M}$ independent also of ε_i's. By the triangle inequality we have

$$\mathbf{E}\sup_{t\in T}\left|\sum_{i\le M} t_i(\delta_i-\delta_i')\right| \le 2\delta(T). \tag{1.28}$$

Observe that by Jensen's inequality we have

$$\mathbf{E}\sup_{t\in T}\left|\sum_{i\le M} t_i|\delta_i-\delta_i'|\right| \ge \mathbf{E}\sup_{t\in T}\left|\sum_{i\le M} t_i\mathbf{E}|\delta_i-\delta_i'|\right|.$$

Since $\mathbf{E}|\delta_i-\delta_i'| = 2\delta$ and $(\varepsilon_i|\delta_i-\delta_i'|)_{i\le M}$ has the same distribution as $(\varepsilon_i(\delta_i-\delta_i'))_{i\le M}$ and $(\delta_i-\delta_i')_{i\le M}$ by (1.28) we can write

$$2\delta\mathbf{E}\left|\sum_{i\le M}\varepsilon_i t_i\right| \le \mathbf{E}\sup_{t\in T}\left|\sum_{i\le M} t_i|\delta_i-\delta_i'|\right| \le 2\delta(T). \tag{1.29}$$

Again, by Jensen's inequality we get

$$\mathbf{E}\sup_{t\in T}\left|\sum_{i\le M}\varepsilon_i t_i(\delta_i-\delta)\right| \le \mathbf{E}\sup_{t\in T}\left|\sum_{i\le M}\varepsilon_i t_i(\delta_i-\delta_i')\right| \le 2\delta(T). \tag{1.30}$$

Combining (1.27), (1.29) and (1.30) yields (1.26). □

Both (1.25) and (1.26) suggest that we want to consider $t(Z_i)$ given by $f(X_i)$ in the empirical case and by $\delta_i t_i$ in the selector case. The importance of next results which are Giné-Zinn type of inequalities [10] will become apparent along the way of proving the Decomposition Theorems.

Theorem 9. *We have*

$$\mathbf{E}\sup_{f\in\mathcal{F}}\sum_{i\le N}|f(X_i)| \le N\sup_{f\in\mathcal{F}}\mu(f) + 4\mathbf{E}\sup_{f\in\mathcal{F}}\left|\sum_{i\le N}\varepsilon_i f(X_i)\right|. \tag{1.31}$$

and

$$\mathbf{E}\sup_{t\in T}\sum_{i\le M}\delta_i|t_i| \le \delta\sup_{t\in T}\sum_{i\le M}|t_i| + 12\delta(T). \tag{1.32}$$

Proof. First,

$$\sum_{i\le N}|f(X_i)| \le \sum_{i\le N}\mathbf{E}|f(X_i)| + \left|\sum_{i\le N}(|f(X_i)|-\mathbf{E}|f(X_i)|)\right|$$

So,

$$\mathbf{E}\sup_{f\in\mathcal{F}}\sum_{i\le N}|f(X_i)| \le \sup_{f\in\mathcal{F}}\sum_{i\le N}\mathbf{E}|f(X_i)| + \mathbf{E}\sup_{f\in\mathcal{F}}\left|\sum_{i\le N}(|f(X_i)|-\mathbf{E}|f(X_i)|)\right| \tag{1.33}$$

Observe that $\sum_{i\le N}\mathbf{E}|f(X_i)| = N\mu(f)$. To deal with the second term consider an independent copy $(X_i')_{i\le N}$ of $(X_i)_{i\le N}$ and note that $(|f(X_i) - |f(X_i')|)$ has the same distribution as $\varepsilon_i(|f(X_i) - |f(X_i')|)$. So, by Jensen's inequality

$$\begin{aligned}\mathbf{E}\sup_{f\in\mathcal{F}}\left|\sum_{i\le N}(|f(X_i)| - \mathbf{E}|f(X_i)|)\right| &\le \mathbf{E}\sup_{f\in\mathcal{F}}\left|\sum_{i\le N}(|f(X_i)| - |f(X_i')|)\right| \\ &= \mathbf{E}\sup_{f\in\mathcal{F}}\left|\sum_{i\le N}\varepsilon_i(|f(X_i)| - |f(X_i')|)\right|.\end{aligned}$$

By triangle inequality

$$\mathbf{E}\sup_{f\in\mathcal{F}}\left|\sum_{i\le N}\varepsilon_i(|f(X_i)| - |f(X_i')|)\right| \le 2\mathbf{E}\sup_{f\in\mathcal{F}}\left|\sum_{i\le N}\varepsilon_i|f(X_i)|\right|.$$

The last step is to apply Corollary 1 for a given value of X_i to obtaian

$$\mathbf{E}\sup_{f\in\mathcal{F}}\left|\sum_{i\le N}\varepsilon_i|f(X_i)|\right| \le 2\mathbf{E}\sup_{f\in\mathcal{F}}\left|\sum_{i\le N}\varepsilon_i f(X_i)\right|.$$

The same argument applies to (1.32). We write

$$\mathbf{E}\sup_{t\in T}\sum_{i\le M}\delta_i|t_i| \le \delta\mathbf{E}\sup_{t\in T}\sum_{i\le M}|t_i| + \mathbf{E}\sup_{t\in T}\left|\sum_{i\le M}(\delta_i - \delta)|t_i|\right|. \tag{1.34}$$

Introduce again $(\delta_i')_{i\le M}$ the independent copy of $(\delta_i)_{i\le M}$ and the Bernoulli sequence $(\varepsilon_i)_{i\le M}$ independent of both sequences. By Jensen's inequality and the fact that $(\varepsilon_i(\delta_i - \delta_i'))_{i\le M}$ has the same distribution as $(\delta_i - \delta_i')_{i\le M}$ we get

$$\mathbf{E}\sup_{t\in T}\left|\sum_{i\le M}(\delta_i - \delta)|t_i|\right| \le \mathbf{E}\sup_{t\in T}\left|\sum_{i\le M}(\delta_i - \delta_i')|t_i|\right| = \mathbf{E}\sup_{t\in T}\left|\sum_{i\le M}\varepsilon_i(\delta_i - \delta_i')|t_i|\right|.$$

Now,

$$\mathbf{E}\sup_{t\in T}\left|\sum_{i\le M}\varepsilon_i(\delta_i - \delta_i')|t_i|\right| \le 2\mathbf{E}\sup_{t\in T}\left|\sum_{i\le M}\varepsilon_i\delta_i|t_i|\right| \le 4\mathbf{E}\sup_{t\in T}\left|\sum_{i\le M}\varepsilon_i\delta_i t_i\right|,$$

where in the last inequality we applied Corollary 1 for given value of δ_i. We use this bound in (1.34) and conclude (1.32) by (1.26). □

1.5 Overview of the following chapters

The heart of this dissertation lies in Chapters 4 and 5, where the three hypotheses posed in [32] concerning infinitely divisible processes, empirical processes and selector processes are settled affirmatively. They can be read straight after the Introduction. Chapter 4 comes from [4]. Chapters 2 and 3 are devoted to study some further properties of the Bernoulli process. Chapter 2 deals with a question posed by K. Oleszkiewicz about comparability of weak and strong moments for Bernoulli series in a Banach space. It provides a partial answer to it and is related to some extension of the comparison principle of a Bernoulli sequence presented in (1.24), which is based on the proof of the Bernoulli Theorem (Theorem 1). The contents of Chapter 2 were published in Bulletin of Polish Academy of Science Mathematics [5]. In Chapter 3 we deal with question posed by W. Szatzschneider about a Lévy-Ottaviani type inequality for a Bernoulli series with monotone coefficients [26]. The contents of Chapter 3 were published in Statistics and Probability Letters [3].

1.6 Acknowledgements

I would like to express my sincere gratitude to my supervisor Witold Bednorz for all his devotion and patience in explaining and clarifying to me the all different parts of the theory of suprema of stochastic processes.
Also, I would like to thank prof. Michel Talagrand whose results on the Bernoulli process and infintely divisible processes play fundamental role in this work. I am most grateful to him for making the contents of the new edition of the book 'Upper and Lower Bounds of Stochastic Processes' available to me as well as for all the comments on the initial version of the obtained results.
Next, I would like to thank all the Probability group in our Faculty for a great work and study environment.
Finally, I acknowledge the National Science Center in Poland for supporting my research with the grant 'Preludium' NCN Grant UMO-2018/31/N/ST1/03982.

Chapter 2

Comparison problem

2.1 Introduction and notation

Throughout this chapter we will use the following notation. For a set A the number of elements in A will be denoted as $|A|$. If $t = (t_i)_{i\geq 1}$ is a sequence of real numbers and $p \geq 1$ then $\|t\|_p = (\sum_{i=1}^{\infty} |t_i|^p)^{\frac{1}{p}}$ and ℓ^p is the space of all sequences t with $\|t\|_p < \infty$. If $S, T \subset \ell^p$ then $S+T = \{s+t : s \in S, t \in T\}$. For a random variable ξ and $p > 0$ we put $\|\xi\|_p = (\mathbf{E}|\xi|^p)^{\frac{1}{p}}$. If $(\xi_i)_{i\geq 1}$ is a sequence of independent, identically distributed random variables such that $\mathbf{E}\xi_i = 0$, $\mathbf{E}\xi_i^2 = 1$ and $t = (t_i)_{i\geq 1} \in \ell^2$ then the random variable

$$X_t = \sum_{i=1}^{\infty} t_i \xi_i \tag{2.1}$$

is well-defined. For each $T \subset \ell^2$ with $0 \in T$ the process $X_T = (X_t)_{t\in T}$ is called canonical. The above series converges in $\|\cdot\|_2$, i.e.

$$\lim_{n\to\infty} \|\sum_{i=1}^{n} t_i \xi_i - X_t\|_2 = 0.$$

Clearly,

$$\|X_t - X_s\|_2 = \|t - s\|_2, \quad \text{for} \quad s, t \in T.$$

Remark 1. The almost sure convergence in (2.1) might be guaranteed also when the independence assumption on ξ_i's is skipped. In that case we may consider a finite-dimensional version of (2.1), where $T \subset \mathbb{R}^d$. The most studied example is when ξ_i's have log-concave tails, i.e. $\mathbf{P}(|\xi_i| > t) = \exp(-N_i(t))$ for $N_i : [0,\infty] \to [0,\infty]$ convex, and may be dependent.

We want to distinguish two types of canonical processes which will be of special interest. If $(\xi_i) = (\varepsilon_i)$ and $\mathbf{P}(\varepsilon_i = 1) = \mathbf{P}(\varepsilon_i = -1) = \frac{1}{2}$ then the process X_T is called canonical Bernoulli and denoted by $B_T = (B_t)_{t\in T}$. As mentioned before this class of processes is of fundamental importance for further applications (see Chapthers 4 and 5). If $(\xi_i) = (g_i)$ and g_i are normally $\mathcal{N}(0,1)$ distributed then the process X_T is called canonical Gaussian and denoted by $G_T = (G_t)_{t\in T}$. In fact, canonical Gaussian processes can be seen as a motivation to study canonical processes in general, the reason being the Karhunen–Loève representation of separable Gaussian processes by means of canonical Gaussian processes (see e.g. [19, Corollary 5.3.4]).

The main object studied will be suprema of canonical processes. For any set T and a stochastic process $(X_t)_{t\in T}$ we define

$$S_X(T) = \sup_{F\subset T} \mathbf{E} \sup_{t\in F} X_t,$$

where F runs through all finite subsets of T. Usually, by considering a separable modification of X_t, $t \in T$, it is possible to guarantee that $\sup_{t\in T} X_t$ is a well-defined random variable (for the definition of a separable version of a process and a discussion of measurability of suprema in the general setting of not necessarily separable Banach spaces see [18, Ch. 2]). In this case $S_X(T)$ coincides with the usual expectation of the supremum of X_t, i.e.

$$S_X(T) = \mathbf{E} \sup_{t\in T} X_t.$$

Let us finish this section with a few important technicalities which will be helpful in dealing with canonical processes. We have $S_X(T) = S_X(T-t)$, where $T - t = \{s - t : s \in T\}$, so we may always require that $0 \in T$. Moreover, $S_X(T) = S_X(\operatorname{Conv} T)$ and $S_X(T) = S_X(\operatorname{cl} T)$, where $\operatorname{Conv} T$ is the convex hull of T and $\operatorname{cl} T$ is the closure of T in ℓ^2.

2.2 Suprema of canonical processes via chaining

First, we recall the basics of the chaining approach to upper bounds for stochastic processes. Let (T, d) be a separable metric space, $t_0 \in T$ a fixed element and X_t, $t \in T$ a process such that $||X_s - X_t|| \leq d(t, s)$ for $t, s \in T$. For each countable, dense $D \subset T$ it is true that $\mathbf{E} \sup_{t\in D} X_t = S_X(T)$. Recall the Definition 1 of a sequence $\mathcal{A} = (\mathcal{A}_n)_{n\geq 0}$ of partitions of T being admissible. Also, these partitions are nested, i.e. for any $A \in \mathcal{A}_n$, $n \geq 1$ there is $B \in \mathcal{A}_{n-1}$ such that $A \subset B$. For $t \in T$ we denote by $A_n(t)$ the unique element of the partition $\mathcal{A}_n$ which contains t. A sequence $\pi = (\pi_n)_{n\geq 0}$ of mappings $\pi_n : T \to T$ is said to be adapted to the partitions $(\mathcal{A}_n)_{n\geq 0}$ if $\pi_n(t) = \pi_n(s)$ for $s, t \in A \in \mathcal{A}_n$, $n \geq 0$ and $\pi_0(t) = t_0$ for $t \in T$ (the common value of π_n on $A \in \mathcal{A}_n$ will be denoted by $\pi_n(A)$). Let $T_n = \{\pi_n(t) : t \in T\}$ and $D = \bigcup_n T_n$. We say that π_n is regular if

$$\lim_{n\to\infty} d(t, \pi_n(t)) = 0 \text{ for each } t \in T. \tag{2.2}$$

For regular π the set D is a dense subset in T. We define

$$\gamma_X(\pi) = \sup_{t\in T} \sum_{n=1}^{\infty} \|X_{\pi_n(t)} - X_{\pi_{n-1}(t)}\|_{2^n}$$

$$\gamma_X(T) = \inf \gamma_X(\pi)$$

where the infimum is taken over all admissible sequence of partitions $(\mathcal{A}_n)_{n\geq 0}$ of T and regular sequence π_n, $n \geq 0$, of mappings T adapted to $(\mathcal{A}_n)_{n\geq 0}$. Furthermore, for $t \in T$ and each $m > 1$ we can write the following chain representation

$$X_{\pi_m(t)} = X_{\pi_0(t)} + \sum_{n=1}^{m} (X_{\pi_n(t)} - X_{\pi_{n-1}(t)}).$$

Therefore for each $t \in D$ we have

$$X_t \leq X_{t_0} + \sum_{n=1}^{\infty} |X_{\pi_n(t)} - X_{\pi_{n-1}(t)}| \tag{2.3}$$

$\gamma_X(T)$ should be compared with Talagrand's functional $\gamma_2(T,d)$ (see [32, Definition 2.2.19] or the formulation due to Latała and Mendelson (see [15], [20]), where it was proved that under suitable regularity assumptions, $S_X(T) \leq K\gamma_X(T)$, where K is a universal constant. Let us give a short argument for a similar upper bound with an improved constant.

Theorem 10. *Under the above assumption on (T,d), $t_0 \in T$ and a process $(X_t)_{t\in T}$ we have*

$$S_X(T) \leq \mathbf{E}X_{t_0} + 3\gamma_X(T).$$

Proof. Let $(\mathcal{A}_n)_{n\geq 0}$ be any admissible sequence of partitions of T and $\pi = (\pi_n)_{n\geq 0}$ a regular and adapted to $(\mathcal{A}_n)_{n\geq 0}$ sequence of mappings of T. For any $A \in \mathcal{A}_n$ and $n \geq 1$ we denote by A' the unique $A' \in \mathcal{A}_{n-1}$ which contains A. In what follows we will use that for any $a, b > 0$ $a \leq b(1 + (\frac{a}{b} - 1)_+)$. We get for each $t \in D$ that

$$\begin{aligned} X_t - X_{t_0} &\leq \sum_{n=1}^{\infty} |X_{\pi_n(t)} - X_{\pi_{n-1}(t)}| \\ &\leq \sum_{n=1}^{\infty} 2\|X_{\pi_n(t)} - X_{\pi_{n-1}(t)}\|_{2^n}\big(1 + (\frac{|X_{\pi_n(t)} - X_{\pi_{n-1}(t)}|}{2\|X_{\pi_n(t)} - X_{\pi_{n-1}(t)}\|_{2^n}} - 1)_+\big) \\ &\leq \big(\sum_{n=1}^{\infty} 2\|X_{\pi_n(t)} - X_{\pi_{n-1}}(t)\|_{2^n}\big)\big(1 + \sum_{n=1}^{\infty}(\frac{|X_{\pi_n(t)} - X_{\pi_{n-1}(t)}|}{2\|X_{\pi_n(t)} - X_{\pi_{n-1}(t)}\|_{2^n}} - 1)_+\big). \end{aligned}$$

The last inequality holds true since $\sum_{n=1}^{\infty} b_n(1 + c_n) \leq (\sum_{n=1}^{\infty} b_n)(1 + \sum_{n=1}^{\infty} c_n)$ for each sequnences $(a_n), (c_n)$ of nonegative real numbers. Hence, it follows that

$$\begin{aligned} &\sup_{t\in D} X_t - X_{t_0} \\ &\leq \big(\sup_{t\in T} \sum_{n=1}^{\infty} 2\|X_{\pi_n(t)} - X_{\pi_{n-1}(t)}\|_{2^n}\big)\big(1 + \sum_{n=1}^{\infty} \sup_{t\in T}(\frac{|X_{\pi_n(t)} - X_{\pi_{n-1}(t)}|}{2\|X_{\pi_n(t)} - X_{\pi_{n-1}(t)}\|} - 1)_+\big) \\ &\leq 2\gamma_X(\pi)\big(1 + \sum_{n=1}^{\infty} \sum_{A\in\mathcal{A}_n} (\frac{|X_{\pi_n(A)} - X_{\pi_{n-1}(A')}|}{2\|X_{\pi_n(A)} - X_{\pi_{n-1}(A')}\|_{2^n}} - 1)_+\big). \end{aligned}$$

We easily see that if $t \in A \in \mathcal{A}_n$ then

$$X_{\pi_n(t)} - X_{\pi_{n-1}(t)} = X_{\pi_n(A)} - X_{\pi_{n-1}(A')}.$$

Therefore, we obtain

$$\mathbf{E}\sup_{t\in D} X_t \le \mathbf{E}X_{t_0} + 2\gamma_X(\pi)\big(1+\sum_{n=1}^{\infty}\sum_{A\in\mathcal{A}_n}\mathbf{E}\frac{|X_{\pi_n(A)} - X_{\pi_{n-1})(A')}|}{2\|X_{\pi_n(A)} - X_{\pi_{n-1}(A')}\|_{2^n}} - 1)_+\big). \quad (2.4)$$

Now, we show that for any nonnegative random variable θ and $p > 2$ we have

$$\mathbf{E}(\frac{\theta}{2\|\theta\|_p} - 1)_+ \le \frac{1}{2}\cdot\frac{1}{p\cdot 2^p}.$$

Obiviously, it is enough to prove that if $\theta \ge 0$ and $\mathbf{E}\theta^p \le 1$ then $\mathbf{E}(\frac{\theta}{2} - 1)_+ \le \frac{1}{2}\cdot\frac{1}{p\cdot 2^p}$. Indeed, if θ is as above and $\xi = \mathbf{E}(\theta|\mathcal{G})$ where $\mathcal{G}$ is the σ field genertated by the single event $C = \{\theta \ge 2\}$ then $\mathbf{E}(\frac{\xi}{2} - 1)_+ = \mathbf{E}(\frac{\theta}{2} - 1)_+$ and by Jensen inequality $\mathbf{E}\xi^p \le \mathbf{E}\theta^p \le 1$. Observe that for the random variable ξ we have $\xi = x\mathbb{1}_C$ for some $x > 2$ with $\mathbf{P}(C) \le \frac{1}{x^p}$. Hence

$$\mathbf{E}(\frac{\theta}{2} - 1)_+ = \mathbf{E}(\xi) \le \max_{x>2}\frac{1}{x^p}(\frac{x}{2} - 1) = (1 - \frac{1}{p})^{p-1}\frac{1}{p\cdot 2^p} \le \frac{1}{2}\cdot\frac{1}{p\cdot 2^p}.$$

Therefore, applying the above inequality for $p = 2^n$ for each $n \ge 1$ the inequality (2.4) yields

$$\mathbf{E}\sup_{t\in D} X_t - \mathbf{E}X_{t_0} \le 2\gamma_X(\pi)(1 + \frac{1}{2}\cdot\sum_{k=1}^{\infty}\frac{N_k}{2^k 2^{2^k}}) \le 3\gamma_X(\pi).$$

Hence taking the infimum over all admissible partitions together with regular and adapted sequences π we conclude the proof.
The same proof as above gives the estimate

$$\mathbf{E}\sup_{t\in T}|X_t| \le \mathbf{E}|X_{t_0}| + 3\gamma_X(T)$$

□

Lemma 1. *Let $X = (X_t), t \in \ell^2$ be a cannonical process and C_1 a constant such that*

$$\|X_t\|_{2^{n+1}} \le C_1\|X_t\|_{2^n}, \forall t \in \ell^2, n \ge 0. \quad (2.5)$$

Then, for all $T_1, T_2 \subset \ell^2$.

$$\gamma_X(T_1 + T_2) \le C_1(\gamma_X(T_1) + \gamma_X(T_2)). \quad (2.6)$$

For canonical Bernoulli and canonical Gaussian processes the above inequality holds with $C_1 = \sqrt{3}$.

Proof. Let for $i = 1, 2$ $(\mathcal{A}_n^i)$ be an admisible sequence of partitions of T_i, together with an adapted and regular sequence $\pi^i = (\pi_n^i)$ of mappings of T_i. We have to construct an admissible sequence of partitions for T_1+T_2 with an associated sequnece of mappings. For that purpose for each $t\in T_1+T_2$ let us choose and fix $t^1\in T_1, t^2\in T_2$ such that $t=t^1+t^2$. For $A^1\subset T_1$, $A^2\subset T_2$ define

$$A^1 * A^2 = \{t\in T_1+T_2\ :\ t^1\in A^1\ ,\ t^2\in A^2\}.$$

Define $\mathcal{B}_n$ in the following way. $\mathcal{B}_0=\{T_1+T_2\}$ and for $n\geq 0$ let $\mathcal{B}_{n+1}$ consist of sets $A^1 * A^2$, where $A^1\in\mathcal{A}^1_n$ and $A^2\in\mathcal{A}^2_n$. It is easy to see that $(\mathcal{B}_n)$ is an admissible sequence of partitions of T_1+T_2. Indeed, $|\mathcal{B}_{n+1}|\leq N_n\cdot N_n\leq N_{n+1}$, the sequence is clearly nested and for $t=t^1+t^2$ we have $B_n(t)=A_n(t^1)*A'_n(t^2)$. Now define the sequence $\pi=(\pi_n)_{n\geq 0}$ by $\pi_0(t)=0$ and $\pi_{n+1}(t)=\pi^1_n(t^1)+\pi^2_n(t^2)$ for $t\in T^1+T^2$. Obviously the sequence (π_n) is regular and adapted to the partition $(\mathcal{B}_n)$. Furthermore,

$$\|t-\pi_{n+1}(t)\|_2\leq \|t^1-\pi^1_n(t^1)\|_2+\|t^2-\pi^2_n(t^2)\|_2,$$

so $\lim_{n\to\infty}\|t-\pi_{n+1}(t)\|_2=0$. In this way we guarantee the regularity condition (2.2) for the sequence of mappings π_n adapted to the partition $\mathcal{B}_n$.
We need to show that for fixed $t\in T_1+T_2$

$$\sum_{n=1}^{\infty}\|X_{\pi_n(t)}-X_{\pi_{n-1}(t)}\|_{2^n}\leq C_1(\gamma_X(T_1)+\gamma_X(T_2)). \tag{2.7}$$

By the above construction and triangle inequality we get

$$\|X_{\pi_{n+1}(t)}-X_{\pi_n(t)}\|_{2^{n+1}}\leq \|X_{\pi^1_n(t^1)}-X_{\pi^1_{n-1}(t^1)}\|_{2^{n+1}}+\|X_{\pi^2_n(t^2)}-X_{\pi^2_{n-1}(t^2)}\|_{2^{n+1}},$$

so by (2.5),

$$\|X_{\pi_{n+1}(t)}-X_{\pi_n(t)}\|_{2^n}\leq C_1(\|X_{\pi^1_n(t^1)}-X_{\pi^1_{n-1}(t^1)}\|_{2^n}+\|X_{\pi^2_n(t^2)}-X_{\pi^2_{n-1}(t^2)}\|_{2^n})$$

and consequently we sum over $n\geq 1$ to obtain (2.7). That (2.5) holds with constant $\sqrt{3}$ for both canonical Bernoulli and Gaussian processes is a result known as the hypercontractivity (cf. [1], [13, Chapter 3.4], [9, Chapter 13]), which states that for $1<q<p<\infty$ we have $\|B_t\|_p\leq\sqrt{\frac{p-1}{q-1}}\|B_t\|_q$ and the same inequality for G_t. □

Let us summarize the available information about the processes for which a full characterization of the supremum (i.e. lower and upper bounds) can be provided with the use of $\gamma_X(T)$. To make the notation consistent within this chapter we will denote the expectations of suprema of Gaussian process and Bernoulli process by $S_G(T)$ and $S_B(T)$ respectively rather than $g(T)$ and $b(T)$. The seminal result of Fernique and Talagrand known as the Majorizing Measure Theorem (see [8], [28] or [32] for a modern formulation) is equivalent to the statement that $S_G(T)$ is comparable with $\gamma_G(T)$ up to a numerical constant. In [31] it was proved that $S_X(T)$ is comparable with a quantity which, in a sense, is equivalent to $\gamma_X(T)$ for a canonical process generated by ξ's which are symmetric and satisfy $\mathbf{P}(|\xi|>t)=\exp(-c_pt^p)$ for a fixed $p\in[1,2]$. A similar result holds for $p > 2$, yet it is only possible to show that there exists a set $T' \subset \ell^2$ (which may significantly differ from T) such that $S_X(T)$ is comparable with $\gamma_X(T')$ up to a numerical constant. Note that the limiting case when $p \to \infty$ is the case of canonical Bernoulli processes. Later, the idea of [31] was slightly generalized by R. Latała [16] to canonical processes generated by ξ with log-concave tails, yet under specific regularity assumptions. Finally, in [17] it was proved that it suffices to assume only certain conditions on the moment growth of ξ. Unfortunately, this result still does not apply to Bernoulli processes.

