CONDITIONAL AND QUANTITATIVE STRONG LAWS OF LARGE NUMBERS

Thesis for the Degree of Doctor of Philosophy (PhD)

Nyanga Honda Masasila
Supervisor : Prof. Dr. István Fazekas

UNIVERSITY OF DEBRECEN

Doctoral Council for Natural Sciences and Engineering

Doctoral School of Informatics

Debrecen, 2026

Hereby I declare that I prepared this thesis within the Doctoral Council for Natural Sciences and Engineering, Doctoral School of Informatics, University of Debrecen in order to obtain a PhD Degree in Engineering at University of Debrecen. The results published in the thesis are not reported in any other PhD theses.

Debrecen, 2026.........

. . . .
Nyanga Honda Masasila

Hereby I confirm that Nyanga Honda Masasila candidate conducted his studies with my supervision within the Theoretical foundation and applications of information technology and stochastic systems Program of the Doctoral School of Informatics between 2022 and 2026. The independent studies and research work of the candidate significantly contributed to the results published in the thesis. I also declare that the results published in the thesis are not reported in any other theses.
I support the acceptance of the thesis.

Debrecen, 2026.............

. . . .
István Fazekas

CONDITIONAL AND QUANTITATIVE STRONG LAWS OF LARGE NUMBERS

Dissertation submitted in partial fulfilment of the requirements for the doctoral (PhD) degree in engineering.

Written by: Nyanga Honda Masasila certified applied mathematician

Prepared in the framework of the Theoretical foundation and applications of information technology and stochastic systems program of the Doctoral School of Informatics of the University of Debrecen.

Dissertation advisor: Prof. Dr. István Fazekas

The official opponents of the dissertation:

Dr. ....

Dr. ....

The evaluation board:

chairperson:	Dr. ....
members:	Dr. ....
	Dr. ....
	Dr. ....
	Dr. ....

The date and venue of the dissertation defence: .... 2026.

Abstract

The Strong Law of Large Numbers (SLLN) is a cornerstone of probability theory, guaranteeing almost sure convergence of empirical averages to expected value. The classical SLLN is possible under assumptions such as independence and identical distribution. However, modern stochastic models frequently operate in settings where information is revealed gradually, dependence structures are complex, or probabilities themselves are uncertain. This dissertation develops a unified and flexible framework for strong laws of large numbers that remains valid under conditioning, weak dependence, multi-indexed structures, and nonlinear expectations.

The first part of the dissertation is devoted to *conditional* strong laws of large numbers. Working on a probability space equipped with a sub-σ-algebra $\mathcal{F}$ representing available information, the thesis extends the abstract maximal inequality-based method of Fazekas and Klesov to the conditional setting. It is shown that a conditional Kolmogorov-type maximal inequality implies a conditional Hájek–Rényi inequality, and that this in turn yields almost sure convergence of suitably normalized partial sums. Importantly, this approach does not rely on specific dependence assumptions such as conditional independence, mixing or others. As a consequence, several known conditional SLLNs including results for $\mathcal{F}$-independent and conditionally negatively associated random variables are recovered as direct corollaries within a single methodological framework, leading to shorter and more general proofs.

The second part addresses *quantitative* strong laws for random variables indexed by two parameters. In many applications, data arise naturally as random fields rather than sequences, and convergence must be analyzed as both indices tend to infinity. The thesis introduces a geometric description of convergence rates using parametrized families of curves in the positive quadrant of $\mathbb{R}^2$. This allows convergence to be controlled via carefully chosen subsequences, from which convergence over the entire lattice can be deduced. Within this framework, explicit probability bounds are obtained for the deviations of normalized double sums, yielding quantitative SLLNs for pairwise independent and quasi-uncorrelated random fields. These results extend earlier one-dimensional quantitative laws and provide one of the first systematic treatments of convergence rates for double-indexed arrays under weak dependence.