Let us study the characterization of $S_B(T)$ settled in [2] (see Theorem 1). Firstly, we want to look at canonical Bernoulli processes from the perspective of distances relating to some properties of Bernoulli-type random variables. We have (see [11], [21] and [12] for the formulation below) for any $p \in \mathbb{N}$, $p \geq 1$,

$$\|B_t\|_p \leq \sum_{i=1}^{p} |t_i^*| + \sqrt{p}(\sum_{i>p} |t_i^*|^2)^{\frac{1}{2}} \leq 4\|B_t\|_p, \tag{2.8}$$

where $(t_i^*)_{i\geq 1}$ is the rearrangement of $(t_i)_{i\geq 1}$ such that $|t_1^*| \geq |t_2^*| \geq \ldots$. Equivalently, we can express this relation as

$$\frac{1}{4} \inf_{t=t^1+t^2} (\|t^1\|_1 + \|G_{t^2}\|_p) \leq \|B_t\|_p \leq \inf_{t=t^1+t^2} (\|t^1\|_1 + \|G_{t^2}\|_p)$$

for $p \geq 2$. This motivates the following interpretation. If we denote by $I \subset \mathbb{N}$ some index set, we can think of (2.8) as a decomposition of the norm $\|B_t\|_p$ into the ℓ^1 part

$$\sum_{i=1}^{p} |t_i^*| = \sup_{|I^c|\leq p} \sum_{i\in I^c} |t_i|$$

and the Gaussian part

$$\sqrt{p}(\sum_{i>p} |t_i^*|^2)^{\frac{1}{2}} = \sqrt{p} \inf_{|I^c|\leq p} (\sum_{i\in I} |t_i|^2)^{\frac{1}{2}}.$$

In fact, a characterization similar to (2.8) can be formulated for a broad class of processes, namely processes with log-concave distributions. In particular, in [15] there is a characterization of $\|X_t - X_s\|_p$ for canonical processes based on one-unconditional log-concave random variables. As already mentioned, the characterization of $S_B(T)$ was known as the Bernoulli conjecture and was finally proved in [2]. It states that similarly to (2.8), $S_B(T)$ can be decomposed into the Gaussian and ℓ^1 parts. More precisely, there must exist a decomposition of T into $T_1, T_2 \subset \ell^2$ such that $T_1 + T_2 \supset T$ and moreover $S_B(T)$ dominates up to a universal constant both $\sup_{t\in T_1} \|t\|_1$ and $S_G(T_2)$. Usually such a decomposition is formulated in terms of existence of a mapping $\pi : T \to \ell^2$ which defines $T_1 = \{t - \pi(t) : t \in T\}$ and $T_2 = \{\pi(t) : t \in T\}$. Recall that we can always assume that $0 \in T$ and $\pi(0) = 0$.

We now prove that Theorem 1 implies that there must exist a subset $T' \subset \ell^2$ such that $\gamma_B(T')$ is comparable to $S_B(T)$. The idea of the proof works also for other classes of canonical processes for which we can characterize $S_X(T)$ in terms of increments (see Remark 3 below).

Theorem 11. *There exists a function $\pi : T \to \ell^2$ such that*

$$K^{-1}(\gamma_B(T_1+T_2)) \leq S_B(T) \leq K(\gamma_B(T_1+T_2)), \tag{2.9}$$

where K is a universal constant, $T_1 = \{t-\pi(t) : t \in T\}$ and $T_2 = \{\pi(t) : t \in T\}$.

Proof. By the main result of [2] we get the existence of $\pi : T \to \ell^2$ and consequently the existence of a decomposition into countable sets $T_1, T_2 \subset \ell^2$ such that $T \subset T_1 + T_2$ and

$$S_B(T) \geq K^{-1}(\sup_{t\in T_1} \|t\|_1 + S_G(T_2)), \tag{2.10}$$

where K is a universal constant. By the famous Fernique-Talagrand majorising measure bound ([28], [22], [8]) we know that $S_G(T)$ is comparable with $\gamma_G(T_2)$. To be precise, we know that $S_G(T) \geq C\gamma_2(T)$, where $\gamma_2(T)$ is Talagrand's γ_2 functional given by

$$\inf \sup_{t\in T_2} \sum_{n=0}^{\infty} 2^{n/2}\Delta_2(A_n(t)),$$

where the infimum runs over all admissible sequences of partitions of T_2 and Δ_2 denotes the diameter of a set in ℓ^2 norm. Obviously, we associate the sequence (π_n) with any admissible partition $(\mathcal{A}_n)_{n\geq 0}$ by choosing $\pi_n(t)$ to be any point in the set $A_n(t)$. Notice that the diameters of sets $A_n(t)$ converge uniformly to 0 so (2.5) is satisfied. To conclude that $\gamma_2(T_2) \geq C\gamma_G(T_2)$ we just estimate the 2^n-th norm of a Gaussian random variable by the 2nd norm. Now, let g be a standard normal variable independent of B_t, $t \in T$. Observe that for any $p \geq 1$,

$$\frac{\sqrt{2}}{\sqrt{\pi}}\|B_t - B_s\|_p = \mathbf{E}|g|\|B_t - B_s\|_p \leq \|G_t - G_s\|_p$$

so we can conclude $\frac{\sqrt{\pi}}{\sqrt{2}}\gamma_G(T_2) \geq \gamma_B(T_2)$.

The next goal is to show that $\sup_{t\in T_1} \|t\|_1 \geq C\gamma_B(T_1)$. To this end we consider a dense countable subset S_1 of T_1. We will start with constructing an admissible sequence of partititions and the associated sequence (π_n) for S_1 and then we will extend the construction to the whole T_1. We choose an admissible sequence $(\mathcal{A}_n(S_1))_{n\geq 0}$ of partitions such that for $n > 0$, $\mathcal{A}_n(S_1)$ consists of $N_n - 1$ single points and one addititonal set that contains all the remaining points. Clearly this partition is nested. Fix t_0 in S_1. Define

$$\pi_n(t) = \begin{cases} t \text{ if } & A_n(t) = \{t\} \\ t_0 & \text{otherwise.} \end{cases}$$

Obviously, the sequence of partitions together with the sequence (π_n) defined above satisfy (2.2). If for some m, $t \in A_m(t) = \{t\}$ but $A_{m-1}(t) \neq \{t\}$, then

$$\sum_{n=1}^{\infty} \|B_{\pi_n(t)} - B_{\pi_{n-1}(t)}\|_{2^n} = \|B_t - B_{t_0}\|_{2^m} \leq \|t - t_0\|_1.$$

Therefore, by the triangle inequality,

$$\sup_{t\in S_1} \sum_{n=1}^{\infty} \|B_{\pi_n(t)} - B_{\pi_{n-1}(t)}\|_{2^n} \leq 2 \sup_{t\in S_1} \|t\|_1.$$

In this way we have proved that $\sup_{t\in S_1} \|t\|_1 \geq \frac{1}{2}\gamma_B(S_1)$. Now, we provide the procedure that allows to extend the construction of the partition and the sequence (π_n) to the whole T_1. Recall that the partition element $\mathcal{A}_n(S_1)$ consisted of $N_n - 1$ singletons which we will denote by S^n. Of course, $S^1 \subset S^2 \subset \dots$. Furthermore, S_1 is countable, so we can refer to some fixed order on S_1. To construct the partition $\mathcal{A}_n(T_1)$ we will proceed by induction. For $n = 0$ we simply put $\mathcal{A}_0(T_1) = \{T_1\}$, $\pi_0(t) = t_0 \in S_1 \cap T_1$, since $S_1 \cap T_1 \neq \emptyset$. Then, suppose we have constructed $\mathcal{A}_{n-1}(T_1)$, $\pi_{n-1}(t)$, $t \in T_1$, $n \geq 2$. Consider $A \in \mathcal{A}_{n-1}(T_1)$. For $t \in A$ we define $\pi_n(t)$ as the element s of S^{n-1} which minimizes $\|t - s\|_2$. In the case of multiple minimizers we choose the smallest in the assumed order. This mapping defines the partition of the set A into not more than $N_{n-1} - 1$ elements. We repeat this procedure for the remaining elements of $\mathcal{A}_{n-1}(T_1)$ to obtain $\mathcal{A}_n(T_1)$, which satisfies the condition of being nested in an obvious way. Moreover, its cardinality does not exceed $N_{n-1} \cdot N_{n-1} = N_n$. Finally, the construction of $\pi_n(t)$ gurantees the regularity condition (2.5). This finishes the construction.

By Lemma 1 we obtain

$$S_B(T) \geq K^{-1}(\gamma_B(T_1) + \gamma_B(T_2)) \geq (KC_1)^{-1}\gamma_B(T_1 + T_2).$$

On the other hand, we have a trivial upper bound

$$S_B(T) \leq S_B(T_1) + S_B(T_2) = S_B(T_1 + T_2) \leq 3\gamma_B(T_1 + T_2),$$

by Theorem 10. □

Remark 2. The natural lower bound in the above Theorem would be of course $\gamma_B(T)$ rather than $\gamma_B(T_1 + T_2)$. However, it is not necessarily true that γ_B is monotone in the sense that $\gamma_B(T) \leq \gamma_B(T_1 + T_2)$ despite the fact that $T \subset T_1 + T_2$. The problem is that $\pi_n(T) \subset T_1 + T_2$ but we cannot easily rearrange π_n so that $\pi_n(T) \subset T$. Obviously, we could reformulate our definition in a way that $\pi_n : T \to \ell^2$. In this setting γ_B is monotonoe but values of π still may stay outside of T. It is a non-trivial question whether is is possible to improve the choice of the decomposition map π on the set T so that $\pi(T) \subset T$.

Let us also observe that for $\mathbf{P}(|\xi_i| > t) = \exp(-c_p t^p)$, $p \geq 2$ we could give a similar proof, since for any p we have Talagrand's [32] characterization of $S_X(T)$.

Remark 3. For the class of canonical processes based on independent symmetric ξ_i such that $\mathbf{P}(|\xi_i| > t) = \exp(-c_p t^p)$, $p \geq 2$, $S_X(T)$ is comparable with $\gamma_X(T_1 + T_2)$ up to a constant for some $T_1 + T_2 \subset \ell^2$ that contains T. The role of T_2 can be again associated with comparing the process X with the Gaussian process, whereas $T_1 \subset \ell^{p^*}$ for $p^* = \frac{p}{p-1}$.

In general, we conjecture that the same is true for canonical processes based on log-concave random variables.

Conjecture 1. *If $(\xi_i)_{i\geq 1}$ is a sequence of independent log-concave random variables with mean 0 and variance 1 then there exist $\pi : T \to \ell^2$ and sets $T_1 = \{t - \pi(t) \in \ell^2 : t \in T\}$ and $T_2 = \{\pi(t) \in \ell^2 : t \in T\}$ such that*

$$K^{-1}(\gamma_X(T_1 + T_2)) \leq S_X(T) \leq K(\gamma_X(T_1 + T_2)),$$

where K is a universal constant.

2.3 Contractions of canonical Bernoulli processes

Suppose we have a map $\varphi : T \to \ell^2$. The main question we treat in this chapter is under what assumptions on X_t, T and φ we can show that $S_X(\varphi(T))$ is bounded by $S_X(T)$ up to a numerical constant. In particular we are interested in the case of canonical Bernoulli processes.
Let us start with classical results concerning comparison of Gaussian processes. It is well-known that if G_t and G'_t, $t \in T$, are centered Gaussian processes and $\mathbf{E}|G_t - G_s|^2 \leq \mathbf{E}|G'_t - G'_s|^2$, then for each finite subset $F \subset T$,

$$\mathbf{E} \sup_{t\in F} G_t \leq \mathbf{E} \sup_{t\in F} G'_t. \tag{2.11}$$

This is a consequence of Slepian's Lemma ([18, Corollary 3.14] provides the proof with constant 2; the proof with the best possible constant 1 is in [8, Corollary 2.1.3]). Note also that by the Majorizing Measure Theorem the result can be generalized to the case where we compare a centered Gaussian process with a centered process for which we only require subgaussianity (see [18, Theorem 12.16]). We start with a discussion of possible extensions of this result. It is natural to ask for other cases when similar comparison results hold. From Theorem 10 it can be easily deduced that if we can compare moments then we can compare γ-type upper bounds.

Corollary 2. *Suppose that $(X_t)_{t\in T}$ is a canonical process and $\varphi : T \to \ell^2$. If there exists a universal constant C such that for each $n \geq 1$*

$$\|X_{\varphi(t)} - X_{\varphi(s)}\|_{2^n} \leq C\|X_t - X_s\|_{2^n}. \tag{2.12}$$

then $S_X(\varphi(T)) \leq 3C\gamma_X(T)$.

Proof. Clearly, by Theorem 10 we have $S_X(\varphi(T)) \leq 3\gamma_X(\varphi(T)) \leq 3C\gamma_X(T)$. □

This means that if we could show that $S_X(t) \geq K^{-1}\gamma_X(T)$, then by Corollary 2 we would get $S_X(\varphi(T)) \leq 3CKS_X(T)$. Unfortunately, in general, there is no proof that $\gamma_X(T)$ is comparable with $S_X(T)$. On the other hand, as discussed before, there are cases where the idea works. In particular, we could use Corollary 2 in order to recover the Gaussian comparison result with some absolute constant. However, in the Gaussian setting, one can simply refer to (2.11), rewriting it in the following way:

$$\text{if } \varphi : T \to \ell^2 \text{ satisfies } \|\varphi(t) - \varphi(s)\|_2 \leq \|t - s\|_2, \text{then } S_G(\varphi(T)) \leq S_G(T). \tag{2.13}$$

We now move to the case of canonical Bernoulli processes. The only known comparison result is Theorem 7([29, Theorem 2.1] or [18, Theorem 4.12]). It states that if $\varphi = (\varphi_i)_{i\geq 1} : T \to \ell^2$, where $\varphi_i : \mathbb{R} \to \mathbb{R}$ are contractions, then $S_B(T)$ dominates $S_B(\varphi(T))$ with constant 1:

$$\text{if } |\varphi_i(x) - \varphi_i(y)| \leq |x - y| \text{ for } i \geq 1, \text{ then } S_B(\varphi(T)) \leq S_B(T). \tag{2.14}$$

Note that if we are interested in comparison up to a numerical constant (not necessarily 1) then the requirement of coordinate contractions is too demanding. However, it is known that the result analogous to (2.11), where we assume that $\varphi : \ell^2 \to \ell^2$ is a Lipschitz contraction, does not hold for Bernoulli processes. Indeed, we can consider

$$T = \{e_1, e_2, \dots\},$$

where e_i are elements of the basis in ℓ^1 and

$$\varphi(T) = \{\frac{1}{\sqrt{n}}(\underbrace{\pm 1, \pm 1, \dots, \pm 1}_{n \text{ terms}}, 0, 0, \dots)\}.$$

Then,

$$\|B_{\varphi(t)} - B_{\varphi(s)}\|_2 = \sqrt{2} = \|B_t - B_s\|_2,$$

$S_B(T) = 1$, but $S_B(\varphi(T)) \geq \sqrt{n}$.

Therefore some additional assumptions on φ or T are required. As we show in this chapter, comparison for canonical Bernoulli depends on a suitable family of distances already present in (2.8). The following comparison result is a straightforward consequence of Theorem 11.

Corollary 3. *Suppose that $\varphi : T \to \ell^2$ can be extended to $T_1 + T_2$ in such a way that for any $p \geq 1$,*

$$\|B_{\varphi(t)} - B_{\varphi(s)}\|_p \leq \|B_t - B_s\|_p \text{ for all } s, t \in T_1 + T_2.$$

Then $S_B(\varphi(T)) \leq KS_B(T)$, where K is a universal constant.

Proof. Clearly, by Theorem 10 we have $S_B(\varphi(T)) \leq 3\gamma_B(\varphi(T))$. Hence, by Theorem 11,

$$S_B(\varphi(T)) \leq 3\gamma_B(\varphi(T)) \leq 3\gamma_B(\varphi(T_1 + T_2)) \leq 3\gamma_B(T_1 + T_2) \leq 3KS_B(T). \qquad \square$$

Note that the problem with application of the above result is that $T_1 + T_2$ may be much larger than T. We conjecture the following generalization of the above result.

Conjecture 2. *Let $\varphi = (\varphi_i)_{i\geq 1} : T \to \ell^2$. If*

$$\|B_{\varphi(t)} - B_{\varphi(s)}\|_p \leq \|B_t - B_s\|_p, \ \ \textit{for all} \ \ p \geq 2, \ \ s,t \in T, \tag{2.15}$$

then $S_B(\varphi(T)) \leq K S_B(T)$ for an absolute constant K.

We prove a weaker form of the conjecture. As explained before, the norm $\|B_t - B_s\|_p$ can be decomposed into the Gaussian and ℓ^1 parts. Our condition states that if the Gaussian part of $\|B_t - B_s\|_p$ dominates the Gaussian part of $\|B_{\varphi(t)} - B_{\varphi(s)}\|_p$, for all $s,t \in T$ and $p \geq 1$ then $S_B(T)$ dominates $S_B(\varphi(T))$ up to an absolute constant.

Theorem 12. *Suppose that for all $s,t \in T$ and all natural $p \geq 0$ we have*

$$\inf_{|I^c|\leq Cp} \sum_{i\in I} |\varphi_i(t) - \varphi_i(s)|^2 \leq C^2 \inf_{|I^c|\leq p} \sum_{i\in I} |t_i - s_i|^2 \tag{2.16}$$

for an absolute constant $C \geq 1$. Then $S_B(\varphi(T)) \leq K S_B(T)$, where K is a universal constant.

Remark 4. The result is stronger than the comparison for Bernoulli processes (2.14). It is easy to see the example of φ for which the contraction on each coordinate will fail, but if for $t \in T$, $\varphi(t)$ is zero for all, but some fixed number of coordinates then C can be chosen to be appropriately large so that (2.16) holds for $p = 0$. Consequently, the comparison will hold true. In this way Theorem 12 supports the conjecture that (2.15) suffices to prove that $S_B(\varphi(T)) \leq K S_B(T)$.

There is an important case for which the conjecture is true: when we assume that the supports $J(t) = \{i \geq 1 : |t_i| > 0\}$ of $t \in T$ are pairwise disjoint for all $t \in T$. It is crucial to understand that in this case the decomposition postulated in the Bernoulli Theorem can have a special form: $\pi(t) = t_{J^1(t)}$ and $t - \pi(t) = t_{J^2(t)}$, where $J^1(t)$ and $J^2(t)$ are disjoint and $J^1(t) \cup J^2(t) = J(t)$. We use the following notation. For $J^k \subset \{1,2,\ldots\}$,$k = 1,2$, we define $t1_{J^i} \in \ell^2$ such that $(t1_{J^k})_i = t_i$ for $i \in J^k$ and $(t1_{J^k})_i = 0$ otherwise. We show this fact when proving the following result.

Theorem 13. *Suppose that (2.15) is satisfied and the supports $J(t) = \{i \geq 1 : |t_i| > 0\}$ are pairwise disjoint for all $t \in T$. Then $S_B(\varphi(T)) \leq K S_B(T)$, where K is a universal constant.*

As we show in the last section, results of this type are of interest when one wants to compare weak and strong moments for random series in a Banach space. The question was proposed by K. Oleszkiewicz in a private communication.

2.4 Proof of the main results

In this section we prove Theorems 12 and 13.

Proof of Theorem 12. The main step in the proof of the Bernoulli Theorem ([2, Proposition 6.2]) is to show the existence of a suitable admissible sequence of partitions. Consequently, if $S_B(T) < \infty$ and $0 \in T$ then it is possible to define nested partitions $\mathcal{A}_n$ of T such that $|\mathcal{A}_n| \le N_n$. Moreover, for each $A \in \mathcal{A}_n$ one can find $j_n(A) \in \mathbb{Z}$ and $\pi_n(A) \in T$ (we use the notation $j_n(t) = j_n(A_n(t))$ and $\pi_n(t) = \pi_n(A_n(t))$, where $t \in A_n(t) \in \mathcal{A}_n$) which satisfy the following conditions for $M > 0$ and $r \ge 2$:

(i) $\|t - s\|_2 \le \sqrt{M} r^{-j_0(T)}$ for $s, t \in T$;

(ii) if $n \ge 1$, $\mathcal{A}_n \ni A \subset A' \in \mathcal{A}_{n-1}$ then either

  (a) $j_n(A) = j_{n-1}(A')$ and $\pi_n(A) = \pi_{n-1}(A')$, or

  (b) $j_n(A) > j_{n-1}(A')$, $\pi_n(A) \in A'$ and

$$\sum_{i \in I_n(A)} \min\{|t_i - \pi_n(A)_i|^2, r^{-2j_n(A)}\} \le M 2^n r^{-2j_n(A)}, \tag{2.17}$$

  where for any $t \in A$,
  $I_n(A) = I_n(t) = \{i \ge 1 : |\pi_{k+1}(t)_i - \pi_k(t)_i| \le r^{-j_k(t)} \text{ for } 0 \le k \le n-1\}$.

(iii) Moreover, the numbers $j_n(A)$, $A \in \mathcal{A}_n$, $n \ge 0$, satisfy

$$\sup_{t \in T} \sum_{n=0}^{\infty} 2^n r^{-j_n(t)} \le L S_B(T), \tag{2.18}$$

  where L is an absolute constant.