The third part of the dissertation moves beyond classical probability by developing a general framework for strong laws of large numbers under conditional sub-additive expectations, motivated by models involving ambiguity, nonlinear risk evaluation, and uncertainty beyond additive probability measures. An axiomatic theory of conditional sub-additive expectations and capacities is established, within which conditional Kolmogorov-type and Hájek–Rényi-type maximal inequalities are derived. These inequalities yield strong laws of large numbers formulated in terms of conditional capacities and quasi-sure convergence, thereby extending classical almost sure convergence to nonlinear expectation spaces. A notion of conditional negative dependence is introduced in this abstract nonlinear setting, and corresponding max-

imal inequalities and strong laws are obtained without requiring classical independence assumptions.

In addition, this framework is further extended in Chapter 4 to sublinear expectation spaces, where strong laws of large numbers are established for φ-sub-Gaussian random variables. By exploiting the structural properties of sublinear expectations and associated capacities, new strong laws are proved for independent sub-Gaussian random variables under nonlinear expectations, providing convergence results under exponential-type tail control rather than moment conditions alone. These results demonstrate that sub-Gaussian concentration behavior is sufficient to ensure quasi-sure convergence of normalized sums in sublinear expectation spaces, thereby extending classical probabilistic strong laws to a broad class of uncertainty-sensitive stochastic models.

Together, these contributions provide a unified theory that integrates conditional probability, nonlinear expectation theory, capacity theory, and weak dependence structures. The results significantly broaden the scope of the strong law of large numbers, showing that maximal inequality methods remain effective in conditional, multi-index, and nonlinear expectation settings, and establishing strong laws under minimal structural assumptions in both additive and non-additive probabilistic frameworks.

Acknowledgments

First and foremost, I express my deepest gratitude to Almighty God for granting me life, strength, wisdom, and perseverance throughout my doctoral studies. Without His grace and guidance, the completion of this work would not have been possible.

I would like to extend my sincere and profound appreciation to my supervisor, Prof. Dr. István Fazekas, for his invaluable guidance, continuous encouragement, and unwavering support throughout the entire course of my research. His insightful comments, constructive criticism, and dedication to academic excellence greatly contributed to the successful completion of this dissertation.

I also wish to express my special appreciation to Dr. Borbála Fazekas, my MSc supervisor, for her excellent mentorship, academic guidance, and encouragement during my master's studies. I am grateful to acknowledge that she is the daughter of my PhD supervisor, Prof. Dr. István Fazekas, and I sincerely thank this remarkable family for providing outstanding supervision and scholarly support throughout both my MSc and PhD journeys. Their combined mentorship has had a profound impact on my academic development.

I am also deeply grateful to the Head of the Doctoral School of Informatics, Prof. Dr. Sándor Baran, for his leadership, academic support, and for providing a conducive research environment. My sincere thanks are further extended to all academic staff members of the Department of Applied Mathematics and Probability Theory at the University of Debrecen for their support, cooperation, and scholarly contributions during my study period.

I gratefully acknowledge the Tempus Public Foundation for sponsoring my studies through the scholarship award, which provided essential financial support and made it possible for me to pursue and successfully complete my doctoral studies. I also thank the management of the Dar es Salaam Institute of Technology, my employer, for granting me study leave and for their institutional support.

My sincere appreciation also goes to my close friends and flatmates, Michael Matonya and Dr. James Kachungwa, with whom I shared my daily life during my studies. Their friendship, encouragement, constructive discussions, and mutual support created a positive and motivating living environment that greatly contributed to my well-being and academic focus.

My heartfelt appreciation goes to my family for their unconditional love, patience, and sacrifices. I am profoundly thankful to my wife, Agness Mbago, for her remarkable tolerance of my frequent absence, her constant encouragement, understanding, and emotional support throughout the demanding years of my studies. I also thank my beloved sons, Pandi Nyanga Honda and Shija Nyanga Honda, for their patience, love, and for accepting the time I had to dedicate away from family life. Their smiles, prayers, and affection continuously motivated me to persevere.

Finally, I express my deepest gratitude to my parents, Honda Masasila Mbushi and Martha Pambe Ngoko, for their lifelong support, prayers, and for laying the foundation of my education and values. Their sacrifices and belief in me have been a constant source of inspiration throughout my academic journey.

To all those who contributed directly or indirectly to the successful completion of this work, I extend my sincere appreciation.

Contents

Introduction **1**

1 A General Approach to Conditional Strong Laws of Large Numbers **20**
- 1.1 Introduction . . . 20
- 1.2 Conditional strong law of large numbers via Hájek-Rényi inequality for expectations . . . 21
- 1.3 Conditional strong law of large numbers via Hájek-Rényi inequality for probabilities . . . 24
- 1.4 Applications . . . 27

2 Quantitative Strong Laws of Large Numbers for Random Variables with Double Indices **31**
- 2.1 Introduction . . . 31
- 2.2 Basic definitions . . . 32
- 2.3 General results . . . 36
- 2.4 Pairwise independent random variables . . . 39
- 2.5 Quasi uncorrelated random variables . . . 43

3 Strong Laws of Large Numbers for General Random Variables under Conditional Sub-additive Expectation and Capacity **46**
- 3.1 Introduction . . . 46
- 3.2 Sub-additive probability and expectation . . . 48
- 3.3 Conditional sub-additive probability and expectation . . . 50
- 3.4 Hájek-Rényi-type maximal inequalities for conditional sub-additive expectations and capacities . . . 54
- 3.5 Strong laws of large numbers in terms of conditional sub-additive expectations and capacities . . . 58
- 3.6 Application to conditionally negatively dependent random variables . . . 61
- 3.7 Discussion . . . 66

4 Strong law of large numbers for φ-sub-Gaussian random variables under sub-linear expectation spaces **68**

4.1 Introduction 68

4.2 Basic Properties of Capacities and Sub-linear Expectations 69

4.3 Sub-Gaussian Random Variables under Sub-linear Expectations 71

4.4 The Main Result 73

4.5 Strong law of large numbers for independent sub-Gaussian variables 74

Summary **78**

Bibliography **88**

Research Conferences and Workshop Participation

List of Publications

Introduction

The laws of large numbers (LLN) form one of the most fundamental pillars of probability theory and mathematical statistics. They formalize the intuitive principle that empirical averages of repeated random experiments stabilize around their expected values when the number of observations increases. Beyond their intrinsic theoretical importance, LLNs provide the mathematical foundation of statistical inference, stochastic modeling, information theory, Monte Carlo methods, econometrics, and modern data sciences.

Historically, the theory has evolved from elementary observations about Bernoulli trials into a rich and multifaceted area encompassing weak and strong modes of convergence, dependence structures, infinite-dimensional spaces, random fields, conditional frameworks and quantitative convergence rates. Each extension reflects the increasing complexity of modern stochastic models and the need for flexible mathematical tools capable of handling dependence, partial information, and multidimensional indexing.

Classical Foundations of the Law of Large Numbers

The earliest rigorous formulation of a law of large numbers is due to Jakob Bernoulli in his *Ars Conjectandi* (1713). In modern terminology, Bernoulli's Law of Large Numbers asserts that for a sequence of independent Bernoulli trials with success probability p, the empirical frequency of successes converges in probability to p. This result, now called the *Bernoulli Law of Large Numbers*, was precursor to what would later be formalized as weak law of large numbers (WLLN).

A major conceptual leap occurred with the development of the *Weak Law of Large Numbers (WLLN)* and the *Strong Law of Large Numbers (SLLN)* in the early twentieth century. The WLLN concerns convergence in probability of normalized sums, while the SLLN strengthens this to almost sure convergence. Khinchin proved the Weak Law of Large Numbers (WLLN) under the weaker assumption of pairwise independence. Specifically, if $\{X_n, n \geq 1\}$ is the sequence of pairwise independent and identically distributed random variables with finite mean μ, then $\frac{1}{n}\sum_{i=1}^{n} X_i \to \mu$ in probability as $n \to \infty$, see, for example, Chung [14] and Rényi [63]. This version of the weak law requires only the existence of the first moment and does not assume finite variance.

The classical form of the Strong Law of Large Numbers (SLLN) is due to Kolmogorov.