As proved in [2, Theorem 3.1], the existence of the quantities $\mathcal{A}_n, j_n(A), \pi_n(A), I_n(A)$ that satisfy conditions (i) and (ii) formulated above implies the existence of a decomposition $T_1, T_2 \subset \ell^2$, $T_1 + T_2 \supset T$ such that

$$\sup_{t^1 \in T_1} \|t^1\|_1 \le LM \sup_{t \in T} \sum_{n=0}^{\infty} 2^n r^{-j_n(t)} \quad \text{and} \quad \gamma_G(T_2) \le L\sqrt{M} \sup_{t \in T} \sum_{n=0}^{\infty} 2^n r^{-j_n(t)}.$$

Together with condition (iii) we get (2.10). Our aim is to use the mapping φ to transport all the required quantities to $\varphi(T)$. Before we do it, we formulate an auxiliary fact about the sets $I_n(A)$: we show that we can get rid of truncation in (2.17) if we skip a well-controlled number of coordinates. We observe that for each $t \in A \in \mathcal{A}_n$ there must exist a set $J_n(t)$ such that $|J_n(t)^c| \le M 2^{n+1}$ and

$$\sum_{i \in J_n(t)} |t_i - \pi_n(t)_i|^2 \le M 2^n r^{-2j_n(t)}. \tag{2.19}$$

The fact will be proved in two steps. First, we show that $|I_n(t)^c| \le M 2^n$. We can only prove that $|I_n(t)| = |I_n(A_n(t))| \le 2^n$ if $\pi_{n-1}(t) \ne \pi_n(t)$, which implies $j_{n-1}(t) \ne j_n(t)$ and $\pi_n(t) \in A_{n-1}(t)$. Therefore, there exists $k \in \{1, \ldots, n\}$ such that

$$j_{n-1}(t) = j_{n-k}(t) > j_{n-k-1}(t), \quad \text{where } j_{-1}(t) = -\infty$$

and hence $\pi_n(t) \in A_{n-1}(t) \subset A_{n-k}(t)$ and $\pi_{n-1}(t) = \pi_{n-k}(t)$, $j_{n-1}(t) = j_{n-k}(t)$, so by the construction of $(\mathcal{A}_n)_{n\geq 0}$,

$$\sum_{i\in I_{n-k}(t)} \min\{(\pi_n(t)_i - \pi_{n-1}(t)_i)^2, r^{-2j_{n-1}(t)}\}$$
$$= \sum_{i\in I_{n-k}(t)} \min\{(\pi_n(t)_i - \pi_{n-k}(t)_i)^2, r^{-2j_{n-k}(t)}\} \leq M2^{n-k}r^{-2j_{n-k}(t)}.$$

Consequently,

$$|\{i \in I_{n-k}(t) : |\pi_n(t)_i - \pi_{n-1}(t)_i| > r^{-j_{n-1}(t)}\}| \leq M2^{n-k}.$$

Obviously,

$$|I_n(t)^c| \leq |I_{n-k}(t)^c| + |\{i \in I_{n-k}(t) : |\pi_n(t)_i - \pi_{n-1}(t)_i| > r^{-j_{n-1}(t)}\}|$$
$$\leq |I_{n-k}(t)^c| + M2^{n-k}.$$

Therefore by induction, $|I_n(t)^c| \leq M\sum_{k=1}^{n} 2^{n-k} \leq M2^n$. Let

$$J_n(t) = \{i \in I_n(A) : |t_i - \pi_n(t)_i| \leq r^{-j_n(A)}\}.$$

The second step is to establish that $|I_n(t) \setminus J_n(t)| \leq M2^n$. Again it suffices to prove the result only for n such that $j_n(t) > j_{n-1}(t)$. Note that by (2.17),

$$|I_n(t) \setminus J_n(t)|r^{-2j_n(t)} = \sum_{i\in I_n(A)\setminus J_n(t)} r^{-2j_n(t)} \leq M2^n r^{-2j_n(t)},$$

and hence the result holds. It remains to observe that

$$|J_n(t)^c| \leq |I_n(t)^c| + |I_n(t) \setminus J_n(t)| \leq M(2^n + 2^n) \leq M2^{n+1}.$$

We turn to constructing an admissible sequence of partitions together with all the related quantities for the set $\varphi(T)$. Let $\mathcal{B}_n$ consist of $\varphi(A)$, $A \in \mathcal{A}_n$. Obviously the partitions $\mathcal{B}_n$ are admissible, nested and $\mathcal{B}_0 = \{\varphi(T)\}$. Moreover, for each $n \geq 0$ and $A \in \mathcal{A}_n$ we define

$$\pi_n(\varphi(A)) = \varphi(\pi_n(A)) \text{ and } j_n(\varphi(A)) = j_n(A)$$

and obviously

$$I_n(\varphi(A)) = I_n(\varphi(t))$$
$$= \{i \geq 1 : |\varphi(\pi_{k+1}(t))_i - \varphi(\pi_k(t))_i| \leq r^{-j_k(\varphi(t))} \text{ for } 0 \leq k \leq n-1\}.$$

As mentioned at the beginning of this proof, in order to use [2, Theorem 3.1] we have to verify conditions (i) and (ii) for the new sequence $\mathcal{B} = (\mathcal{B}_n)_{n\geq 0}$ as well as $j_n(B), \pi_n(B), I_n(B)$ for $B \in \mathcal{B}_n$, $n \geq 0$. For this we need our main condition (2.16). First it is obvious that (2.16) implies for $p = 0$ that

$$\|\varphi(t) - \varphi(s)\|_2 \leq \|t - s\|_2 \leq \sqrt{M}r^{-j_0(T)}.$$

If $A \in \mathcal{B}_n$ and $\varphi(A) \subset \varphi(A') \in \mathcal{B}_{n-1}$ then either

$$j_n(\varphi(A)) = j_n(A) = j_{n-1}(A') = j_{n-1}(\varphi(A'))$$

and

$$\pi_n(\varphi(A)) = \varphi(\pi_n(A)) = \varphi(\pi_{n-1}(A')) = \pi_{n-1}(\varphi(A')),$$

or $j_n(\varphi(A)) = j_n(A) > j_{n-1}(A') = j_{n-1}(\varphi(A'))$. In this case we have $\pi_n(\varphi(A)) = \varphi(\pi_n(A)) \in \varphi(A')$ and it suffices to show that

$$\sum_{i \in I_n(\varphi(A))} \min\{|\varphi(t)_i - \varphi(\pi_n(A))_i|^2, r^{-2j_n(\varphi(A))}\} \le C2^n r^{-2j_n(\varphi(A))}. \tag{2.20}$$

Obviously, the problem now is that we know little about the structure of the set $I_n(\varphi(A))$. Therefore, we simply prove that

$$\sum_{i \ge 1} \min\{|\varphi(t)_i - \varphi(\pi_n(A))_i|^2, r^{-2j_n(\varphi(A))}\} \le C2^n r^{-2j_n(\varphi(A))}.$$

Clearly,

$$\begin{aligned} &\sum_{i \ge 1} \min\{|\varphi(t)_i - \varphi(\pi_n(A))_i|^2, r^{-2j_n(\varphi(A))}\} \\ &\qquad\qquad \le C_2 2^n r^{-2j_n(A)} + \inf_{|I^c| \le C_2 2^n} \sum_{i \in I} |\varphi(t)_i - \varphi(\pi_n(A))_i|^2. \end{aligned} \tag{2.21}$$

We can choose $C_2 \ge 2CM$ in such a way that by (2.16) we get

$$\begin{aligned} &\inf_{|I^c| \le C_2 2^n} \sum_{i \in I} |\varphi(t)_i - \varphi(\pi_n(A))_i|^2 \\ &\le C^2 \inf_{|I^c| \le M 2^{n+1}} \sum_{i \in I} |t_i - \pi_n(A)_i|^2 \le C^2 \sum_{i \in J_n(t)} |t_i - \pi_n(A)_i|^2. \end{aligned}$$

Hence, by (2.19) and (2.21),

$$\sum_{i \ge 1} \min\{|\varphi(t)_i - \varphi(\pi_n(A))_i|^2, r^{-2j_n(\varphi(A))}\} \le (C_2 + C^2 M) 2^n r^{-2j_n(A)}$$

which proves (2.20) with $C_3 = C_2 + C^2 M$. We have proved that the assumptions required in [2, Theorem 3.1] are satisfied for $(\mathcal{B}_n)_{n \ge 0}$ and the related quantities. Consequently, there exists a decomposition $S_1, S_2 \subset \ell^2$ such that $S_1 + S_2 \supset \varphi(T)$ and

$$\sup_{s \in S_1} \|s\|_1 \le LC \sup_{t \in \varphi(T)} \sum_{n \ge 0} 2^n r^{-j_n(t)}, \quad \gamma_G(S_2) \le L\sqrt{C} \sup_{t \in \varphi(T)} \sum_{n \ge 0} 2^n r^{-j_n(t)}.$$

Since $j_n(\varphi(t)) = j_n(t)$ and we have (2.18) for $(\mathcal{A}_n)_{n\geq 0}$, we obtain

$$\sup_{t\in\varphi(T)} \sum_{n\geq 0} 2^n r^{-j_n(t)} \leq L S_B(T).$$

This implies that

$$S_B(\varphi(T)) \leq S_B(S_1) + S_B(S_2) \leq K S_B(T),$$

for a universal constant K and ends the proof. □

The second case we consider is when the supports $J(t) = \{i \geq 1 : |t_i| > 0\}$ are pairwise disjoint for all $t \in T$. Recall the following notation. For any $t \in \ell^2$ and $J \subset \{1, 2, \ldots\}$ we define $t1_J \in \ell^2$ such that $(t1_J)_i = t_i$ for $i \in J$ and $(t1_J)_i = 0$ otherwise.

Proof of Theorem 13. Obviously, we may require that $S_B(T) < \infty$. We additionally assume that $0 \in T$. This simplifies the proof, but the rpoof works also for the general case as we will point out at the end. Recall that by the Bernoulli Theorem [2] there exists a decomposition $T_1 + T_2 \supset T$ such that

$$S_B(T) \geq K^{-1}(\sup_{t\in T_1} \|t\|_1 + \gamma_G(T_2)), \tag{2.22}$$

where K is an absolute constant. Obviously, we may think of K as suitably large. We can represent the decomposition by $\pi : T \to \ell^2$ such that $T_2 = \{\pi(t) : t \in T\}$ and $T_1 = \{t - \pi(t) : t \in T\}$. We show that under the disjoint supports assumption we may additionally require that $\pi(t) = t1_{J^2(t)}$ and $t - \pi(t) = t1_{J^1(t)}$ where $J^1(t)$ and $J^2(t)$ are disjoint subsets of $J(t)$ such that $J^1(t) \cup J^2(t) = J(t)$. Moreover, $J^2(t) = \{i \in J(t) : |t_i| \leq p(t)\}$ for some suitably chosen $p(t) \geq 0$.
In order to prove the result we have to look closer into the definition of $\pi(t)$ in [2, proof of Theorem 3.1]. The definition is based on the construction of admissible sequences of partitions we have described in the proof of Theorem 12 above. Using the notation introduced there, let

$$m(t,i) = \inf\{n \geq 0 : |\pi_{n+1}(t)_i - \pi_n(t)_i| > r^{-j_n(t)}\}, \quad t \in T, i \geq 1. \tag{2.23}$$

Note that $S_B(T)$ is comparable with $\sup_{t\in T} \sum_{n\geq 0} 2^n r^{-j_n(t)}$. Therefore, if $S_B(T)$ is finite then necessarily $\lim_{n\to\infty} j_n(t) = \infty$ for all $t \in T$. From the partition construction used in [2, Section 6] we know that we can additionally assume a regularity condition on $j_n(t)$, $n \geq 0$, namely

$$j_n(t) \leq j_{n-1}(t) + 2 \text{ for all } n \geq 0,$$

and for technical purposes we take $j_{-1}(t) = -\infty$. As in [2, proof of Theorem 3.1] the Bernoulli decomposition $\pi(t)$ is given by $\pi(t)_i = \pi_{m(t,i)}(t)_i$, where if $m(t,i) = \infty$ the definition means that $\pi(t)_i = \lim_{n\to\infty} \pi_n(t)_i$ and the limit exists. Consequently, denoting $J_n(t) = \{i \geq 1 : m(t,i) = n\}$ and $J_\infty(t) = \{i \geq$

$1 : m(t,i) = \infty\}$ we get

$$\pi(t) = \sum_{n\geq 0} \pi_n(t) 1_{J_n(t)} + \pi(t) 1_{J_\infty(t)}.$$

Clearly, $J_n(t)$, $n \geq 0$ and $J_\infty(t)$ are disjoint. Note also that if $m(t,i) = \infty$ and $i \in J(\pi(t))$, then there must exist $n \geq 0$ such that $|\pi_k(t)_i| > 0$ for all $k \geq n$. Due to the disjoint supports assumption this is only possible if there exists $n \geq 0$ such that $\pi_n(t)_i = \pi_{n+1}(t)_i = \cdots$. Now, if there exists $m \geq 0$ such that $A_m(t) = \{t\}$ we define

$$\tau(t) = \inf\{n \geq 0 : A_n(t) = \{t\} = \{\pi_n(t)\},\ j_{n-1}(t) < j_n(t)\},$$

otherwise $\tau(t) = \infty$. The time $\tau(t)$ is of special nature in the sense that without loss of generality we may assume that $j_n(t) = j_{n-1}(t) + 2$ for $n \geq \tau(t)$. This is due to the fact that partitioning ceases after that time. Now, we define

$$J^2(t) = \{i \in J(t) : |t_i| \leq r^{-j_{\tau(t)}(t)-1}\}, \quad J^1(t) = J(t) \setminus J^2(t).$$

We can now introduce the improved version of π denoted by $\bar{\pi}$ and given by

$$\bar{\pi}(t) = t 1_{J^2(t)}.$$

It is clear that

$$\|t - \bar{\pi}(t)\|_1 = \|t 1_{J^1(t)}\|_1.$$

For $n \geq 0$ let

$$L_n(t) = \{i \in J(t) : r^{-j_n(t)} < |t_i| \leq r^{-j_{n-1}(t)}\}.$$

Observe that $J^1(t) = \bigcup_{n<\tau(t)} L_n(t)$. If $i \in L_n(t)$, $n \geq 0$, then we may find $0 \leq m \leq n$ such that $j_{m-1}(t) < j_m(t) = j_{m+1}(t) = \ldots = j_n(t)$. Consequently, by the definition (2.17) of $I_n(t)$, for all $s \in A_m(t)$,

$$\sum_{i\in I_n(t)} \min\{|s_i - \pi_n(t)_i|^2, r^{-2j_n(t)}\} = \sum_{i\in I_m(t)} \min\{|s_i - \pi_m(t)_i|^2, r^{-2j_m(t)}\}$$
$$\leq M 2^m r^{-2j_m(t)} = M 2^m r^{-2j_n(t)} \leq M 2^n r^{-2j_n(t)}.$$

We need to show that the decomposition $\bar{\pi}$ is of the right form, i.e. it satisfies (2.22). To this end we need to investigate a few cases according to different possible paths of approximations π. First suppose that $t \neq \pi_n(t)$. Then we may use the above inequality for $s = t$ and thanks to disjoint supports we have

$$|I_n(t) \cap L_n(t)| r^{-2j_n(t)} \leq \sum_{i\in I_n(t)} \min\{|s_i - \pi_n(t)_i|^2, r^{-2j_n(t)}\} \leq M 2^n r^{-2j_n(t)}$$
$$\text{so } \ |I_n(t) \cap L_n(t)| \leq M 2^n.$$

The same inequality holds if $t = \pi_n(t)$ but $A_m(t) \neq \{t\}$. We show that $L_n(t) \subset I_n(t)$. Indeed, suppose that $i \notin I_n(t)$. This means that for some

$k \in \{0, 1, \dots, n-1\}$ we have $|\pi_{k+1}(t)_i - \pi_k(t)_i| > r^{-j_k(t)}$. This may concern $i \in J(t)$ only if $\pi_{k+1}(t) = t$, $\pi_k(t) \neq t$ or $\pi_k(t) = t$ and $\pi_{k+1}(t) \neq t$, but then it means that $|t_i| > r^{-j_k(t)} \geq r^{-j_{n-1}(t)}$, i.e. $i \notin L_n(t)$. This proves $L_n(t) \subset I_n(t)$. For $1 \leq n < \tau(t)$ this implies that

$$\sum_{i \in L_n(t)} |t_i| \leq M 2^n r^{-j_{n-1}(t)}. \tag{2.24}$$

For $n = 0$ we use simply $|t_i| \leq 2S_B(T)$ and hence

$$\sum_{i \in L_0(t)} |t_i| \leq 2MS_B(T). \tag{2.25}$$

Now suppose that $t = \pi_n(t) = \pi_m(t)$ and $A_m(t) = \{t\}$. If either $t \neq \pi_{m-1}(t)$ or $\{t\} \neq A_{m-1}(t)$, then $\tau(t) = m$. Otherwise $\tau(t) < m$. If $\tau(t) = m$, then by the above argument,

$$\sum_{i \in L_n(t)} |t_i|^2 = \sum_{i \in L_n(t)} \min\{|t_i|^2, r^{-2j_{m-1}(t)}\} \leq M 2^{m-1} r^{-2j_{m-1}(t)},$$

and thus using the fact that $|t_i| \geq r^{-j_m(t)-1}$ and $j_m(t) = j_{m-1}(t) + 2$, we have

$$\sum_{i \in L_n(t)} |t_i| \leq M 2^{m-1} r^{-2j_{m-1}(t) + j_m(t)} \leq M 2^{m-1} r^{-j_{m-1}(t)+2}.$$

We have the remaining bound

$$\sum_{i \in L_n(t)} |t_i| \leq M 2^{\tau(t)-1} r^{-j_{\tau(t)-1}(t)+2}. \tag{2.26}$$

Combining (2.24)–(2.26) we conclude by (2.18)

$$\begin{aligned} \|t 1_{J^1(t)}\|_1 &\leq 2MS_B(T) + 2M \sum_{n=0}^{\tau(t)-2} r^{-j_n(t)} 2^n + M 2^{\tau(t)-1} r^{-j_{\tau(t)-1}(t)+2} \\ &\leq 2MLS_B(T), \end{aligned} \tag{2.27}$$

where L is an absolute constant.

Now consider $s, t \in T$, $s \neq t$. In order to prove that

$$\|\bar{\pi}(s) - \bar{\pi}(t)\|_2 = \|t 1_{J^2(t)} - s 1_{J^2(s)}\|_2 \leq \|\pi(t) - \pi(s)\|_2 \tag{2.28}$$

we have to argue that $J^2(t) \cap J(\pi(s)) = \emptyset$, $J^2(s) \cap J(\pi(t)) = \emptyset$. Note that $J^2(t) \subset J_\infty(t)$ and $J^2(t) \subset J_\infty(s)$. Moreover, $J_\infty(s)$ and $J_\infty(s)$ are disjoint. Obviously, it suffices to show that $J^2(t) \cap J(\pi(s)) = \emptyset$.
First, note that $J^2(t) \cap J_\infty(s) = \emptyset$. Indeed, if the set were non-empty then for a given $n \geq 0$ we would have $t = \pi_n(s) = \pi_{n+1}(s) = \dots$, but then $s \in A_n(t)$ for all $n \geq 0$ and therefore $\tau(t) = \infty$. This would imply $J^2(t) = \emptyset$, which is a contradiction. Suppose that $i \in J^2(t)$ and $i \in J_n(s)$. This is only possible if $\pi_n(s) = t$ and $\pi_{n+1}(s) \neq \pi_n(s) = t$ and $r^{-j_n(s)} < |\pi_n(s)_i|$. Let $m \geq 0$ be

such that $j_{m-1}(s) < j_m(s) = j_{m+1}(s) = \ldots = j_n(s)$. Then either $m = 0$, or $m \geq 1$ and $t = \pi_n(s) = \pi_m(s) \in A_{m-1}(s)$, which means that $A_{m-1}(s) = A_{m-1}(t)$ and $j_{m-1}(s) = j_{m-1}(t)$. Therefore, $\tau(t) \geq m$ and $j_{\tau(t)}(t) > j_{m-1}(t)$. If $i \in J^2(t) \cap J_n(s)$, then

$$r^{-j_{m-1}(t)-2} = r^{-j_{m-1}(s)-2} \leq r^{-j_m(s)} < |t_i| \leq r^{-j_{\tau(t)}(t)} = r^{-j_{\tau(t)-1}(t)-2} \leq r^{-j_{m-1}(t)-2},$$

which is a contradiction. If $m = 0$, then the argument is trivial.
Summing up, by (2.27) we have

$$\sup_{t\in T} \|t - \bar{\pi}(t)\|_1 \leq LS_B(T)$$

and by (2.28) and Gaussian comparison we have $\gamma_G(\pi(T)) \leq \gamma_G(\bar{\pi}(T))$, which means that our improved version of π satisfies

$$S_B(T) \geq K^{-1}(\sup_{t\in T} \|t - \bar{\pi}(t)\|_1 + \gamma_G(\bar{\pi}(T))),$$

where K is a universal constant. In this way we have proved that we may additionally require that $\pi(t) = t1_{J^2(t)}$ and $t - \pi(t) = t1_{J^1(t)}$ for some disjoint $J^1(t), J^2(t)$ such that $J^1(t) \cup J^2(t) = J(t)$. Recall that $J^2(t)$ in each case is of the form $\{i \in J(t) : |t_i| \leq r(t)\}$ for a given $r(t) \geq 0$.
We turn to the main part of the proof. Let $p(t)$ be the smallest positive integer such that

$$\sqrt{p(t)}\|t1_{J^2(t)}\|_2 \geq KS_B(T) \geq \|t1_{J^1(t)}\|_1. \tag{2.29}$$

Note that it is possible that $J^2(t) = \emptyset$, in which case we may think of $p(t)$ as equal to ∞. Since K is large enough and $S_B(T) \geq \frac{1}{2}\sup_{t\in T}\|t\|_2$, it is clear that $p(t)$ must be at least, say, 2. Consequently, by the choice of $p(t)$,

$$\sqrt{p(t)}\|t1_{J^2(t)}\|_2 \leq 2KS_B(T). \tag{2.30}$$

The last step is to define a suitable decomposition for $\varphi(T)$. For each $t \in T$ we define $\pi(\varphi(t)) = t_{J^2(\varphi(t))}$ and $\varphi(t) - \pi(\varphi(t)) = t_{J^1(\varphi(t))}$, where $J^2(\varphi(t))$ and $J^1(\varphi(t))$ are defined by the decomposition of the norm $\|B_{\varphi(t)}\|_{p(t)}$, i.e.

$$\sum_{i\in J^1(\varphi(t))} |\varphi(t)_i| = \sup_{|I^c|\leq p(t)} \sum_{i\in I^c} |\varphi(t)_i|$$

and

$$\sum_{i\in J^2(\varphi(t))} |\varphi(t)_i|^2 = \inf_{|I^c|\leq p(t)} \sum_{i\in I} |\varphi(t)_i|^2.$$

Consequently, by the decomposition (2.8) and the main assumption (2.15),

$$\sum_{i\in J^1(\varphi(t))} |\varphi(t)_i| + \sqrt{p(t)}(\sum_{i\in J^2(\varphi(t))} |\varphi(t)_i|^2)^{\frac{1}{2}}$$
$$\leq 4\|B_{\varphi(t)}\|_{p(t)} \leq 4\|B_t\|_{p(t)} \leq 4(\|t1_{J^1(t)}\|_1 + \sqrt{p(t)}\|t1_{J^2(t)}\|_2).$$

Therefore, by (2.29), (2.30),

$$\sum_{i\in J^1(\varphi(t))} |\varphi(t)_i| \leq K_1 S_B(T).$$

Moreover, by (2.29),

$$(\sum_{i\in J^2(\varphi(t))} |\varphi(t)_i|^2)^{\frac{1}{2}} \leq K_2\|t1_{J^2(t)}\|_2.$$

This implies that

$$\begin{aligned}\|\pi(\varphi(t)) - \pi(\varphi(s))\|_2 &\leq \|\pi(\varphi(t))\|_2 + \|\pi(\varphi(s))\|_2 \\ &\leq K_2(\|t1_{J^2(t)}\|_2 + \|s1_{J^2(s)}\|_2) \leq K_3\|\pi(t) - \pi(s)\|_2.\end{aligned}$$

Therefore, by Gaussian comparison, we get $\gamma_G(\pi(\varphi(T))) \leq K\gamma_G(\pi(T))$ and hence finally

$$S_B(\varphi(T)) \leq K(\sup_{t\in T} \|\pi(\varphi(t))\|_1 + \gamma_G(\pi(\varphi(T)))) \leq KLS_B(T).$$

This ends the proof in the case when $0 \in T$. For the general case the proof follows the same lines, where instead of t we consider $t - \pi_0(t)$. Notice that formally this may not obey the disjoint supports assumption, but it does not qualitatively affect the argument presented above. □

Note that the above proof works since in the case of disjoint supports we have almost perfect knowledge about the decomposition in the Bernoulli Theorem. On the other hand, it is not difficult to give an alternative proof based on the independence of the variables B_t, $t \in T$, but it is worth seeing what the decomposition in [2, Theorem 3.1] should be in order to make Bernoulli comparison possible.

2.5 The Oleszkiewicz problem

In this section we apply our result to compare expectations of norms of random series in a Banach space. First, we prove a general result which concerns $\varphi : T \to \ell^2$ where φ is linear, T is convex and $T = -T$. Then the assumption (2.12) becomes

$$\|B_{\varphi(u)}\|_p \leq C\|B_u\|_p \;\text{ for all } p \geq 1 \;\text{ and }\; u \in \mathrm{cl}\,(\mathrm{Lin}(T)), \tag{2.31}$$

where $\mathrm{Lin}(T)$ is the linear space spanned by the set T. This is because by the assumptions on T any point $u \in \mathrm{Lin}(T)$ can be represented as $c \cdot t$, where $c \in \mathbb{R}$ and $t \in T$. By the linearity of φ,

$$\|B_{\varphi(u)}\|_p = |c|\|B_{\varphi(t)}\|_p \leq C|c|\|B_t\|_p = C\|B_u\|_p.$$

On the other hand, we can easily extend the condition (2.31) to the closure of $\mathrm{Lin}(T)$.

We turn to proving that if $\mathrm{cl}(\mathrm{Lin}(T)) = \ell^2$ then (2.31) implies that $S_B(T)$ dominates $S_B(\varphi(T))$.

Theorem 14. *Suppose that $T = -T$, T is convex and $\mathrm{cl}(\mathrm{Lin}(T)) \supset T_1 + T_2$, where T_1, T_2 are as in Bernoulli Theorem. If φ is linear and satisfies (2.12) then $S_B(\varphi(T)) \leq K S_B(T)$, where K is a universal constant.*

Proof. By the Bernoulli Theorem [2] there exist T_1, T_2 such that $T \subset T_1 + T_2$ and

$$S_B(T) \geq L^{-1}(\sup_{t\in T_1} \|t\|_1 + \gamma_G(T_2)).$$

Since φ is linear, it can be easily extended to $\mathrm{cl}(\mathrm{Lin}(T)) = \ell^2$ and thus we can define $S_i = \varphi(T_i)$, $i \in \{1, 2\}$. Obviously $S_1 + S_2 \supset \varphi(T)$; moreover, (2.31) implies in particular that

$$\|\varphi(u)\|_1 = \|B_{\varphi(u)}\|_\infty \leq C\|B_u\|_\infty = C\|u\|_1.$$

and

$$\|\varphi(u) - \varphi(v)\|_2 = \|B_{\varphi(u-v)}\|_2 \leq C\|B_{u-v}\|_2 = C\|u - v\|_2.$$

Consequently,

$$\sup_{s\in S_1} \|s\|_1 = \sup_{t\in T_1} \|\varphi(t)\|_1 \leq C \sup_{t\in T_1} \|t\|_1$$

and

$$\gamma_G(S_2) = \gamma_G(\varphi(T_2)) \leq C\gamma_G(T_2).$$

Therefore

$$\begin{aligned} S_B(\varphi(T)) &\leq S_B(S_1) + S_B(S_2) \leq K(\sup_{s\in S_1} \|s\|_1 + \gamma_G(S_2)) \\ &\leq CK(\sup_{t\in T_1} \|t\|_1 + \gamma_G(T_2)) \leq CK^2 S_B(T).\square \end{aligned}$$

We aim to study the question, posed by Oleszkiewicz, of comparability of weak and strong moments for Bernoulli series in a Banach space. Let x_i, y_i, $i \geq 1$, be vectors in a Banach space $(B, \|\cdot\|)$. Suppose that for all $x^* \in B^*$ and $u \geq 0$,

$$\mathbf{P}(|\sum_{i\geq 1} x^*(x_i)\varepsilon_i| > u) \leq \bar{C}\mathbf{P}(|\sum_{i\geq 1} x^*(y_i)\varepsilon_i| > \bar{C}^{-1}u). \tag{2.32}$$

This property is called weak tail domination. Weak tail domination can be understood in terms of comparability of weak moments, i.e. for any integer $p \geq 1$ and $x^* \in B^*$,

$$\|\sum_{i\geq 1} x^*(x_i)\varepsilon_i\|_p \leq C\|\sum_{i\geq 1} x^*(y_i)\varepsilon_i\|_p \tag{2.33}$$

Oleszkiewicz asked whether or not this implies comparability of strong moments, that is, whether (2.32) or rather (2.33) implies that

$$\mathbf{E}\|\sum_{i\geq 1} x_i\varepsilon_i\| = \mathbf{E}\sup_{x^*\in B_1^*}\sum_{i\geq 1} x^*(x_i)\varepsilon_i$$
$$\leq K\mathbf{E}\sup_{x^*\in B_1^*}\sum_{i\geq 1} x^*(y_i)\varepsilon_i = K\mathbf{E}\|\sum_{i\geq 1} y_i\varepsilon_i\|, \tag{2.34}$$

where K is an absolute constant. Note that in the Oleszkiewicz problem one may assume that B is a separable space since we can easily restrict the argument to the closure of $\mathrm{Lin}(y_1, x_1, y_2, x_2, \ldots)$. Therefore

$$\mathbf{E}\|\sum_{i\geq 1} y_i\varepsilon_i\| = \sup_{F\subset B_1^*}\mathbf{E}\sup_{x^*\in F}|\sum_{i\geq 1} x^*(y_i)\varepsilon_i|,$$

where F runs through all finite sets contained in $B_1^* = \{x^* \in B^* : \|x^*\| \leq 1\}$. We may assume that $\mathbf{E}\|\sum_{i\geq 1} y_i\varepsilon_i\| < \infty$ since otherwise there is nothing to prove. Consequently, for each $x^* \in B^*$ series $\sum_{i\geq 1} x^*(y_i)\varepsilon_i$ is convergent, which is equivalent to $\sum_{i\geq 1}(x^*(y_i))^2 < \infty$. Let $Q : B^* \to \ell^2$ be defined by $Q(x^*) = (x^*(y_i))_{i\geq 1}$. It is clear that $Q : B^*/\ker Q \to \ell^2$ is a linear isomorphism on the closed linear subspace of ℓ^2. We apply Theorem 14 to get the following result.