In 1933, Kolmogorov proved that if $\{X_n, n \geq 1\}$ is a sequence of independent (not necessarily identically distributed) random variables with mean zero and finite variances, and if the normalized variances satisfy the summability condition

$$\sum_{n=1}^{\infty} \frac{\mathrm{Var}(X_n)}{n^2} < \infty,$$

then

$$\frac{1}{n}\sum_{k=1}^{n} X_k \longrightarrow 0 \quad \text{almost surely.}$$

In the case where the random variables are independent and identically distributed (i.i.d.), the result takes a simpler classical form: if $\{X_n, n \geq 1\}$ is the sequence of i.i.d. random variables with finite mean $\mu = \mathbb{E}[X_1]$, then

$$\frac{1}{n}\sum_{k=1}^{n} X_k \longrightarrow \mu \quad \text{almost surely.}$$

We note that Kolmogorov's proof introduced powerful tools such as Kolmogorov's inequality and the three-series theorem.

Kolmogorov's inequality is a maximal inequality that provides an upper bound on the probability of deviations of partial sums. In its simplest form it states: if $\{X_k, 1 \leq k \leq n\}$ are independent random variables with zero mean and finite variances, then for any $\varepsilon > 0$,

$$\mathbb{P}\left(\max_{1\leq k\leq n} \left|S_k\right| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2}\sum_{j=1}^{n} \mathrm{Var}(X_j),$$

where $S_k = X_1 + \cdots + X_k$. An alternative but more general related tool is the *Hájek–Rényi inequality*, which was originally introduced in [30] by Hájek and Rényi (1955). Specifically, for independent zero-mean variables with finite variances and for $1 \leq m < n$, Hájek–Rényi's inequality asserts

$$\mathbb{P}\left(\max_{m\leq j\leq n} \frac{|S_j|}{b_j} \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2}\left(\sum_{j=m+1}^{n} \frac{\mathrm{Var}(X_j)}{b_j^2} + \frac{1}{b_m^2}\sum_{j=1}^{m} \mathrm{Var}(X_j)\right)$$

for any $\varepsilon > 0$ and any non-decreasing positive sequence b_j. Setting $b_j = 1$ and $m = 0$ recovers Kolmogorov's inequality as a special case. The Hájek–Rényi inequality is extremely useful in proving strong laws, as it controls the probability that any partial sum after time m exceeds a given threshold relative to a growth sequence b_j. In their original paper, Hájek and Rényi used this inequality to give an elegant proof of the SLLN for independent sequences . Later, many authors derived Hájek–Rényi type inequalities for dependent sequences to establish generalized strong laws, see, e.g., [23].

Besides Kolmogorov's classical theorem, several other foundational forms of the Strong Law of Large Numbers (SLLN) were developed. A particularly important advance was obtained by Marcinkiewicz and Zygmund, whose theorem provides a complete characterization of the integrability required for almost sure convergence under appropriate normalization.

In 1937, Marcinkiewicz and Zygmund proved the following result.

THEOREM. Let $0 < r < 2$, and let $\{X_n, n \geq 1\}$ be the sequence of independent and identically distributed random variables. If

$$\mathbb{E}|X_1|^r < \infty,$$

and additionally $\mathbb{E}X_1 = 0$ when $1 \leq r < 2$, then

$$\frac{S_n}{n^{1/r}} \xrightarrow{\text{a.s.}} 0 \qquad \text{as } n \to \infty.$$

Conversely, if this almost sure convergence holds, then necessarily $\mathbb{E}|X_1|^r < \infty$, together with $\mathbb{E}X_1 = 0$ when $1 \leq r < 2$.

Thus, the theorem gives an exact moment condition for strong convergence. This remarkable result contains Kolmogorov's SLLN as the special case $r = 1$. Importantly, the endpoint case $r = 2$ is *not* included in the Marcinkiewicz–Zygmund theorem. At this boundary a fundamentally different phenomenon appears. When

$$\mathbb{E}X = 0, \qquad \mathbb{E}X^2 = \sigma^2 < \infty,$$

the correct normalization no longer yields almost sure convergence but instead leads to convergence in distribution. This is the Central Limit Theorem (CLT): It is described as;

$$\frac{S_n}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, \sigma^2).$$

Thus, while the laws of large numbers describe deterministic stabilization of averages, the CLT is interpreted as a (distributional) rate result.