Corollary 4. *Suppose that Q is onto ℓ^2. Then (2.32) implies (2.34).*

Unfortunately, if Q is not onto ℓ^2 then the above argument fails. Still it is believed that the comparison holds. A partial result can be deduced from Theorem 12:

Corollary 5. *Suppose that for each $x^* \in B^*$ and $p \geq 0$,*

$$\inf_{|I^c|\leq Cp}\sum_{i\in I}|x^*(x_i)|^2 \leq C^2\inf_{|I^c|\leq p}\sum_{i\in I}|x^*(y_i)|^2. \tag{2.35}$$

Then (2.34) holds, i.e.

$$\mathbf{E}\|\sum_{i\geq 1} x_i\varepsilon_i\| \leq K\mathbf{E}\|\sum_{i\geq 1} y_i\varepsilon_i\|.$$

Proof. It suffices to notice that (2.35) implies (2.16) and then apply Theorem 12. □

Chapter 3

Lévy-Ottaviani type inequality

3.1 Introduction

Let $T \subset \mathbb{R}^n$. Recall that $\varepsilon_1, \ldots, \varepsilon_n$ is a sequence of independent Bernoulli random variables i.e. for each $i \geq 1$, $\mathbb{P}(\varepsilon_i = \pm 1) = 1/2$. We will employ a different notation in this chapter than in the rest of the dissertation because it will be more convenient and consistent with the notation of the original problem [26]. For the element $t = (t_1, \ldots, t_n)$ of T we define a random variable $X_t = \sum_{i=1}^n t_i \varepsilon_i$. Obviously, $\mathbf{E} X_t = 0$ and $\mathbf{Var}(X_t) = \sum_{i=1}^n t_i^2 =: \|t\|^2$. Furthermore, let $X = \sup_{t \in T} X_t$.
The main assumption of this work will be the existance of the point $t^0 \in T$ satisfying $\sup_{t \in T} \mathbf{Var}(X_t) = \mathbf{Var}(X_{t^0})$. We will refer to t^0 as the point of maximal variance. The question we want to study concerns the control over X one can expect from knowing t^0. It will be a simple consequence of Theorem 16 and could be also deduced from McDiarmid's inequality (see [34, Problem 3.7]) that the strengthened concentration inequality can be obtained (with constant 2 instead of 8 in the exponent). The more intriguing question is on the tail domination, namely can we expect a Lévy-Ottaviani type of inequality. For this, we define $Y = \sum_{i=1}^n t_i^0 \varepsilon_i$. The main motivation for the study of this question is the following problem posed by W. Szatzschneider in [26]. Suppose that $a_i : [0,1] \to \mathbb{R}_+$, for $i = 1, 2, \ldots, n$ are non-decreasing, right-continuous functions. In the orginal setting it was also assumed that functions a_i satisfy following conditions:

1. for each $t \in [0,1]$, $a_1(t) \geq a_2(t) \geq \ldots \geq a_n(t)$
2. $\sum_{i=1}^n a_i(1) \geq 1 + 2a_1(1)$.

Variables X and Y we defined at the beginning are now of the form $X = \sup_{t \in [0,1]} \sum_{i=1}^n a_i(t) \varepsilon_i$ and $Y = \sum_{i=1}^n a_i(1) \varepsilon_i$. W. Szatzschneider conjectured that under the above conditions the following inequality holds

$$\mathbf{P}(X \geq 1) \leq 2\mathbf{P}(Y \geq 1).$$

Notice that conditions 1 and 2 require that $n \geq 3$. In [26] the conjecture was proved for cases $n = 3$ and $n = 4$ by a simple path analysis. Also, the fact that constant 2 cannot be improved for even n was presented there. Before we state the main result in the direction of Szatzschneider conjecture, let us present a special case when the domination holds, which explains its relation with classic Lévy-Ottaviani inequality i.e. that for independent, symmetric random variables $Z_1, Z_2, \ldots, Z_n$ it holds true that

$$\mathbf{P}(\max_{1\le k\le n}\sum_{i=1}^{k} Z_i \ge u) \le 2\mathbf{P}(\sum_{i=1}^{n} Z_i \ge u).$$

Proposition 1. *Suppose that functions $a_i : [0,1] \to \mathbb{R}_+$ are of the form $a_i(t) = \alpha_i(t)a_i(1)$, where for all $t \in [0,1]$ $0 \le \alpha_n(t) \le \ldots \le \alpha_1(t) \le 1$. Then,*

$$\mathbf{P}(X \ge 1) \le 2\mathbf{P}(Y \ge 1).$$

Proof. Denote $S_k^a = \sum_{i=1}^k a_i(1)\varepsilon_i$. Obviously, $Y = S_n^a$. Then, by the Abel's inequality, we get

$$X = \sum_{i=1}^{n} a_i(1)\alpha_i(t)\varepsilon_i = \sum_{i=1}^{n}(\alpha_i(t) - \alpha_{i+1}(t))S_i^a \le \max_{1\le i\le n} S_i^a,$$

where we put $a_{n+1}(t) = 0$. Hence, by Lévy-Ottaviani inequality, we conclude that

$$\mathbf{P}(X \ge 1) \le \mathbf{P}(\max_{1\le k\le n} S_k^a \ge 1) \le 2\mathbf{P}(S_n^a \ge 1).$$

□

Remark 5. An example of functions satisfying the above condition are $a_i(t) = a_i(1)\mathbb{1}_{[t_i,1]}(t)$ for $0 \le t_1 \le \ldots \le t_n \le 1$.

The approach we propose allows to skip the two mentioned conditions. We will prove the following form of Szatzschneider's conjecture.

Theorem 15. *Let $a_i : [0,1] \to \mathbb{R}_+$, for $i = 1,2,\ldots,n$ be non-decreasing, right-continous functions and $n \ge 5$. Then for $u > 0$*

$$\mathbf{P}(\sup_{t\in[0,1]}\sum_{i=1}^{n} a_i(t)\varepsilon_i \ge 8u) \le 53\mathbf{P}(\sum_{i=1}^{n} a_i(1)\varepsilon_i \ge u).$$

This result is also a consequence of the concentration result (Theorem 16) which we prove in the next section. As we will explain the constant on the left hand side of the above inequality comes from the estimate on the $\mathbf{E}X$ which we obtain by using chaining method in the same manner as we did for Theorem 10. This will be presented in section 3.

Let's finish this section with the important comparison inequalities between the L^p-norms of X_t. Let's denote them by $\|X_t\|_p$. The first one is a hypercontraction (see e.g. [13, Chapter 3.4]) i.e. for $1 < q < p < \infty$

$$\|X_t\|_p \le \sqrt{\frac{p-1}{q-1}}\|X_t\|_q. \tag{3.1}$$

Moreover, we have comparison with the first moment which in the following form is due to Szarek [25]. We have

$$\mathbf{E}|X_t| \geq \frac{1}{\sqrt{2}}\|X_t\|_2 = \frac{1}{\sqrt{2}}\|t\| \tag{3.2}$$

It is easy to see that it extends to X in the sense that

$$\mathbf{E}X \geq (1/2\sqrt{2}) \sup_{t\in T} \|t\|.$$

The aim of section 3.3 is to prove that $\mathbf{E}X$ is actually comparable with $\|t^0\|$ in the Szatzschneider setting. It is an interesting and yet unfulfilled task to provide a geometrical description of sets T for which such a comparison occurs.

3.2 Concentration

We aim to prove a special form of concentration result.

Theorem 16. *Let $T = [0, t_1^0] \times \cdots \times [0, t_n^0]$ and $\varphi : \mathbb{R} \to [0, \infty)$ be any convex, increasing function. Then*

$$\mathbf{E}\varphi(X - \mathbf{E}X) \leq \mathbf{E}\varphi(Y). \tag{3.3}$$

Proof. Consider numbers $(b(t))_{t\in T}$ and define $\tilde{X} = \sup_{t\in T} \left(\sum_{i=1}^n t_i\varepsilon_i + b(t)\right)$. We will prove that

$$\mathbf{E}\varphi(\tilde{X} - \mathbf{E}\tilde{X}) \leq \mathbf{E}\varphi(Y).$$

and apply this result for $b \equiv 0$. We will proceed by induction. For $n = 0$ both sides equal 0. For $n \geq 1$, we will condtion on ε_1. To this end we define

$$\tilde{X}_+ = \sup_{t\in T}\left(t_1 + b(t) + \sum_{i=2}^n t_i\varepsilon_i\right) \quad \text{and} \quad \tilde{X}_- = \sup_{t\in T}\left(-t_1 + b(t) + \sum_{i=2}^n t_i\varepsilon_i\right).$$

Notice that $\mathbf{E}\tilde{X} = (\mathbf{E}\tilde{X}_- + \mathbf{E}\tilde{X}_+)/2$, so we can write

$$\mathbf{E}\varphi(X - \mathbf{E}X)$$
$$= \frac{1}{2}\left(\mathbf{E}\varphi\left(\tilde{X}_+ - \mathbf{E}\tilde{X}_+ + \frac{\mathbf{E}\tilde{X}_+ - \mathbf{E}\tilde{X}_-}{2}\right) + \mathbf{E}\varphi\left(\tilde{X}_- - \mathbf{E}\tilde{X}_- + \frac{\mathbf{E}\tilde{X}_- - \mathbf{E}\tilde{X}_+}{2}\right)\right). \tag{3.4}$$

Therefore, by the induction assumption used for convex increasing functions $x \mapsto \varphi(x + (\mathbf{E}\tilde{X}_+ - \mathbf{E}\tilde{X}_-)/2)$ and $x \mapsto \varphi(x + (\mathbf{E}\tilde{X}_- - \mathbf{E}\tilde{X}_+)/2)$ we have

$$
\begin{aligned}
&\mathbf{E}\varphi(\tilde{X} - \mathbf{E}\tilde{X}) \\
&\le \frac{1}{2}\left(\mathbf{E}\varphi\left(\sum_{i=2}^{n} t_i^0\varepsilon_i + \frac{\mathbf{E}\tilde{X}_+ - \mathbf{E}\tilde{X}_-}{2}\right) + \mathbf{E}\varphi\left(\sum_{i=2}^{n} t_i^0\varepsilon_i + \frac{\mathbf{E}\tilde{X}_- - \mathbf{E}\tilde{X}_+}{2}\right)\right) \\
&= \mathbf{E}\varphi\left(\frac{|\mathbf{E}\tilde{X}_+ - \mathbf{E}\tilde{X}_-|}{2}\varepsilon_1 + \sum_{i=2}^{n} t_i^0\varepsilon_i\right). \qquad (3.5)
\end{aligned}
$$

Observe that

$$\frac{|\mathbf{E}\tilde{X}_+ - \mathbf{E}\tilde{X}_-|}{2} \le \sup_{t\in T} t_1 = t_1^0$$

and thus using the contraction principle (Theorem 7) in the special case, when we condition on $\varepsilon_2, \dots, \varepsilon_n$ and consider a supremum over a single point we get

$$\mathbf{E}\varphi\left(\frac{|\mathbf{E}\tilde{X}_+ - \mathbf{E}\tilde{X}_-|}{2}\varepsilon_1 + \sum_{i=2}^{n} t_i^0\varepsilon_i\right) \le \mathbf{E}\varphi\left(\sum_{i=1}^{n} t_i^0\varepsilon_i\right). \qquad (3.6)$$

Combining (3.4),(3.5),(3.6) completes the proof. □

There are two functions which are of special interest. The first one will recover the strenghened concentration, while the other will lead to the main result of this work.

Corollary 6. *We have*

$$\mathbf{P}(|X - \mathbf{E}X| \ge u) \le 2e^{-\frac{u^2}{2\|t^0\|^2}}. \qquad (3.7)$$

Proof. Apply (3.3) for $\varphi(x) = e^{\lambda x}$, $\lambda \in \mathbb{R}$. □

Corollary 7. *Let $0 < \alpha \le 1$ and $u > 0$. Then,*

$$\mathbf{P}(X \ge \mathbf{E}X + (1+\alpha)u) \le \frac{4}{\alpha u}\mathbf{P}(Y \ge u)\mathbf{E}(Y)_+. \qquad (3.8)$$

Proof. Consider $\varphi(x) = (x-u)_+$. Then, by (3.3) we get that $\mathbf{E}(X - \mathbf{E}X - u)_+ \le \mathbf{E}(Y - u)_+$. We will show that

$$\alpha u \mathbf{P}(X \ge \mathbf{E}X + (1+\alpha)u) \le \mathbf{E}(X - \mathbf{E}X - u)_+ \qquad (3.9)$$

and

$$\mathbf{E}(Y - u)_+ \le 4\mathbf{P}(Y \ge u)\mathbf{E}(Y)_+. \qquad (3.10)$$

Notice that (3.9) follows simply from

$$\mathbf{E}(X - \mathbf{E}X - u)_+ \ge \mathbf{E}(X - \mathbf{E}X - u)_+ \mathbb{1}_{\{X - \mathbf{E}X \ge (1+\alpha)u\}} \ge \alpha u \mathbf{P}(X - \mathbf{E}X \ge (1+\alpha)u).$$

Moreover, (3.10) can be deduced from the Kahane's inequality (see e.g. [13, Proposition 1.4.1]). Indeed,

$$\begin{aligned}\mathbf{E}(Y-u)_+ &= \int_u^\infty \mathbf{P}(Y \geq t)dt = \int_0^\infty \mathbf{P}(Y \geq u+t)dt \\ &\leq 4\mathbf{P}(Y \geq u)\int_0^\infty \mathbf{P}(Y \geq t)dt = 4\mathbf{P}(Y \geq u)\mathbf{E}(Y)_+.\end{aligned}$$

□

Let's state the main result of this chapter.

Theorem 17. *Consider a subset $T \subseteq [0, t_1^0] \times \cdots \times [0, t_n^0]$ of $\mathbb{R}^n$. Let X and Y be as in Theorem 16. Suppose that there exists a positive constant C_1 such that $\mathbf{E}X \leq C_1\|t^0\|$. Then, for $u > 0$, $\alpha \in (0, 1]$, $\theta \in (0, 1)$*

$$\mathbf{P}(X \geq (\frac{C_1}{\sqrt{\theta}} + 1 + \alpha)u) \leq C_{\alpha,\theta}\mathbf{P}(Y \geq u), \tag{3.11}$$

where $C_{\alpha,\theta} = \max\{\frac{18}{(1-\theta)^2}, \frac{2}{\alpha\sqrt{\theta}}\}$.

Proof. Suppose that $u \leq \sqrt{\theta}\|t^0\|$. Notice that by (3.1) we have $(\mathbf{E}|Y|^2)^2/\mathbf{E}|Y|^4 \geq 1/9$. Hence, by the Paley-Zygmund inequality we get

$$\begin{aligned}\mathbf{P}(Y \geq u) &\geq \mathbf{P}(Y \geq \sqrt{\theta}\|t^0\|) = \frac{1}{2}\mathbf{P}(|Y| \geq \sqrt{\theta}\|t^0\|) = \frac{1}{2}\mathbf{P}(|Y|^2 \geq \theta\|t^0\|^2) \\ &= \frac{1}{2}\mathbf{P}(|Y|^2 \geq \theta\mathbf{E}|Y|^2) \geq \frac{1}{2}(1-\theta)^2\frac{(\mathbf{E}|Y|^2)^2}{\mathbf{E}|Y|^4} \geq \frac{(1-\theta)^2}{18},\end{aligned}$$

so trivially

$$\mathbf{P}(X \geq (\frac{C_1}{\sqrt{\theta}} + 1 + \alpha)u) \leq 1 \leq \frac{18}{(1-\theta)^2}\mathbf{P}(Y \geq u).$$

Now, consider $u \geq \sqrt{\theta}\|t^0\|$. Notice that $\mathbf{E}(Y)_+ = 1/2\mathbf{E}|Y| \leq 1/2\sqrt{\mathbf{E}|Y|^2} = 1/2\|t^0\|$. Hence by Corollary 7

$$\begin{aligned}\mathbf{P}(X \geq (\frac{C_1}{\sqrt{\theta}} + 1 + \alpha)u) &\leq \mathbf{P}(X \geq (\frac{\mathbf{E}X}{\sqrt{\theta}\|t^0\|} + 1 + \alpha)\sqrt{\theta}\|t^0\|) \\ &\leq \frac{4}{\alpha\sqrt{\theta}\|t^0\|}\mathbf{P}(Y \geq u)\frac{1}{2}\|t^0\|. \end{aligned} \tag{3.12}$$

This finishes the proof. □

Remark 6. Instead of using Kahane's inequality in Corollary 7 one can use [6, Lemma 7] to obtain that $\mathbf{P}(X \geq \mathbf{E}X + (1+\alpha)u) \leq \frac{16}{\alpha}\mathbf{P}(Y \geq u)$. Then by considering cases when u is less or greater than $(1/2\sqrt{2})\|t^0\|$ and applying (3.2) one can get

$$\mathbf{P}(X \geq (2\sqrt{2}C_1 + 2)u) \leq 16\mathbf{P}(Y \geq u). \tag{3.13}$$

3.3 Chaining

Theorem 18. *The following inequality holds* $\mathbf{E}X \le C\|a(1)\|$, *where* $C \le 4.45$.

Proof. The proof is based on the special choice of approximation nets T_k, $k \ge 0$. We denote the number of elements $|T_k| = N_k$, where N_k are numbers which we choose later. Define $T_k = \{u_0^k, u_1^k, \dots, u_{N_k-1}^k\}$ in the following way

$$u_l^k = \inf\{t \in [0,1] : \|a(t)\|^2 \ge \frac{l}{N_k}\|a(1)\|^2\}.$$

Since $a_i(t)$ are right continuous we have that

$$\frac{l}{N_k}\|a(1)\|^2 \le \|a(u_l^k)\|^2 \le \frac{l+1}{N_k}\|a(1)\|^2. \tag{3.14}$$

Moreover, $T_k \subset T_{k+1}$. Let us define $\pi_k(t) \in T_k$ as $\max\{u_l^k \in T_k : \ u_l^k \le t\}$. Therefore, if $t \in T_k$, $k \ge 1$ and $\pi_{k-1}(t) = u_l^{k-1}$ then

$$\frac{l}{N_{k-1}} \le \|a(\pi_{k-1}(t))\|^2 \le \|a(t)\|^2 < \frac{l+1}{N_{k-1}}.$$

As a consequence of the above inequality and monotonicity of each a_i we get the following crucial fact

$$\|a(t) - a(\pi_{k-1}(t))\|^2 \le \|a(\pi_k(t))\|^2 - \|a(\pi_{k-1}(t))\|^2 \le \frac{\|a(1)\|^2}{N_{k-1}}. \tag{3.15}$$

It is clear that $\bigcup_k T_k$ is dense in T. Fix K and consider points $t \in T_K$. Obviously, $\pi_K(t) = t$. Using backward induction we define t_k for $k = 0, 1, 2, \dots, K$ as $t_K = \pi_K(t) = t$ and for $k < K$, $t_k = \pi_k(t_{k+1})$. Note that $t_0 = 0$ for all $t \in T_K$. Before we state the main chaining argument we present two helpful inequalities. First, recall that from (3.1) we can bound any norm of X_t by $\|t\|$, namely $\|X_t\|_p \le \sqrt{p-1}\|t\|$. Also, (see proof of Theorem 10), we have for any constant $C \ge 1$ and $p \ge 2$

$$\mathbf{E}(\frac{X_t}{\|X_t\|_p} - C)_+ = \frac{1}{2}\mathbf{E}(\frac{|X_t|}{\|X_t\|_p} - C)_+ \le \frac{1}{2}\max_{x \ge C}\frac{1}{x^p}(x - C) \le \frac{1}{2}C\frac{1}{p-1}\left(\frac{p-1}{Cp}\right)^p. \tag{3.16}$$

We proceed to chaining (it follows the same lines as the proof of Theorem 10).

$$\begin{aligned}
\mathbf{E}X &= \lim_{K\to\infty} \mathbf{E} \sup_{t\in T_K} (X_t - X_0) = \lim_{K\to\infty} \mathbf{E} \sup_{t\in T_K} \sum_{k=1}^{K} (X_{t_k} - X_{t_{k-1}}) \\
&\leq \lim_{K\to\infty} \mathbf{E} \sup_{t\in T_K} \sum_{k=1}^{K} C_k \|X_{t_k} - X_{t_{k-1}}\|_{p_k} \left(1 + \left(\frac{X_{t_k} - X_{t_{k-1}}}{C_k \|X_{t_k} - X_{t_{k-1}}\|_{p_k}} - 1\right)_+\right) \\
&\leq \|a(1)\| \lim_{K\to\infty} \mathbf{E} \sup_{t\in T_K} \sum_{k=1}^{K} C_k \frac{(p_k-1)^{1/2}}{|T_{k-1}|^{1/2}} \left(1 + \left(\frac{X_{t_k} - X_{t_{k-1}}}{C_k \|X_{t_k} - X_{t_{k-1}}\|_{p_k}} - 1\right)_+\right) \qquad (3.17) \\
&\leq \|a(1)\| \lim_{K\to\infty} \sum_{k=1}^{K} C_k \frac{(p_k-1)^{1/2}}{|T_{k-1}|^{1/2}} \left(1 + \sum_{u\in T_k} \mathbf{E} \left(\frac{X_u - X_{\pi_{k-1}(u)}}{C_k \|X_u - X_{\pi_{k-1}(u)}\|_{p_k}} - 1\right)_+\right) \\
&= \|a(1)\| \lim_{K\to\infty} \sum_{k=1}^{K} C_k \frac{(p_k-1)^{1/2}}{|T_{k-1}|^{1/2}} \left(1 + \sum_{u\in T_k} \frac{1}{2C_k} \mathbf{E} \left(\frac{|X_u - X_{\pi_{k-1}(u)}|}{\|X_u - X_{\pi_{k-1}(u)}\|_{p_k}} - C_k\right)_+\right) \\
&\leq \|a(1)\| \lim_{K\to\infty} \sum_{k=1}^{K} C_k \frac{(p_k-1)^{1/2}}{|T_{k-1}|^{1/2}} \left(1 + \frac{1}{2C_k} |T_k| C_k \frac{1}{p_k - 1} \left(\frac{p_k - 1}{C_k p_k}\right)^{p_k}\right) \qquad (3.18)
\end{aligned}$$

where in (3.17) we used (3.15) and (3.1), while (3.18) follows from (3.16). It remains to choose parameters C_k, p_k and $|T_k|$ in the optimal way. For this we pick $C_1 = 1$ and $C_k = 2$ for $k \geq 2$. For each k we choose $p_k = 2^k$. We define $|T_k|$ iteratively so that $|T_0| = 1$ and $|T_k|$ it is the multiple of $|T_{k-1}|$ (to satisfy $T_{k-1} \subset T_k$) closest to the minimizer of the function

$$f(x) = \left(\frac{2^k - 1}{|T_{k-1}|}\right)^{\frac{1}{2}} \frac{1}{2^k - 1} \left(\frac{2^k - 1}{2^{k+1}}\right)^{2^k} x + 2\left(\frac{2^{k+1} - 1}{x}\right)^{\frac{1}{2}},$$

which is

$$x = ((2^{k+1} - 1)(2^k - 1))^{\frac{1}{3}} \left(\frac{2^k}{2^k - 1}\right)^{\frac{2}{3}2^k} |T_{k-1}|^{\frac{1}{3}} 2^{2^k}.$$

The result then follows by substituting values of C_k, p_k and T_k and a simple estimation. □

Proof of Theorem 15. We apply Theorem 17 with $\theta = (C_1/(7-\alpha))^2$ for $\alpha = 0.1$ and $C_1 = 4.45$. □

Remark 7. Constant inside the probability on the left hand side in Theorem 15 can be reduced to 6 in exchange for $C_{\alpha,\theta} \leq 430$. Alternatively, we can apply (3.13) to reduce constant on the right hand side to 16 with constant on the left equal to 14.6.

Remark 8. Notice that Corollary 7 implies that for big u (say $u > \mathbf{E}X/\epsilon$, $\epsilon > 0$ small) the result is close to the original conjecture. Namely, for $\alpha = \epsilon$ we get that

$$\mathbf{P}(X \geq (1+2\epsilon)u) \leq \mathbf{P}(X \geq \mathbf{E}X + (1+\alpha)u) \leq \frac{4\mathbf{E}(Y)_+}{\epsilon u} \mathbf{P}(Y \geq 1).$$

The constant $4\mathbf{E}(Y)_+/(\epsilon u)$ gets smaller with larger u we take. Obviously, the estimate works until u exceeds $\sum_{i=1}^n |a_i(t)|$.

Chapter 4

Infinitely divisible processes

4.1 Overview of the result

It is a long-term program to investigate properties of general Lévy-type processes. It is unexpectedly hard to provide some unified approach to this matter and what is usually done in practice is to explore the behaviour of specific examples separately. Initial attempt to deal with this task was done by M. Talagrand in 1993 [29], where the regularity of infinitely divisible processes was proved under additional technical assumption on the Lévy measure. It was conjectured that the result holds without it, however little progress on this problem has been done since then. The main difficulty was related to the fact that the Bernoulli Conjecture (see [2]) was still open then. Our goal is to prove that the Talagrand's conjecture holds true.
The essential technical tool in the study of infinitely divisible processes is its' series representation due to J. Rosiński [23]. Consider a σ-finite measure space (Ω, ν) and a Poisson point process of intensity ν on Ω (see e.g. [7], [14] for the introduction to the subject of such processes). Recall it is a random subset Π of Ω such that for any measurable subset A of finite intensity measure a cardinality of $A \cap \Pi$ denoted by $|A \cap \Pi|$ is a Poisson random variable of expectation $\nu(A) < \infty$ and for disjoint measurable subsets $A_1, \ldots, A_k$ random variables $(|A_i \cap \Pi|)_{1 \le i \le k}$ are independent. We denote elements of Π by $(Z_i)_{i \ge 1}$. By $(\varepsilon_i)_{i \ge 1}$ we denote a sequence of random signs (Bernoulli sequence) i.e. $\mathbb{P}(\varepsilon_i = \pm 1) = 1/2$ independent of $(Z_i)_{i \ge 1}$. It is a delicate matter whether the construction of elements of the random set is well-defined. However, we just refer to [23] or [32, Theorem 11.2.7] for details since what is crucial for our work is the existence of the series representation of infinitely divisible processes. For this reason and the fact that we want to work with such processes in full generality we use the following definition. It is different from a classical terminology used in this area as might be found for example in [24].

Definition 3. An infinitely divisible (symmetric, without Gaussian component) process is a collection $(X_t)_{t \in T}$ where T is a set of functions on Ω satisfying $\int_\Omega t^2 \wedge 1 d\nu < \infty$ for $t \in T$ and where $X_t = \sum_{i \ge 1} \varepsilon_i t(Z_i)$.

In order to see that the definition is legit we need a following property of a Poisson point process (see [32, Lemma 11.3.1]), which we also use later.

$$\mathbf{E} \sum_{i \ge 1} t(Z_i) = \int_\Omega t(\omega) \nu(d\omega). \tag{4.1}$$

To show that $\sum_{i\geq 1}\varepsilon_i t(Z_i)$ converges almost surely, we use (4.1) and see that $\mathbf{E}\sum_{i\geq 1} t(Z_i)^2\wedge 1<\infty$ by the assumption $\int_\Omega t^2\wedge 1 d\nu<\infty$. Moreover, this assumption implies that $\nu(\{\omega : |t(\omega)|\geq 1\})<\infty$, so there are finitely many i's with $|t(Z_i)|\geq 1$.

The main question concerns the behaviour of $\mathbf{E}\sup_{t\in T} X_t$. Let us remind that formally we consider

$$\sup_{F\subset T}\mathbf{E}\sup_{t\in F} X_t,$$

where the supremum runs over all finite subsets F of T. Usually, by considering a separable modification of X_t, $t\in T$, it is possible to guarantee that $\sup_{t\in T} X_t$ is a well-defined random variable (for the definition of a separable version of a process and a discussion of measurability of suprema in the general setting of not necessarily separable spaces see [18, Ch. 2]). In this case $\sup_{F\subset T}\mathbf{E}\sup_{t\in F} X_t$ coincides with the usual expectation of the supremum of X_t. We will assume that T is separable so the formal definition coincides with $\mathbf{E}\sup_{t\in T} X_t$. Our study is a part of the theory of suprema of stochastic processes developed by M. Talagrand which can be summarized as relating the size of the process to the appropriately measured size of the index set T. For this let us recall γ numbers. First, consider distances on T

$$d_\infty(s,t)=\sup_{\omega\in\Omega}|s(\omega)-t(\omega)|$$

and

$$d_2^2(s,t)=\int_\Omega (s(\omega)-t(\omega))^2\nu(d\omega).$$

Consider an admissible (see Definition 1) sequence of partitions $(\mathcal{A}_n)_{n\geq 0}$ of set T. By $A_n(t)$ we will denote the unique element of partition $\mathcal{A}_n$ that contains $t\in T$ and by $\Delta(\cdot)$ the diameter of set in distance d. Given $\alpha>0$ we have

$$\gamma_\alpha(T,d)=\inf\sup_{t\in T}\sum_{n\geq 0} 2^{n/\alpha}\Delta(A_n(t)),$$

where the infimum is taken over all admissible sequences.

We are ready to formulate upper bounds for $\mathbf{E}\sup_{t\in T} X_t$. We will follow the same steps as in the case of empirical and selector processes (see Section 1.3). One obvious bound follows from the inequality

$$|X_t|\leq\sum_{i\geq 1}|t(Z_i)|.$$

It motivates the definition of the process $(|X|_t)_{t\in T}$, where $|X|_t=\sum_{i\geq 1}|t(Z_i)|$ and a set T of functions on Ω satisfying $\int_\Omega |t|\wedge 1 d\nu<\infty$. Therefore, if it happens that $(|X|_t)_{t\in T}$ is bounded then obviously $\mathbf{E}\sup_{t\in T} X_t$ is bounded and we say that its boundedness owes nothing to cancellation. The other reason for the boundedness follows from the chaining, which we explain next. The fundamental idea of Talagrand is that these are the only reasons for the boundedness of the process. This applies not only to infinitely divisible proceeses and we give an account of this fact in Chapter 5.

An important feature of infinitely divisible processes is that they obey the Bernstein inequality. Before we state it we will need some properties of Poisson Point Processes (cf. [33, Chapter 10.2]). Firstly, if X is a Poisson random variable with mean a, then

$$\mathbf{E}\exp(\lambda X) = \exp(a(\exp(\lambda) - 1)).$$

Now, if A is a measurable set of finite measure, $t = c\mathbb{1}_A$ and $X = |A \cap \Pi|$, then

$$\mathbf{E}\exp\left(\lambda \sum_{i\geq 1} t(Z_i)\right) = \mathbf{E}(\lambda c)^X = \exp(\nu(A)(\lambda c - 1)).$$

Therefore, by independence of $|A_i \cap \Pi|$ for step functions t we have the following identity

$$\mathbf{E}\exp\left(\lambda \sum_{i\geq 1} t(Z_i)\right) = \exp\left(\int (\exp(\lambda t(\omega)) - 1)\nu(d\omega)\right), \tag{4.2}$$

which extends to bounded functions t satisfying $\int |t| \wedge 1 d\nu < \infty$ by approximation. We will see that the formula (4.2) in some sense replaces the independence of Z_i assumption as we had in the case of empirical processes for example. We move towards the Bernstein inequality. Firstly, for $t \in T$, denoting by $\mathbf{E}_\varepsilon$ the expectation in ε_i's, we have

$$\mathbf{E}_\varepsilon \exp\left(\lambda \sum_{i\geq 1} \varepsilon_i t(Z_i)\right) = \exp\left(\sum_{i\geq 1} \log(\cosh(\lambda t(Z_i)))\right),$$

so by taking the expectation and applying (4.2) we get

$$\mathbf{E}\exp\left(\lambda \sum_{i\geq 1} \varepsilon_i t(Z_i)\right) = \exp\left(\int (\cosh(\lambda t(\omega)) - 1)\nu(d\omega)\right).$$

Observe that $\cosh(\lambda t(\omega)) - 1 \leq \lambda^2 t^2(\omega)$ for $|\lambda t(\omega)| \leq 1$. Consequently, arguing just as in the proof of Theorem 6 we get

$$\mathbf{P}\left(\left|\sum_{i\geq 1} \varepsilon_i (s(Z_i) - t(Z_i))\right| \geq v\right) \leq 2\exp\left(-\frac{1}{L}\min\left(\frac{v^2}{d_2(s,t)^2}, \frac{v}{d_\infty(s,t)}\right)\right).$$

This togehter with Theorem 3 yields

Theorem 19. *We have*

$$\mathbf{E}\sup_{t\in T} X_t \leq L(\gamma_2(T, d_2) + \gamma_1(T, d_\infty)). \tag{4.3}$$

As mentioned, the main concept due to M. Talagrand is that $\mathbf{E}\sup_{t\in T} X_t$ is actually comparable with the two quantities described above. However, the initial version of this result requires the following condition on measure ν.

Definition 4 ($H(C_0, \delta)$ condition)**.** Consider $\delta > 0$, $C_0 > 0$. We say that a measure ν satisfies a condition $H(C_0, \delta)$ if for all $s, t \in T$, and all $u > 0$, $v \geq 1$ we have

$$\nu(\{\beta; |\beta(s) - \beta(t)| \geq uv\}) \leq C_0 v^{-1-\delta} \nu(\{\beta; |\beta(s) - \beta(t)| \geq u\}).$$

All α-stable distributions for $\alpha \in (1, 2)$ obey this condition and one can construct a large class of measures for which it also holds, however it excludes for example Dirac deltas. Such measures, whose mass is carried by a single point, are precisely the source of difficulty in going beyond the above tail condition. Dealing with them requires the same amount of effort as with Bernoulli processes (see [2]) which we will discuss in Section 2.
The following result is known as the Decoposition Theorem for infinitely divisible processes.

Theorem 20. *Suppose that $H(C_0, \delta)$ holds. Then we can decompose $T \subset T_1 + T_2$ in such a manner that*

$$\gamma_2(T_1, d_2) + \gamma_1(T_1, d_\infty) \leq L\mathbf{E} \sup_{t \in T} X_t \tag{4.4}$$

and

$$\mathbf{E} \sup_{t \in T_2} |X|_t \leq L\mathbf{E} \sup_{t \in T} X_t. \tag{4.5}$$

It was orginally proved in [29]. Despite the control of the measure ν the argument there is still very complex and demanding. The proof of Bernoulli Conjecture [2] provided a new hope for approaching infinitely divisible processes, but it has been only recently to make the proof in [29] more comprehensible. We will outline the new approach in the next section.
The main goal of this chapter is to prove Theorem 20 without $H(C_0, \delta)$ condition. Before describing the method for this let us explain briefly the meaning of Theorem 20. Observe that having the decomposition of $t \in T$ given by $t = t_1 + t_2$, $t_1 \in T_1, t_2 \in T_2$ and $T \subset T_1 + T_2$ we can write

$$\mathbf{E} \sup_{t \in T} X_t \leq \mathbf{E} \sup_{t \in T} X_{t_1} + \mathbf{E} \sup_{t \in T} X_{t_2} \leq \mathbf{E} \sup_{t \in T_1} X_t + \mathbf{E} \sup_{t \in T_2} X_t.$$

Theorem 20 states that infinitely divisible process can be splitted into two parts. $\mathbf{E} \sup_{t \in T_1} X_t$ is the one explained through chaining (by (4.4) and (4.3)) and $\mathbf{E} \sup_{t \in T_2} X_t$ is bounded since $\mathbf{E} \sup_{t \in T_2} |X_t| \leq \mathbf{E} \sup_{t \in T_2} |X|_t$ and then by (4.5). The first part should be considered as the one where cancellations between terms occur while in the second there are no cancellations. It is a general phenomenon which we will discuss further in Chapter 5 by looking at empirical processes and selector processes (see [32, Chapter 9 and Chapter 12]).

4.2 Tools

4.2.1 Lower bounds from partition

Let us define mappings which are square of distances, but we will refer to them as distances for simplicity. For $j \geq 1$, $s, t \in T$ let

$$\varphi_j(s,t) = \int_\Omega r^{2j}|s(\omega) - t(\omega)|^2 \wedge 1\nu(d\omega) \tag{4.6}$$

and we denote by $B_j(t, r)$ a ball in the distance φ_j centred at t with radius r. Let us also assume that

$$s \neq t \implies \lim_{j\to\infty} \varphi_j(s,t) = \infty. \tag{4.7}$$

The first tool is the following (see [33, Theorem 6.6.1], [32, Theorem 5.2.6]). We do not provide the proof of this Theorem here. However, the proof of Theorem 29 mimics this argument.

Theorem 21. *Consider a countable set T of measurable functions on Ω, a number $r \geq 4$ and assume $0 \in T$. Consider an admissible sequence of partitions $(\mathcal{A}_n)_{n\geq 0}$ of T, and for $A \in \mathcal{A}_n$ consider $j_n \in \mathbb{Z}$, with the following properties, where $u > 0$ is a parameter*

$$A \in \mathcal{A}_n, B \in \mathcal{A}_{n-1}, A \subset B \implies j_n(A) \geq j_{n-1}(B),$$

$$\forall s,t \in A \in \mathcal{A}_n,\ \ \varphi_{j_n(A)}(s,t) \leq u2^n.$$

Then we can write $T \subset T' + T'' + T'''$ where $0 \in T'$ and

$$\gamma_2(T', d_2) \leq L\sqrt{u} \sup_{t\in T} \sum_{n\geq 0} 2^n r^{-j_n(A_n(t))}, \tag{4.8}$$

$$\gamma_1(T', d_\infty) \leq L \sup_{t\in T} \sum_{n\geq 0} 2^n r^{-j_n(A_n(t))} \tag{4.9}$$

$$\forall t \in T'', \|t\|_1 \leq Lu \sup_{t\in T} \sum_{n\geq 0} 2^n r^{-j_n(A_n(t))}. \tag{4.10}$$

Moreover,

$$\forall t \in T''', \exists s \in T,\ \ |t| \leq 5|s|\mathbb{1}_{\{2|s| \geq r^{-j_0(A_0(t))}\}}. \tag{4.11}$$

What Theorem 21 effectively says is that if we can provide an admissible sequence of partitions of the set T together with the sequence (j_n) satisfying described properties and most importantly that if

$$\sup_{t\in T} \sum_{n\geq 0} 2^n r^{-j_n(A_n(t))} \leq \mathbf{E} \sup_{t\in T} X_t \tag{4.12}$$

then we are basically done with the proof of Theorem 20. The remaining steps are relatively simple. First, we show the Giné-Zinn type of result ([32, Theorem 11.5.1], [33, Theorem 9.7.1]) the proof of which should be compared with the one of Theorem 9. Again we will need to deal with the fact that Z_i's are not independent, so we cannot simply repeat the proof in the case of empirical processes. We will need some extra tool which is the following Lemma (see [33, Lemma 10.2.1], cf.[14, Proposition 3.8]).

Lemma 2. *Consider a Poisson point process of intensity ν and a set A with $\nu(A) > 0$. Given $|A \cap \Pi| = N$, the set $A \cap \Pi$ has the same distribution as the set $\{X_1, X_2, \ldots, X_N\}$, where the variables X_i are independent and distributed according to the probability λ on A given by $\lambda(B) = \nu(A \cap B)/\nu(A)$ for $B \subset A$.*

Theorem 22.

$$\mathbf{E} \sup_{t \in T} |X|_t \leq \sup_{t \in T} \int_\Omega |t(\omega)| \nu(d\omega) + 4\mathbf{E} \sup_{t \in T} |X_t|. \tag{4.13}$$

Proof. Consider a measurable subset $A \subset \Omega$ with $\nu(A) < \infty$. Consider a sequence $\{X_1, \ldots, X_N\}$ of independent random variables distributed according to probability λ on A given for $B \subset A$ by $\lambda(B) = \nu(A \cap B)/\nu(A)$. As usual, by $(\varepsilon_i)_{i \geq 1}$ we denote an independent Bernoulli sequence, independent of X_i's. Then, by the same argument as for empirical processes (1.31) we get

$$\mathbf{E} \sup_{t \in T} \sum_{i \leq N} |t(X_i)| \leq \frac{N}{\nu(A)} \sup_{t \in T} \int_A |t(\omega)| \nu(d\omega) + 4\mathbf{E} \sup_{t \in T} \left| \sum_{i \leq N} \varepsilon_i t(X_i) \right|. \tag{4.14}$$

Now, we want to apply Lemma 2. For this consider a Poisson point process $(Z_i)_{i \geq 1}$ of intensity measure ν. Given $N = |\{i \geq 1 : Z_i \in A\}|$ we take the expectation in (4.14) and use the fact that $\mathbf{E}N = \nu(A)$ to get

$$\mathbf{E} \sup_{t \in T} \sum_{i \leq N} |t(X_i)| \mathbb{1}_A(Z_i) \leq \sup_{t \in T} \int_A |t(\omega)| \nu(d\omega) + 4\mathbf{E} \sup_{t \in T} \left| \sum_{i \leq N} \varepsilon_i t(X_i) \mathbb{1}_A(Z_i) \right|. \tag{4.15}$$

Obviously,

$$\int_A |t(\omega)| \nu(d\omega) \leq \int_\Omega |t(\omega)| \nu(d\omega).$$

For the second part we apply Jensen's inequality by taking the expectation in those ε_i's at which $\mathbb{1}_A(Z_i) = 0$ outside the absolute value and the supremum and get

$$\mathbf{E} \sup_{t \in T} \left| \sum_{i \leq N} \varepsilon_i t(X_i) \mathbb{1}_A(Z_i) \right| \leq \mathbf{E} \sup_{t \in T} \left| \sum_{i \leq N} \varepsilon_i t(X_i) \right|,$$

which finishes the proof by (4.15).

□

We then use Theorem 21 to write $T_1 = T'$ and $T_2 = T'' + T'''$ and the last piece is to show that

$$\sup_{t \in T_2} \int_\Omega |t(\omega)| \nu(d\omega) \leq L\mathbf{E} \sup_{t \in T} X_t. \tag{4.16}$$

As introduced in Chapter 1 (see Theorem 2), providing an admissible sequence of partitions such that the series

$$\sum_{n\geq 0} 2^n r^{-j_n(A_n(t))}$$

is a lower bound for some functional of the set T is a central tool in Talagrand's machinery (see [33, Theorem 7.1.2], [32, Theorem 10.1.2]). Let emphasize again that the existance of the admissible sequence depends on whether the appropriate functional satisfies the growth condition of Definition 2 or not. Let us repeat this notion (see also [33, Section 7] and [32, Section 10]) so that we can expose the importance of the $H(C_0, \delta)$ condition in the Talagrand's initial approach and motivate our strategy. Recall the distance φ_j given by (4.6).

Definition 5. We say that functionals $F_{n,j}$ satisfy the growth condition for $r = 2^\kappa$, $\kappa \in \mathbb{Z}$, if the following occurs. Consider any $j \in \mathbb{Z}$, any $n \geq 1$ and $m = N_n$. Consider any sets $(H_l)_{1\leq l\leq m}$ that are separated in the following sense: there exist points $u, t_1, \ldots, t_m$ in T for which $H_l \subset B_{j+2}(t_l, 2^{n+\kappa})$ and

$$\forall l, l' \leq m, \;\; l \neq l', \;\; \varphi_{j+1}(t_l, t_{l'}) \geq 2^{n+1}$$

$$\forall l \leq m, \;\; t_l \in B_j(u, 2^n).$$

Then,

$$F_{n,j}\Big(\bigcup_{l\leq m} H_l\Big) \geq 2^n r^{-j-1} + \min_{l\leq m} F_{n+1,j+1}(H_l).$$

In order to define the functional used in the proof of Theorem 20 we need more structure related to the Bernoulli process. The next section is devoted to explain how techniques used by Talagrand in his new proof of Theorem 20 [33, Section 10.8] are helpful for us to establish the general form of Theorem 20.

4.2.2 Partition for Bernoulli process

First, observe that conditionally on $(Z_i)_{i\geq 1}$, X_t is a Bernoulli process defined as a collection $(Y_t)_{t\in S}$, where $Y_t = \sum_{i\geq 1} \varepsilon_i t_i$ and $S \subset \ell^2$. Let us restate here the fundamental result due to W. Bednorz and R. Latała called Bernoulli Theorem [2] because we will use a little different notation than in Chapter 1.

Theorem 23. *Let $S \subset \ell^2$. There exists a universal constant L such that*

$$\inf\{\gamma_2(S_1, d) + \sup_{t\in S_2} \|t\|_1; S \subset S_1 + S_2\} \leq L\mathbf{E}\sup_{t\in S} Y_t,$$

where d denotes the Euclidean distance and $\|\cdot\|_1$ is a ℓ^1-norm.

What might be not noticed by looking at the above statement is the special construction crucial for its proof. Namely, it provides the admissible sequence of partitions $(\mathcal{A}_n)_{n\geq 0}$ and numbers $(j_n(A_n(t)))_{n\geq 0}$ such that

$$\sup_{t\in S} \sum_{n\geq 0} 2^n r^{-j_n(A_n(t))} \leq L\mathbf{E}\sup_{t\in S} Y_t \tag{4.17}$$

Recent developments of M. Talagrand (see [33, Chapter 8.6]) allow to provide a lower bound for $\mathbf{E}\sup_{t\in S} Y_t$ in terms of a functional which depends only on the set S, not on decomposition $S_1 + S_2$. The significance of this result should be compared with constructing the majorising measure for Gaussian processes (cf. [28], [33, Chapters 2.6, 4.1], [32, Chapter 2.4]). In particular it allows to prove Theorem 20 without $H(C_0, \delta)$. For this reason and for the clarity of our argument we restate this construction here. It can be summarized as follows. Recall that by N_n we denote a number 2^{2^n}.

Theorem 24. *Let $S \subset \ell^2$ and μ be a probability measure on S. Fix a positive number $r \geq 4$. Consider the following sequence of square distances for $j \in \mathbb{Z}$*

$$\forall s,t \in S \ \ \tilde{\varphi}_j(s,t) = \sum_{i\geq 1} (r^{2j}|t_i - s_i|^2) \wedge 1. \tag{4.18}$$

Now, let k_0 be the largest integer such that the diameter of S in the Euclidean distance ($\Delta(S, d_2)$) does not exceed r^{-k_0}. For $t \in S$ define $k_0(t) = k_0$ and for $n \geq 1$

$$k_n(t) = \sup\{j \in \mathbb{Z}\ ;\ \mu(\{s \in T, \tilde{\varphi}_j(s,t) \leq 2^n\}) \geq N_{n-1}^{-1}\}. \tag{4.19}$$

Let

$$I_\mu(t) = \sum_{n\geq 0} 2^n r^{-k_n(t)}. \tag{4.20}$$

Then there exists a universal constant L such that

$$\int_S I_\mu(t)\mu(dt) \leq L\mathbf{E}\sup_{t\in S} Y_t. \tag{4.21}$$

The proof of Theorem 24 involves two steps. We will adapt the notation from [33]. Put $b^*(S) = \inf\{\gamma_2(S_1, d) + \sup_{t\in S_2}\|t\|_1; S \subset S_1 + S_2\}$. Let $\bar{b}(S)$ be the infimum of numbers $\sup_{t\in S}\sum_{n\geq 0} 2^n r^{-j_n(A_n(t))}$ such that there exists an admissible sequence $(\mathcal{A}_n)_{n\geq 0}$ of partititons of S, and for $A \in \mathcal{A}_n$ an integer $j_n(A)$ satisfying

$$s,t \in A \implies \tilde{\varphi}_{j_n(A)}(s,t) \leq 2^n \tag{4.22}$$

and

$$\Delta(S, d_2) \leq r^{-j_0(T)}.$$

The first step is to show that [33, Theorem 8.6.2]

$$\bar{b}(S) \leq Lb^*(S). \tag{4.23}$$

It follows from the subadditivity of $\bar{b}$ (cf. proof of (2.6)) i.e. for $T, T' \subset \ell^2$

$$\bar{b}(T + T') \leq \bar{b}(T) + \bar{b}(T')$$

and then the following estimates. First, $\bar{b}(B_1) \leq Lr$, where $B_1 = \{t \in \ell^2 : \sum_{i\geq 1}|t_i| \leq 1\}$ (see [33, Theorem 17.2.1], which we do not quote here since it adds a little to the whole story and is quite technical). Moreover, $\bar{b}(T) \leq L\gamma_2(T, d_2)$, which is a consequence of the fact that $\gamma_2(T, d_2)$ satisfies the growth condition of Definition 2 combined with Theorem 2. The second step is the following Lemma (see [33, Proposition 8.6.7]), which we quote together with the proof for the instructional purposes.

Lemma 3. *Given any probability measure μ on S we have*

$$\int_S I_\mu(t)\mu(dt) \le L\bar{b}(S), \tag{4.24}$$

where $I_\mu(t)$ is as in (4.20).

Proof. Consider an admissible sequence $(\mathcal{A}_n)_{n\ge 0}$ of partitions of S and for $A \in \mathcal{A}_n$ an integer $j_n(A)$ as in (4.22) such that

$$\sup_{t\in S}\sum_{n\ge 0} 2^n r^{-j_n(A_n(t))} \le 2\bar{b}(S).$$

By the definition of k_0 it follows that $\tilde{\varphi}_{k_0+1}(s,t) > 1$ and since $\tilde{\varphi}_{k_0+1}(s,t) \le r^{2(k_0+1)}d_2(s,t)^2$ we have $r^{-k_0-1} < \Delta(S,d_2)$. Hence, since $\Delta(S,d_2) \le r^{-j_0(T)}$,

$$r^{-k_0} \le r^{-j_0(T)} \le \sup_{t\in S}\sum_{n\ge 0} 2^n r^{-j_n(A_n(t))}.$$

Now, for $n \ge 1$, $A \in \mathcal{A}_n$ and $t \in A$

$$A \subset \{s \in S:\ \tilde{\varphi}_{j_n(A)}(s,t) \le 2^n\} \subset \{s \in S:\ \tilde{\varphi}_{j_n(A)}(s,t) \le 2^{n+1}\}.$$

So, if $\mu(A) \ge N_{n+1}^{-1}$ then $k_{n+1}(t) \ge j_n(A)$, hence

$$\int_A 2^{n+1} r^{-k_{n+1}(t)}\mu(dt) \le 2\int_A 2^n r^{-j_n(A_n(t))}\mu(dt).$$

On the other hand, if $\mu(A) < N_{n+1}^{-1}$

$$\int_A 2^{n+1} r^{-k_{n+1}(t)}\mu(dt) < 2^{n+1} r^{-k_0} N_{n+1}^{-1}.$$

Summation over $A \in \mathcal{A}_n$ and then over $n \ge 0$ yields

$$\int_T \sum_{n\ge 1} 2^n r^{-k_n(t)} \le L\sup_{t\in S}\sum_{n\ge 0} 2^n r^{-j_n(A_n(t))} + Lr^{-k_0},$$

which finishes the proof, because for $n = 0, 1$

$$2^n r^{-k_n(t)} \le Lr^{-k_0} \le L\sup_{t\in S}\sum_{n\ge 0} 2^n r^{-j_n(A_n(t))}.$$

□

Having prepared the lower bound for the Bernoulli process we are ready to come back to infinitely divisible processes. Consider $I_{\mu,Z}$ be defined as in (4.20) but for $s,t \in S$ and

$$\tilde{\varphi}_j(s,t) = \sum_{i\ge 1}(r^{2j}|t(Z_i) - s(Z_i)|^2) \wedge 1$$

by taking expectation with respect to the process $(Z_i)_{i\geq 1}$ in (4.21) we have

$$\mathbf{E}\int_T I_{\mu,Z}(t)\mu(dt) \leq L\mathbf{E}\sup_{t\in T} X_t. \tag{4.25}$$

The next step is to repeat the construction provided in Theorem 24 for the set T. Using the distance (4.6) we define numbers $(j_k)_{k\geq 0}$ associated to each point $t\in T$. Let

$$j_0 = \sup\{j\in\mathbb{Z} : \forall s,t\in T, \varphi_j(s,t)\leq 4\}. \tag{4.26}$$

The use of constant 4 in the above will become apparent with the next Lemma. Given any probability measure μ on T we define for any integer $n\geq 0$ and $t\in T$

$$j_0^\mu(t) = j_0$$

and

$$j_n^\mu(t) = \sup\{j\in\mathbb{Z} : \mu(B_j(t,2^n)) > N_n^{-1}\}. \tag{4.27}$$

We also define

$$J_\mu(t) = \sum_{n\geq 0} 2^n r^{-j_n^\mu(t)}.$$

The crucial fact (see [33, Lemma 9.3.2]) is the following Lemma the proof of which we include again for the sake of completeness.

Lemma 4. *For each $t\in T$ we have $J_\mu(t)\leq L\mathbf{E}I_{\mu,Z}(t)$.*

Proof. Fix probability measure μ on T. To simplify the notation put $j_n(t) = j_n^\mu(t)$. Observe that by (4.1) $\mathbf{E}\tilde{\varphi}(s,t) = \varphi(s,t)$. It is easy to verify (we postpone the proof of this fact until Chapter 5, see also [32, Lemma 7.4.3], [33, Lemma 10.2.2]) that for any function $0\leq f\leq 1$ and constant A such that $4A\leq \int f d\nu$ we have

$$\mathbf{P}\left(\sum_{i\geq 1} f(Z_i)\leq A\right)\leq \exp(-A). \tag{4.28}$$

We aim to prove that

$$\mathbf{P}(k_0\leq j_0)\geq \frac{1}{L} \tag{4.29}$$

and for $n\geq 3$

$$\mathbf{P}(k_{n-2}(t)\leq j_n(t))\geq \frac{1}{2}. \tag{4.30}$$

These relations imply respectively that $\mathbf{E}r^{-k_0}\geq r^{-j_0}/L$ and for $n\geq 3$,

$$L\mathbf{E}2^{n-3}r^{-k_{n-3}(t)}\geq 2^n r^{-j_n(t)}.$$

Summing these inequalities and using for $n\leq 2$ that $2^n r^{-j_n(t)}\leq Lr^{-j_0}\leq \mathbf{E}2^n r^{-k_n(t)}$ yields the result.
To prove (4.29) we use (4.28). By the definition of j_0, $\varphi_{j_0+1}(s,t) > 4$ and therefore

$$\mathbf{P}(\tilde{\varphi}_{j_0+1}(s,t) > 1)\leq 1-\exp(-1).$$

By the definition of k_0 the event $\tilde{\varphi}_{j_0+1}(s,t) > 1$ implies that $k_0 \leq j_0$.
Now we argue that (4.30) holds. By the definition of $k_n(t)$ we have

$$\mu(\{s \in T:\ \varphi_{k_n(t)+1}(s,t) \leq 2^n\}) \leq N_n^{-1}.$$

On the other hand, by (4.28), $\varphi_{k_n(t)+1}(s,t) \geq 2^n$ implies that

$$\mathbf{P}(\tilde{\varphi}_{k_n(t)+1}(s,t) \leq 2^{n-2}) \leq \exp(-2^{n-2}) \leq N_{n-2}^{-1}.$$

Hence,

$$\mathbf{E}\mu(\{s \in T:\ \varphi_{k_n(t)+1}(s,t) \geq 2^n, \tilde{\varphi}_{k_n(t)+1}(s,t) \leq 2^{n-2}\}) \leq N_{n-2}^{-1}$$

and by Markov inequality we have

$$\mathbf{P}(\mu(\{s \in T:\ \varphi_{k_n(t)+1}(s,t) \geq 2^n, \tilde{\varphi}_{k_n(t)+1}(s,t) \leq 2^{n-2}\}) \leq 2N_{n-2}^{-1}) \geq \frac{1}{2}.$$

Therefore, conditioned on such event we get that

$$\mu(\{s \in T:\ \tilde{\varphi}_{k_n(t)+1}(s,t) \geq 2^{n-2}\}) \leq N_n^{-1} + 2N_{n-2}^{-1} < N_{n-3}^{-1}$$

and in turn $k_{n-3}(t) \leq j_n(t)$. This finishes the proof.

□

Lemma 4 together with (4.25) imply that

$$\int_T J_\mu(t)\mu(dt) \leq L \int_T \mathbf{E} I_{\mu,Z}(t)\mu(dt) = L\mathbf{E} \int_T I_{\mu,Z}(t)\mu(dt) \leq L\mathbf{E} \sup_{t\in T} X_t. \quad (4.31)$$

The functional that can be used for the proof of Theorem 20 is given for $A \subset T$ by

$$F_n(A) = F_{n,j}(A) = \sup_{\mu(A)=1} \inf_{t\in A} J_{\mu,n}(t),$$

where

$$J_{\mu,n}(t) = \sum_{k\geq n} 2^k r^{-j_k^\mu(t)}.$$

The idea was to prove that above functionals satisfy the growth condition of Definition 2, which in turn will induce the existence of the admissible sequence. The critical point is that to show the growth condition (see [32, Proposition 11.6.10]) one needs to show that $B_{j+2}(t_l, 2^{n+\kappa}) \subset B_{j+1}(t_l, 2^{n-4})$, which is a consequence of $H(C_0,\delta)$ condition.
Since there is little hope for proving the growth condition without $H(C_0,\delta)$ condition the other strategy is imposed for the proof of general Decomposition Theorem. First, we deduce from the fact that (4.31) holds for any probability measure that there is some measure say μ_0 for which we have

$$\sup_{t\in T} J_{\mu_0}(t) \leq L\mathbf{E}\sup_{t\in T} X_t.$$

This could be done by a standard argument using Hahn-Banach theorem provided that $J_{\mu,0}(t)$ is convex as a function of μ. Proving the convexity is highly non-trivial task which therefore is replaced by a modified reasoning giving a rise to the most technical part of the proof. It is a subject of the next section.

4.3 The Separation Theorem

We start with the main consequence of (4.31). It should be compared with [30, Propodition 3.2], [33, Lemma 13.3.9], [32, Lemma 13.1.4]. Fix a finite subset F of T,

Lemma 5. *Consider a number $a > 0$. Suppose that $\mathcal{C}$ is a closed, convex subset of real-valued functions on a finite set F. Assume that $f \in \mathcal{C}$, $g \geq f \implies g \in \mathcal{C}$ and for each probability measure μ on F there exists $f \in \mathcal{C}$ such that $\int f d\mu \leq a$. Then a constant function $\mathbf{a} \equiv a$ belongs to $\mathcal{C}$.*

Proof. Suppose that $\mathbf{a} \notin \mathcal{C}$. Then by the Hahn-Banach theorem there exists a linear functional φ on the space of functions on T such that for each $f \in \mathcal{C}$,

$$\varphi(f) > \varphi(\mathbf{a}). \tag{4.32}$$

Consider a function g on T with $g \geq 0$ and a number $\lambda > 0$ so that $f + \lambda g \in \mathcal{C}$ and $\varphi(f) + \lambda\varphi(g) > \varphi(\mathbf{a})$. This implies that φ is non-negative, so there is a probability measure μ_0 on T and a number $\beta \geq 0$ such that for each non-negative function g on T we have $\varphi(g) = \beta \int g d\mu_0$, in particular $\varphi(\mathbf{a}) = \beta a$. By the assumption there is $f \in \mathcal{C}$ such that $\int f d\mu_0 \leq a$, so $\varphi(\mathbf{a}) \leq \beta a = \varphi(\mathbf{a})$ contradicting (4.32). Hence, $\mathbf{a} \in \mathcal{C}$.

□

Theorem 25. *There exists a positive integer M, sequence of non-negative numbers $(\alpha_i)_{i\leq M}$ with $\sum_{i\leq M} \alpha_i = 1$ and a sequence $(\mu_i)_{i\leq M}$ of probability measures on a finite subset F of T such that for each $t \in F$.*

$$\sum_{i\leq M} \alpha_i J_{\mu_i}(t) \leq L\mathbf{E} \sup_{t\in T} X_t. \tag{4.33}$$

Proof. Consider a closed, convex hull $\mathcal{C}$ of the set of functions f for which there exists a probability measure μ on T such that for each $t \in T$, $f(t) \geq J_\mu(t)$. Given any probability measure μ on T, the function $f(t) = J_\mu(t)$ belongs to $\mathcal{C}$ and by (4.31) we have $\int_F f(t)\mu(dt) \leq a := L\mathbf{E}\sup_{t\in T} X_t$. The conclusion follows from Lemma 5.

□

So far, we have proved that for any finite subset $F \subset T$ we have

$$\sup_{t\in F} \sum_{i\leq M} \alpha_i \sum_{n\geq 0} 2^n r^{-j_n^{\mu_i}(t)} \leq L\mathbf{E} \sup_{t\in T} X_t.$$

We aim to show that there is a single probability measure on F and integers j_n depending only on this measure such that $\sup_{t\in F} \sum_{n\geq 0} 2^n r^{-j_n(t)} \leq K\mathbf{E}\sup_{t\in T} X_t.$. For a fixed finite set $F \subset T$ and the seqeuces (α_i) of non-negative numbers and (μ_i) of probability measures from Theorem 25 we define

$$\mu = \sum_{i \le M} \alpha_i \mu_i \tag{4.34}$$

and for each $t \in T$ numbers $j_n(t)$ with j_0 as in (4.26) and for $n \ge 1$

$$r^{-j_n(t)-1} < \sum_{i \le M} \alpha_i r^{-j_n^{\mu_i}(t)} \le r^{-j_n(t)}. \tag{4.35}$$

Then, it is a matter of changing the order of summation to notice that

$$t \in F \implies \sum_{n \ge 0} 2^n r^{-j_n(t)} \le r \sum_{i \le M} \alpha_i \sum_{n \ge 0} 2^n r^{-j_n^{\mu_i}(t)} \tag{4.36}$$

since j_0 is fixed. What is crucial for the further construction is that j_n's defined in (4.35) preserve the defining property of $j_n^{\mu_i}$'s given by (4.27).

Lemma 6. *Fix $n \ge 0$ and $t \in T$. For μ as in (4.34) and $j_n(t)$ as in (4.35) we have*

$$\mu(B_{j_n(t)}(t, 2^n)) \ge \frac{2}{3}\frac{1}{N_n} \ge \frac{1}{N_{n+1}}. \tag{4.37}$$

Proof. For $n = 0$ it is straightforward from (4.26). For $n \ge 1$ observe that if $j_n^{\mu_i}(t) \ge j_n(t)$, then $\mu_i(B_{j_n(t)-1}(t, 2^n)) \ge 1/N_n$. Define $\beta_j = \sum_i \alpha_i \mathbb{1}_{\{j_n^{\mu_i}(t)=j\}}$. Certainly, $\sum_{j \ge 1} \beta_j = 1$. Moreover,

$$\mu(B_{j_n(t)}(t, 2^n)) \ge \sum_{j \ge j_n(t)} \beta_j \frac{1}{N_n}.$$

Now, we will argue that $\sum_{j < j_n(t)} \beta_j$ can be bounded from above so that (4.37) follows. From the definition of $j_n(t)$ we have

$$\sum_{j < j_n(t)} \beta_j r^- j \le r^{-j_n(t)},$$

which implies that $\beta_j r^{-j} \le r^{-j_n(t)}$ for each $j < j_n(t)$ and in turn that $\beta_j \le r^{-j_n(t)+j}$. Hence, since $r \ge 4$,

$$\sum_{j < j_n(t)} \beta_j \le \sum_{j < j_n(t)} r^{j - j_n(t)} \le \sum_{l \ge 1} r^{-l} \le \frac{1}{3}$$

which finishes the argument.

□

Now, consider numbers $j_n^{\mu}(t)$ defined in (4.27) for the measure μ of (4.34). By (4.37) we have that $j_{n+1}^{\mu}(t) \ge j_n(t)$. Hence, by (4.36), (4.33) and the fact that $j_0^{\mu} = j_0 = j_0^{\mu_i}$ it follows that for any $t \in F$

$$\sum_{n \ge 0} 2^n r^{-j_n^{\mu}(t)} \le L \sum_{n \ge 0} 2^n r^{-j_n(t)} \le L\mathbf{E} \sup_{t \in T} X_t. \tag{4.38}$$

4.4 The Decomposition Theorem

What we achieved in the previous section is the lower bound from a majorising measure μ which depends on the finite set F. Namely,

$$\sup_{t\in F} J_\mu(t) \le L\mathbf{E}\sup_{t\in T} X_t.$$

We aim to replace the lower bound by the sum depending on the admissible sequence of partitions of the whole set T as in Theorem 21. Before stating the result observe (using $(a+b)^2 \le 2(a^2+b^2)$) that by the definition (4.6) we have the following form of the triangle inequality. For $s,t,x\in T$

$$\varphi_j(s,t) \le 2(\varphi_j(s,x)+\varphi_j(x,t)) \tag{4.39}$$

and as a consequnce

$$\forall s,t\in T,\ \varphi_j(s,t) > 4a > 0 \implies B_j(s,a)\cap B_j(t,a)=\emptyset. \tag{4.40}$$

First, let us refine the definition of $j_n^\mu(t)$ so that they form a non-decreasing sequence which can increase by at most 1. Define

$$\tilde{j}_n^\mu(t) = \min_{0\le p\le n}(j_p^\mu(t)+n-p).$$

In this way we have $\tilde{j}_0^\mu(t) = j_0^\mu(t)$ and for $n\ge 1$

$$\tilde{j}_n^\mu(t) \le \tilde{j}_{n+1}^\mu(t) \le \tilde{j}_n^\mu(t)+1. \tag{4.41}$$

Moreover, $\tilde{j}_n^\mu(t) \le j_n^\mu(t)$, so

$$\mu(B_{\tilde{j}_n^\mu(t)}(t,2^n)) \ge N_n^{-1}. \tag{4.42}$$

Finally, for $t\in F$, since $r\ge 4$ we have

$$\sum_{n\ge 0} 2^n r^{-\tilde{j}_n^\mu(t)} \le \sum_{n\ge 0} 2^n \sum_{0\le p\le n} r^{-j_p^\mu(t)-n+p} = \sum_{p\ge 0} 2^p r^{-j_p^\mu(t)} \sum_{n\ge p}\left(\frac{2}{r}\right)^{n-p} \le 2J_\mu(t). \tag{4.43}$$

Theorem 26. *There exists an admissible sequence of partititons $(\mathcal{A}_n)_{n\ge 0}$ of F and for $A\in\mathcal{A}_n$ there exists an integer $j_n(A)$ such that for each $t\in T$*

$$\sum_{n\ge 0} 2^n r^{-j_n(A_n(t))} \le L\sum_{n\ge 0} 2^n r^{-\tilde{j}_n^\mu(t)}. \tag{4.44}$$

Moreover,

$$s,t\in A\in\mathcal{A}_n \implies \varphi_{j_n(A)}(s,t) \le 2^{n+2}. \tag{4.45}$$

The partitioning procedure is the content of the next Lemma.

Lemma 7. *Consider $A \subset F$. There exists a partition $\mathcal{A}$ of A such that $|\mathcal{A}| \leq N_n$ and for each $B \in \mathcal{A}$*

$$s, t \in B \implies \varphi_{\tilde{j}_n^\mu(t)}(s,t) \leq 2^{n+4}. \tag{4.46}$$

Proof. Consider $U \subset T$ such that $\forall s,t \in U$, $\varphi_{\tilde{j}_n^\mu(t)}(s,t) > 2^{n+2}$. By (4.40) balls $B_{\tilde{j}_n^\mu(t)}(t, 2^n)$ are disjoint for each $t \in U$. By (4.42) it follows that $|U| \leq N_n$. Take U with maximal cardinality. Then $A \subset \bigcup_{t \in U} B_{\tilde{j}_n^\mu(t)}(t, 2^{n+2})$ and each of these balls satisfy (4.46). If we list elements of U denoting them by $t_1, \ldots, t_M$, $M \leq N_n$, then the partition $\mathcal{A}$ consists of sets

$$D_1 = A \cap B_{\tilde{j}_n^\mu(t)}(t_1, 2^{n+2}),$$

$$D_2 = (A \backslash D_1) \cap B_{\tilde{j}_n^\mu(t)}(t, 2^{n+2}),$$

$$\vdots$$

$$D_M = (A \backslash \bigcup_{k \leq M-1} D_k) \cap B_{\tilde{j}_n^\mu(t)}(t_M, 2^{n+2}).$$

Since each element of this partition is contained in the ball of radius 2^{n+2} (4.46) follows.

$\square$

Proof of Theorem 26. We proceed by induction. Set $\mathcal{A}_0 = \mathcal{A}_1 = \mathcal{A}_2 = \{T\}$ and for $n \leq 2$, $j_n(T) = j_0$. For $A \in \mathcal{A}_n$, $n \geq 2$, there exists an integer, which we denote by $j_n(A)$, such that

$$t \in A \implies \tilde{j}_{n-2}^\mu(t) = j_n(A). \tag{4.47}$$

Suppose we have constructed $\mathcal{A}_n$. By (4.47) and (4.41) for $t \in A \in \mathcal{A}_n$ we have $\tilde{j}_{n-1}^\mu \in \{j_n(A), j_n(A)+1\}$. Set

$$A_0 = \{t \in A : \tilde{j}_{n-1}^\mu(t) = j_n(A)\} \text{ and } A_1 = \{t \in A : \tilde{j}_{n-1}^\mu(t) = j_n(A)+1\}.$$

Now, we apply Lemma 7 for $n-1$ rather than n to get partitions of A_0 and A_1 into at most N_{n-1} elements, so that we obtain at most $2N_{n-1} \leq N_n$ sets. For the element B of A_0 we put $j_{n+1}(B) = j_n(A)$ and for the element B of A_1 we put $j_{n+1}(B) = j_n(A)+1$. Apply this procesdure to each set $A \in \mathcal{A}_n$ to get partitition $\mathcal{A}_{n+1}$ which is obviously nested. Clearly $|\mathcal{A}_{n+1}| \leq N_n^2 \leq N_{n+1}$. Condition (4.47) holds for $n+1$ by the construction as well as (4.45). Since for $n \leq 2$, $j_n(A_n(t)) = j_0$ and for $n \geq 3$, $j_n(A_n(t)) = \tilde{j}_{n-2}^\mu(t)$ also (4.44) follows. $\square$

The last step is to level up the partition to the whole set T and to formulate the main lower bound. Before the formal statement let us describe the idea for building the partition elements. We will follow the intuition that the partition of large finite subset of T should not vary too much from the partition of T itself. To formalize this concept we will consider ascending sequence of finite subsets $F_N \subset T$ and either assign the element of T to already existing partition element of F_N or let it define a new partition element. Alongside, we will need to guarantee that integers j defined for each partition of finite subset converge. This will be done by the procedure of choosing consecutively appropriate subsequences of the indices of F_N.

Theorem 27. *Assume that T is countable. Then, there exists an admissible sequence $(\mathcal{A}_n)_{n\geq 0}$ of partitions of T and for $A \in \mathcal{A}_n$ an integer $j_n(A)$ such that the following holds*

$$\forall t \in T, \ \sum_{n\geq 0} 2^n r^{-j_n(A_n(t))} \leq L\mathbf{E}\sup_{t\in T} X_t, \tag{4.48}$$

$$A \in \mathcal{A}_n, \ C \in \mathcal{A}_{n-1}, \ A \subset C \implies j_{n-1}(C) \leq j_n(A) \leq j_{n-1}(C) + 1 \tag{4.49}$$

and

$$s, t \in A \in \mathcal{A}_n \implies \varphi_{j_n(A)}(s,t) \leq 2^{n+2}. \tag{4.50}$$

Proof. Assume first that T is finite. Then, the result follows from Theorem 26 combined with (4.43).

Now, let T be countable so that $T = \bigcup_{N\geq 1} F_N$, where $(F_N)_{N\geq 1}$ is ascending sequence of finite subsets of T. We can enumerate elements of T i.e. $T = \{t_1, t_2, \dots\}$ and put $F_N = \{t_1, t_2, \dots, t_N\}$. Our aim is to construct the admissible sequence of partitions $(\mathcal{A}_n)_{n\geq 0}$ of T and verify conditions (4.48), (4.49) and (4.50). This will end the proof. Certainly, $\mathcal{A}_0 = \{T\}$. The approach is based on the analysis of partitions $(\mathcal{A}_{n,N})_{n\geq 0}$ of F_N and use them for defining the limiting partitions of T. Recall that $A_{n,N}(t_i)$ denotes the element of n-th partition of F_N that contains t_i and $j_n(A_{n,N}(t_i))$ is the associated integer. To simplify the notation we put $j_n(A_{n,N}(t_i)) = j_{n,N}(t_i)$.

For t_1 we obtain a sequence of sets $(A_{n,N}(t_1))_{N\geq 1}$ and the sequence of integers $(j_{n,N}(t_1))_{N\geq 1}$. Note that by (4.41)

$$j_0 \leq j_{n,N}(t_1) \leq j_0 + n. \tag{4.51}$$

The main step in the construction is to describe appropriate limiting procedure allowing to define $(A_n(t_1))_{n\geq 0}$ and $(j_n((t_1))_{n\geq 0}$. It relies highly on the boundedness of $j_{n,N}(t_i)$ since we can expect a stabilization on certain infinite subsequences. For the sake of simplicity all the sequences used in the proof will be identified with certain subsets of positive integers. Then the fact that M is a subsequence of M' is equivalent to $M \subset M'$. The first task is to define subsequences $N_n(t_1)$ for $n \geq 0$. We know that for $n \leq 2$ and each N, $j_{n,N}(t_1) = j_0$, so we put $N_0(t_1) = N_1(t_1) = N_2(t_1) = \{1, 2, \dots\}$. In this case, $\mathcal{A}_n = \{T\}$. Next we choose a subsequence $N_3(t_1)$ of $N_2(t_1)$ such that $(j_{3,N}(t_1))_{N\in N_2(t_1)}$ converges to some limit (which is guaranteed by (4.51)) and we denote its' limit by $j_3(t_1)$. We proceed in this way to obtain nested sequences $(N_n(t_1))_{n\geq 0}$ and finally by the diagonal procedure we can select a sequence $N(t_1)$ such that for each n the sequence $(j_{n,N}(t_1))_{N\in N(t_1)}$ converges to $j_n(t_1)$.

Let $N_0(t_2) = N(t_1)$. For $n > 0$ consider a set $A_n(t_1)$. We describe when t_2

should belong to $A_n(t_1)$, namely

$$t_2 \in A_n(t_1) \iff \exists N_n(t_2) \subset N_{n-1}(t_2) \text{ such that } \forall N \in N_n(t_2),\ t_2 \in A_{n,N}(t_1). \tag{4.52}$$

There are two possible cases. Firstly, if the condition (4.52) is satisfied for each n we obtain $N_0(t_2) \supset N_1(t_2) \supset \ldots$. Due to the property that $j_{n,N}(t_1)$ converges to $j_n(t_1)$ for $N \in N(t_1)$ we derive that $j_{n,N}(t_2) = j_{n,N}(t_1)$ converges to the same limit for $N \in N_n(t_2)$. Secondly, the condition (4.52) might not be satisfied for some $n = \bar{n} \geq 1$, which means that $t_2 \in A_{n,N}(t_1)$ for finitely many $N \in N_{n-1}(t_2)$. In this case we start a new element of partition (which will be denoted by $A_n(t_2)$) and consrtuct the subsqence $N_n(t_2)$ in order to stabillize $j_{n,N}(t_2)$ so that for $N \in N_n(t_2)$ it converges to $j_n(t_2)$, which is guaranteed by (4.51). We continue in this way obtaining new parition elements $(A_n(t_2))_{n\geq\tilde{n}}$, subsequences $N_n(t_2)$ and limits $j_n(t_2)$ for all $n \geq \tilde{n}$. Hence, again we obtain nested family of subsequences $N_0(t_2) \supset N_1(t_2) \supset \ldots$. Finally, in both cases we can define $N(t_2)$ by the diagonal method.

The procedure described for t_1 and t_2 should help getting the intuition and clarify the general construction which we provide now. The argument goes by induction. Consider t_i for $i \geq 1$ and suppose we have dealt with $t_1, \ldots, t_{i-1}$. We aim to define an inductive procedure for constructing $N_n(t_i)$ and deciding whether t_i belongs to already existing partition element or starts a new one. Put $N_0(t_i) = N(t_{i-1})$. Note that $N(t_{i-1})$ is already provided. For $n > 0$, consider $N_{n-1}(t_i)$ and suppose that there exists a subsequence M of $N_{n-1}(t_i)$ such that

$$t_i \in \bigcup_{j=1}^{i-1} A_{n,N}(t_j) \ \ \forall \ \ N \in M.$$

If so, then we select the smallest j for which there exists $N_n(t_i) \subset N_{n-1}(t_i)$ with the propety that $t_i \in A_{n,N}(t_j)$ for $N \in N_n(t_j)$. In this way we obtain an infinite subsequence $N_n(t_i) \subset N_{n-1}(t_i)$ and put t_i into the partition element $A_n(t_j)$. Similarly as we have argued for t_2, we have that $j_{n,N}(t_i)$ converges to $j_n(t_j)$. Secondly, it may happen that t_i belongs to $\bigcup_{j=1}^{i-1} A_{n,N}(t_j)$ only for finitely many $N \in N_{n-1}(t_i)$. In this case we start a new parititon element, $A_n(t_i)$. When choosing $N_n(t_i) \subset N_{n-1}(t_i)$ we care only for the stabilization of $j_{n,N}(t_i)$ namely we require that $j_{n,N}(t_i)$ converges to a limit which we denote by $j_n(t_i)$. Once again, (4.51) implies the existence of this limit. Following the above scheme we decide whether t_i starts a new partition element for $\mathcal{A}_n$ or not, construct $j_n(t_i)$ and $N_0(t_i) \supset N_1(t_i) \ldots$. We complete the procedure by choosing $N(t_i)$ from the family $N_0(t_i) \supset N_1(t_i) \ldots$ by the diagonal method. Note that it does not affect the convergence of $j_{n,N}(t_i)$ to $j_n(t_i)$ for $N \in N(t_i)$.

Now we check that the defined sequence of partitions is admissible. Namely, that the sequence of partitions $\mathcal{A}_n = \{A_n(t_i) : i \in I\}$, where I is the index set gathering those points in T which start a partition element, is nested and satisfies $|\mathcal{A}_n| \leq 2^{2^n}$. We have $A_n(t_i) = \{t_i\} \cup \{t_j \in T :\ j > i,\ t_j \in A_{n,N}(t_i)\ \forall\ N \in N_n(t_j)\}$. The crucial property of the constructed partition is following. Fix m and consider $F_m = \{t_1, ..., t_m\}$. For any n there exists a constant $K_{n,m}$ large

enough such that for $N > K_{n,m}$ and $N \in N(t_m)$ we have $A_{n,N}(t_i) = A_n(t_i) \cap F_m$. Hence, $|\mathcal{A}_n \cap F_m| = |\mathcal{A}_{n,N}| \leq 2^{2^n}$ and since m is arbitrary we conclude that $|\mathcal{A}_n| \leq 2^{2^n}$. By the same reason we argue that $A_{n+1}(t_i) \subset A_n(t_i)$, because $(A_{n+1}(t_i) \cap F_m) = A_{n+1,N}(t_i) \subset A_{n,N}(t_i) = (A_n(t_i) \cap F_m)$.
Finally, we verify (4.48), (4.49) and (4.50). They are all straightforward consequence of the fact that for $N > K(n,m)$ and $N \in N_n(t_m)$ we have $j_{n,N}(t_i) = j_n(t_i)$.

□

Remark 9. The proof easily extends to separable T, that is to such T which has a countable subset which is dense in T with respect to some metric and also each function $\varphi_j(s,t)$ is continuous in this metric. Let $\mathcal{A}_n$ be the admissible sequence of partitions of this subset given by Theorem 27. Let $\bar{A}$ be the closure of the set A, so by continuity of $\varphi_{j(A)}(s,t)$, $A \in \mathcal{A}_n$, we have for each $s,t \in \bar{A}$, $\varphi_{j(A)}(s,t) \leq 2^{n+2}$. Sets $\bar{A}$ cover T, but they do not have to be disjoint. We construct an admissible sequence $\mathcal{B}_n$ of T such that for each $B \in \mathcal{B}_n$ there exists $A \in \mathcal{A}_n$ such that $B \subset \bar{A}$. It is possible because $\mathcal{A}_n$ is admissible so in particular $\bar{A} = \bigcup\{\bar{C} : \; C \subset A, \; A \in \mathcal{A}_{n+1}\}$.

Assume with no loss in generality that $0 \in T$ which implies that (see [32, Lemma 2.2.1])

$$\mathbf{E}\sup_{t\in T}|X_t| \leq 2\mathbf{E}\sup_{t\in T}X_t. \tag{4.53}$$

We need one more ingredient for the final result.

Lemma 8. *Assume $0 \in T$. For each $t \in T$ we have*

$$\int_\Omega |t|\mathbb{1}_{\{2|t|\geq r^{-j_0(t)}\}}d\nu \leq L\mathbf{E}\sup_{t\in T}X_t. \tag{4.54}$$

Proof. By (4.1) we have $\int_\Omega |t|\mathbb{1}_{\{2|t|\geq r^{-j_0(t)}\}}d\nu = \mathbf{E}\sum_{i\geq 1}|t(Z_i)|\mathbb{1}_{\{2|t(Z_i)|\geq r^{-j_0(t)}\}}$. Define

$$N_k = \sum_{i\geq 1}\mathbb{1}_{\{r^k\leq 2|t(Z_i)|<r^{k+1}\}}$$

and observe that N_k is a Poisson random variable with mean $\nu(\{\omega : r^k \leq 2|t(\omega)| < r^{k+1}\})$. Recall that $\varphi_{j_0}(t,0) \leq 4$. Again by (4.1),

$$r^{-2j_0}\varphi_{j_0}(t,0) = \mathbf{E}\sum_{i\geq 1}\min(|t(Z_i)|^2, r^{-2j_0}).$$

Hence,

$$\frac{r^{-2j_0}}{4}\mathbf{E}\sum_{k\geq -j_0}N_k \leq \mathbf{E}\min(|t(Z_i)|^2, r^{-2j_0}) \leq r^{-2j_0}.$$

It means that for each $k \geq -j_0$, $\mathbf{E}N_k \leq 1$ and since for $\lambda \leq 1$ it holds that $\lambda \leq e(1-e^{-\lambda})$ we have

$$\mathbf{E}N_k \leq e\mathbf{P}(N_k > 0).$$

The above leads to

$$\mathbf{E}\sum_{i\geq 1}|t(Z_i)|\mathbb{1}_{\{2|t(Z_i)|\geq r^{-j_0(t)}\}} \leq \mathbf{E}\sum_{k\geq -j_0}\frac{r^{k+1}}{2}N_k \leq 2e\mathbf{E}\sum_{k\geq -j_0} r^{k+1}\mathbb{1}_{\{N_k>0\}}.$$

Observe that

$$\begin{aligned}\mathbf{E}\sum_{k\geq -j_0}(r^{2(k+1)}\mathbb{1}_{\{N_k>0\}})^{\frac{1}{2}} &\leq \mathbf{E}\left(\sum_{k\geq -j_0} r^{2(k+1)}N_k\right)^{\frac{1}{2}} \leq L\mathbf{E}\left(\sum_{i\geq 1}|t(Z_i)|^2\right)^{\frac{1}{2}}\\ &\leq L\sup_{t\in T}\mathbf{E}\left(\sum_{i\geq 1}|t(Z_i)|^2\right)^{\frac{1}{2}} \leq L\sup_{t\in T}\mathbf{E}|\sum_{i\geq 1}\varepsilon_i t(Z_i)|\\ &\leq L\mathbf{E}\sup_{t\in T}|\sum_{i\geq 1}\varepsilon_i t(Z_i)|.\end{aligned}$$

In particular, $\mathbf{E}\sum_{k\geq -j_0}(r^{2(k+1)}\mathbb{1}_{\{N_k>0\}})^{\frac{1}{2}} < \infty$, so there exists maximal k for which $N_k > 0$. Thus, we can deduce that the sum is controlled by the last term and we can write

$$\mathbf{E}\sum_{k\geq -j_0} r^{k+1}\mathbb{1}_{\{N_k>0\}} \leq \frac{r}{r-1}\mathbf{E}(\sum_{k\geq -j_0} r^{2(k+1)}\mathbb{1}_{\{N_k>0\}})^{\frac{1}{2}}.$$

The result follows by (4.53).

□

Theorem 27 and Theorem 21 lead to the main result of this dissertation, which is the following.

Theorem 28. *Consider a countable set T of measurable functions on Ω as in Definition 3 and assume $0 \in T$. Then we can write $T \subset T_1 + T_2$ in such a manner that*

$$\gamma_2(T_1, d_2) + \gamma_1(T_1, d_\infty) \leq L\mathbf{E}\sup_{t\in T} X_t \tag{4.55}$$

and

$$\mathbf{E}\sup_{t\in T_2}|X|_t \leq L\mathbf{E}\sup_{t\in T} X_t. \tag{4.56}$$

Proof. Let $s(T) = \mathbf{E}\sup_{t\in T} X_t$. Consider T', T'', T'' as in Theorem 21. Put $T_1 = T'$ and $T_2 = T'' + T'''$. By (4.48) we have $\gamma_2(T_1, d_2) \leq Ls(T)$ and $\gamma_1(T_1, d_\infty) \leq Ls(T)$ using (4.8) and (4.9) respectively, so we showed (4.55). Now, by replacing T_2 by $T_2 \cap (T - T_1)$, it follows from (4.13) that

$$\begin{aligned}\mathbf{E}\sup_{t\in T_2}|X|_t &\leq \sup_{t\in T_2}\int_\Omega |t(\omega)|\nu(d\omega) + 2\mathbf{E}\sup_{t\in T_2}|X_t|\\ &\leq \sup_{t\in T'''}\int_\Omega |t(\omega)|\nu(d\omega) + L(\mathbf{E}\sup_{t\in T}|X_t| + \mathbf{E}\sup_{t\in T_1}|X_t|),\end{aligned}$$

where we used that for $t \in T''$ it is straightforward from (4.10) and (4.48) that $\|t\|_1 \leq Ls(T)$. Now, notice that (4.3) and (4.53) together with (4.55) give that $\mathbf{E}\sup_{t\in T_1}|X_t| \leq Ls(T)$. Hence, the last piece we need to show is that for each

$t \in T'''$, $\|t\|_1 \le Ls(T)$. By the construction (4.11) we know that for each $t \in T'''$ there is $s \in T$ such that

$$|t| \le 5|s|\mathbb{1}_{\{2|s| \ge r^{-j_0}\}}.$$

Combining this with Lemma 8 we finally get that

$$\sup_{t \in T'''} \int_\Omega |t(\omega)|\nu(d\omega) \le Ls(T).$$

□

4.5 Comments

4.5.1 Approximation net for T

There is another proof of Theorem 28 which is based on a direct construction of approximation net of T rather than admissible sequence of partitions. Namely, one can show the following form of Theorem 21. For the sake of clarity in presenting the result in what follows the j_n^μ of (4.38) will be denoted by $j_n(t)$. We will assume with no loss in generality that $0 \in T$ which implies that $\mathbf{E}\sup_{t\in T}|X_t| \le 2\mathbf{E}\sup_{t\in T} X_t$. Let us point out straight ahead that the approach with partitions has a huge advantage over the one presented below when it comes to extending the result beyond the finite T case. Nevertheless, it seems instructive to see the following construction.

Theorem 29. *Consider a compact set T of measurable functions on Ω and assume that $0 \in T$. Let μ be the probability measure on T and $j_n(t) \in \mathbb{Z}$ both as in (4.36). Then we can write $T \subset T_1 + T_2 + T_3$ with $0 \in T_1$, where for some parameter $u \ge 0$*

$$\gamma_2(T_1, d_2) \le L\sqrt{u}\sup_{t\in T}\sum_{n\ge 0} 2^n r^{-j_n(t)} \tag{4.57}$$

$$\gamma_1(T_1, d_\infty) \le L\sup_{t\in T}\sum_{n\ge 0} 2^n r^{-j_n(t)} \tag{4.58}$$

$$\forall t \in T_2, \quad \|t\|_1 \le Lu\sup_{t\in T}\sum_{n\ge 0} 2^n r^{-j_n(t)} \tag{4.59}$$

and

$$\forall t \in T_3 \quad \exists s \in T, \quad |t| \le 4|s|\mathbb{1}_{\{2|s|\ge r^{-j_0(t)}\}}, \tag{4.60}$$

where $\|\cdot\|_1$ is the L^1 norm with respect to the measure ν.

Before we proceed to the proof of Theorem 29 let us record the essential tool which enables to provide a lower bounds in terms of γ-numbers.

Lemma 9. *[32, Theorem 2.3.1] Consider a metric space (T, d), and for $n \ge 0$, consider subsets T^n of T with $|T_0| = 1$ and for $n \ge 1$ suppose there is $\tau \ge 0$ such $|T_n| \le N_{n+\tau}$. Consider a number A and let*

$$U = \{t \in T : \sum_{n\ge 0} 2^{n/\alpha} d(t, T^n) \le A\}.$$

Then $\gamma_\alpha(U,d) \le LA$.

The constant in the assertion depends on α and τ only. Lemma 9 clarifies the task of proving Theorem 29. Together with the set T_1 we have to provide approximating sets T_1^n such that for each element $t \in T_1$ we have

$$\sum_{n=0}^{\infty} 2^{n/2} d_2(t, T_n^1) \le \sum_{n\ge 0} 2^n r^{-j_n(t)}$$

as well as

$$\sum_{n\ge 0} 2^n d_\infty(t, T_n^1) \le \sum_{n=0}^{\infty} 2^n r^{-j_n(t)}.$$

T_2 and T_3 will be obtained by splitting $T - T_1$ with respect to the size of its elements and then we will need to verify (4.59) and (4.60).
The construction is as follows. It should be compared with the construction provided for the majorising measures (see [27, Theorem 2.2]). First, observe (using $(a+b)^2 \le 2(a^2+b^2)$) that we have the following form of the triangle inequality. For $s,t,x \in T$

$$\varphi_j(s,t) \le 2(\varphi_j(s,x) + \varphi_j(x,t)). \tag{4.61}$$

Fix $n \ge 0$. We will define an approximating nets $\tilde{T}_n = \{t_0, \dots, t_M\}$ and we will show that $M \le N_{n+\tau}$ for some $\tau \ge 0$. Let $t_0 = 0$. We choose t_1 so that it maximizes $j_n(t)$ over the whole set T. Then we define set

$$D_1 = \{s \in T : \varphi_{j_n(s)-1}(t_1, s) \le 2^{n+2}\}$$

and for $k \ge 2$ we choose $t_k \in T \setminus \bigcup_{l=1}^{k-1} D_l$ which maximizes $j_n(t)$ over the set $T \setminus \bigcup_{l=1}^{k-1} D_l$. Then, we define

$$D_k = \{s \in T : \varphi_{j_n(s)-1}(t_k, s) \le 2^{n+2}\}.$$

Now, we need to argue that the procedure of choosing points t_k will cease after at most $N_{n+\tau'}$ steps for some $\tau' \ge 0$. For this, let $k' < k$ and consider balls centred at $t_{k'}, t_k$ of radius 2^n in distance $\varphi_{j_n(t_{k'})-1}$ and $\varphi_{j_n(t_k)-1}$ respectively. If we show that they are disjoint, then the claim will follow since j_n's are defined so that $\mu(B_{j_n(t)-1}) \ge 1/2N_n$ implying that $M \le 2N_n \le 2^{2^{n+1}}$. Suppose that there is $x \in B_{j_n(t_{k'})-1}(t_{k'}, 2^n) \cap B_{j_n(t_k)-1}(t_k, 2^n)$. Since $t_k \notin D_{k'}$ we have that

$$\varphi_{j_n(t_k)-1}(t_k, t_{k'}) > 2^{n+2}. \tag{4.62}$$

On the other hand, we have $\varphi_{j_n(t_{k'})-1}(t_{k'}, x) \le 2^n$ and $\varphi_{j_n(t_k)-1}(t_k, x) \le 2^n$. Observe that the way of choosing points t_k implies that $j_n(t_{k'}) > j_n(t_k)$, so by (4.6) we get

$$\varphi_{j_n(t_k)-1}(t_k, x) \le \varphi_{j_n(t_{k'})-1}(t_k, x),$$

which together with the triangle inequality (4.61) implies that

$$\varphi_{j_n(t_{k'})-1}(t_k, t_{k'}) \le 2(\varphi_{j_n(t_{k'})-1}(t_k, x) + \varphi_{j_n(t_{k'})-1}(t_{k'}, x)) \le 2^{n+2},$$

which contradicts (4.62) so $B_{j_n(t_{k'})-1}(t_{k'}, 2^n) \cap B_{j_n(t_k)-1}(t_k, 2^n) = \emptyset$. This proves the claim that $|\tilde{T}_n| \leq N_{n+1}$. Now, at each level $n \geq 0$ we define

$$\tilde{\pi}_n(t) = t_l, \text{where } \ l = \min\{i \geq 1 : t \in D_i\}. \tag{4.63}$$

and

$$\pi_n(t) = \begin{cases} \tilde{\pi}_n(t) & \text{if } j_n(t) > j_{n-1}(t) \\ \pi_{n-1}(t) & \text{if } j_n(t) = j_{n-1}(t), \end{cases}$$

where $\pi_0(t) = 0$. In order to prove Theorem 29 we will follow closely the main steps of Theorem 6.2.6 in [33] (cf. [32, Theorem 5.2.6]). We have to define $t^1(\omega), t^2(\omega)$ and $t^3(\omega)$ such that

$$t(\omega) = t^1(\omega) + t^2(\omega) + t^3(\omega).$$

For $t \in T$, $\omega \in \Omega$ define

$$m(t,\omega) = \inf\{n : |\pi_{n+1}(t)(\omega) - \pi_n(t)(\omega)| > r^{-j_n(t)+1}\} \tag{4.64}$$

if the set on the right is not empty and put $m(t,\omega) = \infty$ otherwise. Let

$$\Omega_n(t) = \{\omega \in \Omega : m(t,\omega) \geq n\}. \tag{4.65}$$

Lemma 10. *If $n < m(t,\omega)$, then*

$$\sum_{n \leq m < m(t,\omega)} |\pi_{m+1}(t)(\omega) - \pi_m(t)(\omega)| \leq 2r^{-j_n(t)+1}. \tag{4.66}$$

Proof. For each $m < m(t,\omega)$ we have

$$|\pi_{m+1}(t)(\omega) - \pi_m(t)(\omega)| \leq r^{-j_m(t)+1},$$

so

$$\sum_{n \leq m < m(t,\omega)} |\pi_{m+1}(t)(\omega) - \pi_m(t)(\omega)| \leq \sum_{n \leq m} r^{-j_m(t)+1} \leq r^{-j_n(t)+1} \frac{1}{1 - 1/r},$$

where we used the fact that $j_m(t)$ is non-decreasing sequence of positive integers and since $r \geq 4$ the result follows.

□

Let us present the decomposition of $t \in T$. Define $t^1(\omega) = \pi_{m(t,\omega)}(t)(\omega)$ and in the case when $m(t,\omega) = \infty$ we define $t^1(\omega) = \lim_{n\to\infty} \pi_n(t)(\omega)$ as the existence of the limit is guaranteed from (4.66) and since we can assume that

$$\sup_{t \in T} \sum_{n \geq 0} 2^n r^{-j_n(t)} < \infty. \tag{4.67}$$

Notice that since $\pi_0(t) = 0$ from (4.66) applied to $n = 0$ it follows that for each ω

$$|t^1(\omega)| \leq 2r^{-j_0}. \tag{4.68}$$

Recall that by (4.26) j_0 is independent of t. Intuitively, the definition of t_2 should be simply $t - t^1$, however we will need some control over its size. Therefore, define

$$\Xi(t) = \{\omega \in \Omega : |t(\omega)| \leq r^{-j_0(t)}/2\}$$

and then $t^2(\omega) = (t(\omega) - t^1(\omega))\mathbb{1}_{\Xi(t)}$ and $t^3(\omega) = (t(\omega) - t^1(\omega))\mathbb{1}_{\Xi(t)^c}$. We define

$$T_1 = \{t^1 : t \in T\} \ ; \ T_2 = \{t^2 : t \in T\} \ ; \ T_3 = \{t^3 : t \in T\}.$$

It is easy to see that (4.60) holds true. Indeed, for any $t \in T$ we have

$$|(t(\omega) - t^1(\omega))\mathbb{1}_{\Xi(t)^c}| \leq (|t(\omega)| + |t^1(\omega)|)\mathbb{1}_{\Xi(t)^c} \leq (|t(\omega)| + 2r^{-j_0})\mathbb{1}_{\Xi(t)^c} \leq 4|t(\omega)|\mathbb{1}_{\Xi(t)^c},$$

since on the set $\Xi(t)^c$, $r^{-j_0(t)} < 2|t(\omega)|$.
We proceed to L^∞ and L^2 estimates for the set T_1 i.e. (4.58) and (4.57) respectively. The proof of (4.58) is the subject of the next two lemmas.

Lemma 11. *Let $t^1_n(\omega) = \pi_{m(t,\omega)\wedge n}(t)(\omega)$ and $T^n_1 = \{t^1_n : t \in T\}$. Then, $|T^n_1| \leq N_{n+\tau}$ for $\tau \geq 0$.*

Proof. It is clear from the construction (4.63) that $\pi_{m(t,\omega)\wedge n}(t)(\omega)$ is one of the points t_l, therefore $T^n_1 \subset \bigcup_{k\leq n} \tilde{T}_k$, so $|T^n_1| \leq \sum_{k\leq n} 2^{2^{k+1}} \leq 2^{2^{n+2}}$.

□

Lemma 12. *We have $d_\infty(t^1, T^n_1) \leq 2r^{-j_n(t)+1}$.*

Proof. For $n \geq m(t,\omega)$, $|t^1(\omega) - t^1_n(\omega)| = 0$. For $n < m(t,\omega)$,

$$\begin{aligned} |t^1(\omega) - t^1_n(\omega)| &= |\sum_{n\leq m<m(t,\omega)} (\pi_{m+1}(t)(\omega) - \pi_m(t)(\omega))| \\ &\leq \sum_{n\leq m<m(t,\omega)} r^{-j_m(t)+1} \leq 2r^{-j_n(t)+1} \end{aligned}$$

by the same argument as in Lemma 10. So, $\|t^1(\omega) - t^1_n(\omega)\|_\infty \leq 2r^{-j_n(t)+1}$ and therefore $d_\infty(t^1, T^n_1) \leq 2r^{-j_n(t)+1}$.

□

Corollary 8. *The bound (4.58) holds true.*

Proof. Conditions of Lemma 9 are verified in Lemmas 11 and 12, therefore we apply it with $\alpha = 1$.

□

Now, we present the control over T_1 in L^2 norm i.e. (4.57)

Lemma 13. *For $(t^1_n)_{n\geq 0}$ as in Lemma 11 and some number $u \geq 0$ we have*

$$d_2(t^1_n, t^1_{n+1}) \leq \sqrt{u} r^{-j_n(t)} 2^{n/2}. \tag{4.69}$$

Proof. Let $p(n,t) = \inf\{p \geq 0 : j_n(t) = j_p(t)\}$. Recall that the construction of $\pi_n(t)$ implies that

$$\varphi_{j_{p(n,t)}(t)-1}(\pi_{p(n,t)}(t), t) = \int r^{2j_{p(n,t)}(t)-2}|\pi_{p(n,t)}(t) - t|^2 \wedge 1\mu(d\omega) \leq 2^{p(n,t)+2}.$$

Recall that by definition of $m(t,\omega)$ and Ω_n we have

$$|t^1_{n+1} - t^1_n| \le |\pi_{n+1}(t) - \pi_n(t)| \mathbb{1}_{\Omega_n \cap \{|\pi_{n+1}(t)-\pi_n(t)| \le r^{-j_n(t)+1}\}},$$

so

$$\begin{aligned}
d_2^2(t^1_{n+1}, t^1_n) &= \int |t^1_{n+1}(\omega) - t^1_n(\omega)|^2 \nu(d\omega) \\
&\le \int_{\Omega_n} |\pi_{n+1}(t) - \pi_n(t)|^2 \wedge r^{-2j_n(t)+2} \nu(d\omega) \\
&\le r^{-2j_n(t)+2} \varphi_{j_n(t)-1}(\pi_{n+1}(t), \pi_n(t)) \\
&\le r^{-2j_n(t)+2} \varphi_{j_{p(n,t)}-1}(\tilde{\pi}_{n+1}(t), \tilde{\pi}_{p(n,t)}(t)) \\
&\le 2r^{-2j_n(t)+2} (\varphi_{j_{p(n,t)}-1}(\tilde{\pi}_{n+1}(t), t) + \varphi_{j_{p(n,t)}-1}(\tilde{\pi}_{p(n,t)}(t), t)) \\
&\le 2r^{-2j_n(t)+2} (\varphi_{j_{n+1}(t)-1}(\tilde{\pi}_{n+1}(t), t) + \varphi_{j_{p(n,t)}-1}(\tilde{\pi}_{p(n,t)}(t), t)) \\
&\le 24 r^{-2j_n(t)+2} 2^n.
\end{aligned}$$

□

Corollary 9. *The bound (4.57) holds true.*

Proof. Apply again Lemma 9. First, by the assumption (4.67) and (4.69) we notice that $(t^1_n)_{n\ge 0}$ is a Cauchy sequence in L^2. Notice that for all $\omega \in \Omega$, $\lim_{n\to\infty} t^1_n(\omega) = t^1(\omega)$, so

$$\lim_{q\to\infty} \|t^1_q - t^1_n\|_2 = \|t^1 - t^1_n\|_2.$$

Hence for any element t^1 of T_1,

$$d_2(t^1, U_n) \le \|t^1 - t^1_n\|_2 = \lim_{q\to\infty} \|t^1_q - t^1_n\|_2 \le \sum_{m\ge n} \|t^1_{m+1} - t^1_m\|_2 \le \sqrt{u} \sum_{m\ge n} 2^{m/2} r^{-j_m(t)}.$$

Finally,

$$\begin{aligned}
\sum_{n\ge 0} 2^{n/2} d_2(t^1, U_n) \le \sqrt{u} \sum_{n\ge 0} 2^{n/2} \sum_{m\ge n} 2^{m/2} r^{-j_m(t)} &= \sqrt{u} \sum_{m\ge 0} 2^{m/2} r^{-j_m(t)} \sum_{n\le m} 2^{n/2} \\
&\le L\sqrt{u} \sum_{m\ge 0} 2^{m/2} r^{-j_m(t)}.
\end{aligned}$$

Inequality (4.57) then follows by Lemma 9 applied for $\alpha = 2$.

□

The last part concerns the L^1 control of t^2 i.e. (4.59). Recall, that by the construction we have for each $n \ge 0$

$$\varphi_{j_n(t)-1}(\pi_n(t), t) = \int_\Omega r^{2j_n(t)-2} |\pi_n(t)(\omega) - t(\omega)|^2 \wedge 1 \nu(d\omega) \le 2^{n+2}.$$

Similarly to $m(t,\omega)$ of (4.64) we define for each $t \in T$

$$r(t,\omega) = \inf\{n \geq 0 : |\pi_{n+1}(t)(\omega) - t(\omega)| \geq \frac{1}{2} r^{-j_{n+1}(t)+1}\}, \tag{4.70}$$

and $r(t,\omega) = \infty$ if the above set is empty. Notice that for $\omega \in \Xi(t)$ and $n < r(t,\omega)$ it holds that

$$|\pi_{n+1}(t)(\omega) - \pi_n(t)(\omega)| \leq |\pi_{n+1}(t)(\omega) - t(\omega)| + |\pi_{n+1}(t)(\omega) - \pi_n(t)(\omega)| \leq r^{-j_n(t)+1},$$

since $\pi_0(t) = 0$ and $j_n(t)$ is increasing. This means that $r(t,\omega) \leq m(t,\omega)$. Now, put

$$t_n^2(\omega) = (t - t^1)\mathbb{1}_{\{r(t,\omega)=n\}\cap\Xi(t)}(\omega).$$

The proof of the next result is exactly the same as of [33, Lemma 6.6.6] (cf. [32, proof of Theorem 5.2.6]). We repeat it for the sake of completenes.

Lemma 14. *We have that*

$$t^2 = \sum_{n \geq 0} t_n^2 \tag{4.71}$$

and

$$\|t_2^n\|_1 \leq 3r^{-j_n(t)+1}\nu(\{\omega \in \Omega : r(t,\omega) = n\} \cap \Xi(t)) \tag{4.72}$$

Proof. Fix $\omega \in \Xi(t)$. Recall that $t^1(\omega) = \pi_{m(t,\omega)}(t)(\omega)$ if $m(t,\omega) < \infty$ and $\lim_{n\to\infty} \pi_n(t)(\omega)$ otherwise. We use the fact that $r(t,\omega) \leq m(t,\omega)$. First, suppose that $r(t,\omega) < \infty$, then trivially $t - t^1 = (t - t^1)\mathbb{1}_{\{r(t,\omega)=n\}}$. If $r(t,\omega) = \infty$, then $m(t,\omega) = \infty$ and $|\pi_{n+1}(t) - t| \leq r^{-j_{n+1}(t)+1}$ for each n, so since $\lim_{n\to\infty} j_n(t) = \infty$, we have $t^1(\omega) = \lim_{n\to\infty} \pi_n(t)(\omega) = t(\omega)$. This holds for every $\omega \in \Xi(t)$, so (4.71) follows. Now, if $r(t,\omega) = n$ then $m(t,\omega) \geq n$, so by the proof of Lemma 12 $|\pi_n(t)(\omega) - t^1(\omega)| \leq 2r^{-j_n(t)+1}$. Also, $r(t,\omega) = n$ implies that $|\pi_n(t)(\omega) - t(\omega)| \leq r^{-j_n(t)+1}/2$, therefore

$$|t(\omega) - t^1(\omega)| \leq |t(\omega) - \pi_n(t)(\omega)| + |\pi_n(t)(\omega) - t^1(\omega)| \leq 3r^{-j_n(t)+1}$$

so (4.72) follows. □

Corollary 10. *The bound (4.59) holds true.*

Proof. By (4.72) we have to bound $\nu(\{\omega \in \Omega : r(t,\omega) = n\} \cap \Xi(t))$ by 2^n. Recall, that by the construction we have for each $n \geq 0$

$$\varphi_{j_n(t)-1}(\pi_n(t), t) = \int_\Omega r^{2j_n(t)-2}|\pi_n(t)(\omega) - t(\omega)|^2 \wedge 1\nu(d\omega) \leq 2^{n+2}.$$

Now,

$$\begin{aligned} \nu(\{\omega &\in \Omega : r(t,\omega) = n\} \cap \Xi(t)) \\ &\leq 4\int_\Omega (r^{2j_{n+1}(t)-2}|\pi_{n+1}(t)(\omega) - t(\omega)| \wedge 1)\mathbb{1}_{\{r(t,\omega)=n\}\cap\Xi(t)}\nu(d\omega) \leq 2^{n+5}. \end{aligned}$$

□

4.5.2 Initial result

As mentioned before the initial proof of Theorem 28 was based on the approximation net rather than building the admissible sequence of partitions which was then more technically involved to be extended to countable T. Also the original result was a bit weaker because it contained an extra term $L\int_\Omega |t|\mathbb{1}_{\{2|t|\geq r^{-j_0(t)}\}}d\nu$ in the bound (4.56). Only after communicating the convexity argument of Section 4.3 to M. Talagrand he suggested Lemma 8 providing a proof similar to the one of Lemma 16 from the next Chapter.

Chapter 5

Generalized Bernoulli Conjecture

As noticed by M. Talagrand the contents of Chapter 4 apply not only to infintely divisible processes. Our goal now is to extract those parts of the proof of the Decomposition Theorem for infinitely divisible processes which do not depend on the structure of such processes. Although the name of this chapter indicates the decomposition for selector processes we will prove here quite general result from which decompositions of both empirical and selector processes will follow. Let us recall the general framework. We consider $S = \mathbf{E}\sup_{t\in T}\sum_{i\le N}\varepsilon_i Z_i(t)$ and assume that $Z_i = 0$ for $i > N$ and that Z_i are independent. Each Z_i is distributed according to the law λ_i and $\nu = \sum_{i\le N}\lambda_i$. We see T as functions on the measured space $(\mathcal{F},\nu)$, $\mathcal{F} = \mathbb{R}^T$ and the distances d_2 and d_∞ are induced by $L^2(\nu)$ and $L^\infty(\nu)$ norm respectively. We will prove the following general Decomposition Theorem [33, Theorem 9.6.3].

Theorem 30. *For any independent sequence $(Z_i)_{i\le N}$ of random functions there is a decomposition $Z_i = Z_i^1 + Z_i^2$ such that*

$$\gamma_2(T,d_2) + \gamma_1(T,d_\infty) \le LS, \tag{5.1}$$

where the distances are induced by $L^2(\nu^1)$ and $L^\infty(\nu^1)$ norm respectively and $\nu^1 = \sum_{i\le N}\lambda_i^1$, λ_i^1 being the law of Z_i^1, and that

$$\mathbf{E}\sup_{t\in T}\sum_{i\le N}|Z_i^2(t)| \le LS. \tag{5.2}$$

The first pillar of the proof of Decomposition Theorem for infinitely divisible processes was Lemma 4. Remarkably, it stems out of the following simple yet critical observation. Let us define

$$\psi_{j,\omega}(s,t) = \sum_{i\ge 1} r^{2j}|Z_i(s) - Z_i(t)|^2 \wedge 1,$$

which should be compared with $\tilde{\varphi}_j(s,t)$ we defined in Chapter 4. The ω indicates the randomness of Z_i's. Also, define

$$\varphi_j(s,t) = \mathbf{E}\psi_{j,\omega}(s,t) = \int_{\mathcal{F}} r^{2j}|s(\omega) - t(\omega)|\nu(d\omega).$$

Lemma 15. *With the above notation we have* $\forall s,t \in T$ *and* $\forall j \in \mathbb{Z}$

$$\mathbf{P}(\psi_{j,\omega}(s,t) \leq \varphi_j(s,t)/4) \leq \exp(-\varphi_j(s,t)/4) \tag{5.3}$$

Proof. It is a consequence of [33, Lemma 6.5.2] (also [32, Lemma 7.4.3]). Let us repeat it here. Suppose that W_i's are independent random variables such that $0 \leq W_i \leq 1$ and that there is a constant A such that $4A \leq \sum_{i\geq 1} \mathbf{E}W_i$. We will show that

$$\mathbf{P}\left(\sum_{i\geq 1} W_i \leq A\right) \leq \exp(-A). \tag{5.4}$$

Firstly, we have $\mathbf{P}(\sum_{i\geq 1} W_i \leq A) \leq \exp(A)\mathbf{E}\exp(-\sum_{i\geq 1} W_i)$. Now, notice that for $0 \leq x \leq 1$ we have $1 - x \leq e^{-x} \leq 1 - x/2$, so we have

$$\mathbf{E}\exp(-W_i) \leq 1 - \mathbf{E}W_i/2 \leq \exp(-\mathbf{E}W_i/2)$$

and by independence of W_i

$$\mathbf{E}\exp\left(-\sum_{i\geq 1} W_i\right) \leq \exp\left(-\sum_{i\geq 1} \mathbf{E}W_i/2\right) \leq \exp(-2A),$$

which finishes the argument for (5.4). To get (5.3) we apply (5.4) with $W_i = r^{2j}|Z_i(s) - Z_i(t)|^2 \wedge 1$ and $A = \varphi_j(s,t) = \sum_{i\geq 1} \mathbf{E}W_i$. □

Remark 10. In the case of infinitely divisible processes the proof of (4.28) follows the same lines because of the fact that due to (4.2) we have

$$\begin{aligned}\mathbf{E}\exp\left(-\sum_{i\geq 1} f(Z_i)\right) &= \exp\left(\int (\exp(-f(\omega)) - 1)\nu(d\omega)\right) \\ &\leq \exp\left(-1/2\int f(\omega)\nu(d\omega)\right) \leq \exp(-2A).\end{aligned}$$

It might come as a surprise, but that's almost all we need to restate the main results of sections 4.3 and 4.4. The minor remaining component is to argue that we can assume that 0 is in T. Also we will need a version of Lemma 8 and the following Giné-Zinn inequality

$$\mathbf{E}\sup_{t\in T}\sum_{i\leq N}|t(Z_i)| \leq \sup_{t\in T}\sum_{i\leq N}\int_{\mathcal{F}}|t|d\lambda_i + 4\bar{S}, \tag{5.5}$$

where we used the notation

$$\bar{S} = \mathbf{E}\sup_{t\in T}\left|\sum_{i\leq N}\varepsilon_i t(Z_i)\right|$$

the proof of which is exactly the same as of (1.31).

So, first choose any $t_0 \in T$ and consider $Z_i'(t) = Z_i(t) - Z_i(t_0)$, so that $Z_i'(t_0) \equiv 0$

and $\mathbf{E}\sup_{t\in T}\sum_{i\geq 1}\varepsilon_i Z_i'(t) = S$. Hence, if $Z_i' = Z_i'^1 + Z_i'^2$ is the decomposition of Theorem 30, then $Z_i^1(t) = Z_i'^1(t) + Z_i(t_0)$ and $Z_i^2 = Z_i'^2$ is the required decomposition of Z_i. Thus, with no loss in generality we may assume that $Z_i(t_0) \equiv 0$ for each $i \leq N$ and therefore $t_0 = 0$ ν-a.e. So, $\bar{S} \leq 2S$.
Next, we formulate the analogue of Lemma 8. Recall the definition (4.26) of the number j_0.

Lemma 16. *For $t \in T$ we have*

$$\int_{\mathcal{F}} |t| \mathbb{1}_{\{2|t|\geq r^{-j_0}\}} d\nu \leq L\bar{S}. \tag{5.6}$$

Proof. By the definition of j_0 and the fact that $t_0 = 0$ ν-a.e. we have

$$\int r^{2j_0}|t|^2 \wedge 1 d\nu \leq 4.$$

Let $U = \{2|t| \geq r^{-j_0}\} = \{4r^{2j_0}|t|^2 \geq 1\}$ and note that for positive numbers a, b we have $4a \wedge b \leq 4(a \wedge b)$, so $\nu(U) \leq 16$, which by the definition of ν means that $\sum_{i\leq N}\lambda_i(U) \leq 16$. Define $I = \{i \leq N : \lambda_i(U) \geq 1/2\}$. Notice that since $\nu(U) \leq 16$ Markov inequality implies that $|I| \leq 32$. For $i \notin I$ define an event Ξ_i given by $Z_i \in U$ and for $j \neq i$, $Z_j \notin U$. Then,

$$\mathbf{P}(\Xi_i) = \lambda_i(U)\prod_{j\neq i}(1 - \lambda_j(U)) \geq \lambda_i(U)/L.$$

Now, conditioned on Ξ_i, Z_i is distributed according to the restriction of λ_i to U, so

$$\frac{1}{\mathbf{P}(\Xi_i)}\mathbf{E}\mathbb{1}_{\Xi_i}|t(Z_i)| = \frac{1}{\lambda_i(U)}\int_U |t| d\lambda_i.$$

Hence, since $\lambda_i(U)/\mathbf{P}(\Xi_i) \leq L$, we get

$$\int_U |t| d\lambda_i \leq L\mathbf{E}\mathbb{1}_{\Xi_i}|Z_i(t)| = L\mathbf{E}\mathbb{1}_{\Xi_i}\left|\varepsilon_i Z_i(t) + \sum_{j\neq i}\mathbf{E}_\varepsilon \varepsilon_j Z_j(t)\right| \leq L\mathbf{E}\mathbb{1}_{\Xi_i}\left|\sum_{j\leq N}\varepsilon_j Z_j\right|,$$

where in the last inequality we used Jensen's inequality by taking the expectation in ε_j's for $j \neq i$ outside the absolute value. To finish the proof we summate above inequalities for $i \notin I$, use the fact that Ξ_i are disjoint. Moreover, for $i \in I$ we simply write

$$\int_U |t| d\lambda_i \leq \int_{\mathcal{F}} |t| d\lambda_i = \mathbf{E}|Z_i(t)| \leq \mathbf{E}\left|\sum_{j\leq N}\varepsilon_j Z_j\right|,$$

which again follows from Jensen's inequality and since I is finite we just combine both of the above inequalities to get (5.6). ∎

Remark 11. The above argument can serve as another proof of Lemma 8.

We will need two more intermediate steps towards the proof of the general Decomposition Theorem. The first one will produce decomposition of the set T and the other will translate it into the decomposition of the measure ν.

Theorem 31. *Consider a set T of measurable functions on a measured space $(\mathcal{F}, \nu)$ and on $T \times T$ let $\varphi(s,t) = \int_{\mathcal{F}} r^{2j}|s(\omega) - t(\omega)|\nu(d\omega)$. There is a decomposition $T \subset T_1 + T_2$, where the set T_1 satisfies*

$$\gamma_2(T_1, d_2) \leq LS \quad \text{and} \quad \gamma_1(T_1, d_\infty) \leq LS \tag{5.7}$$

and where

$$\int_{\mathcal{F}} |t| d\nu \leq LS \tag{5.8}$$

for each $t \in T_2$.

Proof. It follows from Theorem 27 and then by Theorem 21 combined with (5.6) by defining $T_1 = T'$ and $T_2 = T'' + T'''$. □

Theorem 32. *There is a decomposition $Z_i = Z_i^1 + Z_i^2$ such that (5.1) holds true together with*

$$\forall t \in T, \quad \int_{\mathcal{F}} |t| d\nu^2 \leq LS, \tag{5.9}$$

where $\nu^2 = \sum_{i \leq N} \lambda_i^2$ and λ_i^2 is the law of Z_i^2.

Proof. We apply Theorem 31 to T seen as space of functions on $\mathcal{F}$. It means that we can write $t = t^1 + t^2$, where $t^1 \in T_1$, $t^2 \in T_2$ so that T_1 satisfies (5.7) and $\int_{\mathcal{F}} |t^2| d\nu \leq LS$. We define the decomposition of $Z_i(t)$ as follows

$$Z_i^1(t) = t^1(Z_i) \quad \text{and} \quad Z_i^2(t) = t^2(Z_i).$$

Equivalently, by looking at t as function on $\mathcal{F}$ we can rewrite the above equations as

$$t(Z_i^1) = t^1(Z_i) \quad \text{and} \quad t(Z_i^2) = t^2(Z_i).$$

Now, if λ_i^1 is a law of Z_i^1 and λ_i^2 is a law of Z_i^2, then we have

$$\int_{\mathcal{F}} |t| d\lambda_i^2 = \mathbf{E}|t(Z_i^2)| = \mathbf{E}|t^2(Z_i)| = \int_{\mathcal{F}} |t^2| d\lambda_i^2. \tag{5.10}$$

By summing (5.10) over $i \leq N$ and applying (5.8) we obtain (5.9). To get (32) we deduce the equivalence of $\gamma_2(T_1, d_2)$ and $\gamma_2(T, d_2)$, where d_2 is induced by ν^1 by exact the same reasoning as in (5.10) and then apply (5.7). Similarly, for $\gamma_2(T, d_\infty)$. □

Proof of Theorem 30. By Theorem 32 it now follows in the same way as Theorem 28 using (5.5) and (1.23). □

It is now easy to see that the following result concerning empirical processes (see [33, Theorem 5.8.3], [32, Research problem 9.1.3]) follows in the same way. The only extra tool we need is (1.25) and since we have already proved the Giné-Zinn inequality (1.31) the following result is straightforward.

Theorem 33. *Consider a class $\mathcal{F}$ of functions in $L^2(\mu)$ and assume $\mu(f) = 0$ for $f \in \mathcal{F}$. For a number N, there exists decomposition $\mathcal{F} \subset \mathcal{F}_1 + \mathcal{F}_2$ with $0 \in \mathcal{F}_1$ and such that*

$$\gamma_2(\mathcal{F}_1, d_2) \le \frac{L}{\sqrt{N}} S_N(\mathcal{F}),$$

$$\gamma_1(\mathcal{F}_1, d_\infty) \le L S_N(\mathcal{F})$$

and

$$\mathbf{E} \sup_{f \in \mathcal{F}_2} \sum_{i \le N} |f(X_i)| \le L S_N(\mathcal{F}).$$

Furthermore, the Generalized Bernoulli Conjecture [32, Conjecture 12.3.3] also follows from this approach (see [33, Theorem 9.11.1]).

Theorem 34. *For a set T of sequences there exists decomposition $T \subset T_1 + T_2$ such that*

$$\gamma_2(T_1, d_2) \le \frac{L}{\sqrt{\delta}} \delta(T),$$

$$\gamma_1(T_1, d_\infty) \le L\delta(T)$$

and

$$\mathbf{E} \sup_{t \in T_2} \sum_{i \le M} |t_i| \delta_i \le L\delta(T).$$

Proof. Let us summarize all required steps. First, we consider $Z_i(t) = \delta_i t_i$ and the distances needed for the critical property (5.3) are given by

$$\psi_{j,\omega}(s,t) = \sum_{i \le M} \delta_i r^{2j} |t_i - s_i|^2 \wedge 1 \quad \text{and} \quad \varphi_j(s,t) = \delta \sum_{i \le M} r^{2j} |t_i - s_i|^2 \wedge 1,$$

which holds because δ_i's are independent. This means that we are in the position of applying Theorem 31, which in this case reads that for $S = \mathbf{E} \sup_{t \in T} \sum_{i \le M} \varepsilon_i t_i \delta_i$ we can decompose $T \subset T_1 + T_2$ so that

$$\gamma_2(T_1, d_2) \le \frac{LS}{\sqrt{\delta}} \quad \text{and} \quad \gamma_1(T_1, d_\infty) \le LS$$

together with

$$\sup_{t \in T_2} \sum_{i \le M} |t_i| \le \frac{LS}{\delta}.$$

The remaining steps are to use Giné-Zinn inequality (1.32) and conclude with (1.22) and (1.26) just as in the proof of Theorem 28. □

Bibliography

[1] W. Beckner. “Inequalities in Fourier Analysis”. In: *Ann. of Math.* 102 (1998), pp. 159–182.

[2] W. Bednorz and R. Latała. “On the boundedness of Bernoulli processes”. In: *Ann. of Math.* 180 (2014), pp. 1167–1203.

[3] W. Bednorz and R. Martynek. “A Lévy–Ottaviani type inequality for the Bernoulli process on an interval”. In: *Stat Probab Lett* 162.C (2020).

[4] W. Bednorz and R. Martynek. “A note on infinitely divisible processes”. arXiv:2008.09876. 2020.

[5] W. Bednorz and R. Martynek. “On the contraction property of Bernoulli canonical processes”. In: *Bulletin Polish Acad. Sci. Math.* 67.2 (2019), pp. 187–209.

[6] W. Bendorz and T. Tkocz. “Stochastic Dominance and Weak Concentration for Sums of Independent Symmetric Random Vectors”. In: *IMRN* (2018), pp. 187–209.

[7] D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes. Vol. 1.* New York: Springer-Verlag, 2003.

[8] X. Fernique. “Régularité des trajectoires des fonctions aléatoires gaussiennes”. In: *École d'Été de Probabilités de Saint-Flour, IV-1974, Lecture Notes in Math.* 480. Berlin: Springer, 1975, pp. 1–96.

[9] D. J. Garling. *Inequalities: A Journey into Linear Analysis.* Cambridge: Cambridge University Press, 2007.

[10] E. Giné and J. Zinn. “Some limit theorems for empirical processes”. In: *Ann. Probab.* 12 (1984), pp. 929–989.

[11] P. Hitczenko. “Domination inequality for martingale transforms of a Rademacher sequence”. In: *Israel J. Math.* 84 (1993), pp. 161–178.

[12] P. Hitczenko and S. Kwapień. “On the Rademacher series”. In: *Probability in Banach Spaces.* Ed. by J. Hoffmann-Jørgensen et al. Vol. 9. 2050. Birkäuser Boston, Inc., 1994, pp. 31–36.

[13] S. Kwapień and W.A.Woyczyński. *Random Series and Stochastic Integrals: Single and Multiple. Probability and Its Applications.* Boston, MA: Birkäuser Boston, Inc., 1992.

[14] G. Last and M. Penrose. *Lectures on the Poisson Process.* Institute of Mathematical Statistics Textbooks. Cambridge: Cambridge Univ. Press, 2018.

[15] R. Latała. "Moments of unconditional logarithmically concave vectors". In: *Geometric Aspects of Functional Analysis, Israel Seminar, 2006-2010*. 2050. Springer, 2012, pp. 301–315.

[16] R. Latała. "Sudakov minoration principle and supremum of some processes". In: *Geom. Funct. Anal.* 7 (1997), pp. 936–953.

[17] R. Latała and T. Tkocz. "Sudakov minoration principle and supremum of some processes". In: *Electron. J. Probab.* 20 (2015).

[18] M. Ledoux and M. Talagrand. *Probability in Banach spaces. Isoperimetry and processes*. Ergeb. Math. Grenzgeb. 23. Berlin: Springer, 1991.

[19] M. B. Marcus and J. Rosen. *Markov processes, Gaussian processes, and local times*. Cambridge Stud. Adv. Math. 100. Cambridge: Cambridge Univ. Press, 2006.

[20] S. Mendelson and G. Paouris. "On generic chaining and the smallest singular value of random matrices with heavy tails". In: *J. Funct. Anal.* 262 (2012), pp. 3775–3811.

[21] S. Montgomery-Smith. "Comparison of sums of independent identically distributed random vectors". In: *Probab. Math. Statist.* 14 (1992), pp. 281–285.

[22] G. Pisier. "Conditions d'entropie assurant la continuité de certains processus et applications à l'analyse harmonique". In: *Séminaire d'Analyse Fonctionnelle 1979-1980*. 13-14. Palaiseau: École Polytech.

[23] J. Rosiński. "On series representations of infinitely divisible random vectors". In: *Ann. Probab.* 18 (1990), pp. 405–430.

[24] K. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge Studies in Advanced Mathematics. 68. Cambridge: Cambridge University Press, 2013.

[25] S. J. Szarek. "On the best constants in Khinchin inequality". In: *Studia Math.* 58 (1976), pp. 197–208.

[26] W. Szatzschneider. "An Inequality". In: *SIAM Rev.* 33.1 (1991), pp. 116–118.

[27] M. Talagrand. "Majorising measures without measures". In: *Ann. Probab.* 29.1 (2001), pp. 411–417.

[28] M. Talagrand. "Regularity of Gaussian processes". In: *Acta Math.* 159 (1987), pp. 99–149.

[29] M. Talagrand. "Regularity of infinitely divisible processes". In: *Ann. Probab.* 20.1 (1993), pp. 362–432.

[30] M. Talagrand. "Sample Boundedness of Stochastic Processes Under Increment Conditions". In: *Ann. Probab.* 18.1 (1990), pp. 1–49.

[31] M. Talagrand. "The supremum of some canonical processes". In: *Amer. J. Math.* 116 (1994), pp. 283–325.

[32] M. Talagrand. *Upper and Lower Bounds for Stochastic Processes. Modern Methods and Classical Problems*. Ergeb. Math. Grenzgeb. 60. New York: Springer, 2014.

[33] M. Talagrand. *Upper and Lower Bounds for Stochastic Processes. Some Dreams Came True*. Preprint, 2020.

[34] R. Van Handel. "Probability in High Dimensions". Lecture Notes. 2016.

[35] R. Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge: Cambridge Univ. Press, 2018.