Between laws of large numbers and CLT lies one of the deepest results of probability theory, the Law of the Iterated Logarithm (LIL). Under the same assumptions

$$\mathbb{E}X = 0, \qquad \mathbb{E}X^2 = \sigma^2 < \infty,$$

the LIL determines the precise almost sure magnitude of the fluctuations:

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n \log\log n}} = \sigma, \quad \liminf_{n\to\infty} \frac{S_n}{\sqrt{2n \log\log n}} = -\sigma \qquad \text{a.s.}$$

The Law of the Iterated Logarithm therefore provides a parabolic bound on how large the oscillations of partial sums may be as the function of the number of summands. In fact, the SLLN shows that $S_n/n \to 0$ almost surely, the CLT describes fluctuations of size $\sqrt{n}$ in distribution, and the LIL identifies the exact maximal almost sure growth rate of these fluctuations, thereby forming a bridge between strong laws and limit distributions. For details see, e.g., [29] and [67].

In a seminal paper, Kolmogorov [41] proved the following LIL for independent, not necessarily identically distributed, random variables.

THEOREM (Kolmogorov). Suppose that $\{X_n, n \geq 1\}$ is the sequence of independent random variables with mean 0 and finite variances σ_k^2, $k \geq 1$. Set

$$S_n = \sum_{k=1}^{n} X_k, \qquad s_n^2 = \sum_{k=1}^{n} \sigma_k^2, \quad n \geq 1.$$

If

$$|X_n| = o\left(\frac{s_n}{\sqrt{\log\log s_n}}\right) \qquad \text{for all } n, \tag{0.0.1}$$

then

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2s_n^2 \log\log s_n^2}} = 1, \qquad \liminf_{n\to\infty} \frac{S_n}{\sqrt{2s_n^2 \log\log s_n^2}} = -1 \quad \text{a.s.}$$

Remark 0.0.1. Marcinkiewicz and Zygmund [47] provide an example which shows that the growth condition (0.0.1) cannot in general be weakened.

Condition (0.0.1) looks strange if one also assumes that the random variables are identically distributed; the condition should disappear.

The sufficiency part of the following LIL is due to Hartman and Wintner [31] while the necessity is due to Strassen [70].

THEOREM (Hartman–Wintner). Suppose that $\{X_n, n \geq 1\}$ is the sequence of independent random variables with mean 0 and finite variance σ^2, and set

$$S_n = \sum_{k=1}^{n} X_k, \qquad n \geq 1.$$

Then

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2\sigma^2 n \log\log n}} = 1, \qquad \liminf_{n\to\infty} \frac{S_n}{\sqrt{2\sigma^2 n \log\log n}} = -1 \quad \text{a.s.} \tag{0.0.2}$$

Conversely, if

$$\mathbb{P}\left(\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{n\log\log n}} < \infty\right) > 0,$$

then $\mathbb{E}X_1^2 < \infty$, $\mathbb{E}X_1 = 0$, and (0.0.2) holds.

Generalizations of the Law of Large Numbers

In modern probability theory, law of large numbers does not refer only to the classical scenario of i.i.d. sequences. The theorem has been extended and adapted to numerous settings that relax the assumptions of identical distribution or independence, consider different index sets (e.g. multiple indices or continuous indices), or even broaden the state space of the random variables (e.g. vector valued or Banach space-valued random elements). We outline some of these important generalizations, along with key contributions and conditions.

Relaxation of Independence. One major line of research has been to identify the weakest dependence assumptions under which a strong law still holds. Classical SLLN proofs break down if the random variables are not independent, but many dependence conditions have been studied that still allow for a law of large numbers. For example, one can consider sequences that are pairwise independent (any two distinct X_i, X_j are independent, but the sequence as a whole may not be mutually independent). This case was studied by Etemadi [19], who established the following result. Let $\{X_n, n \geq 1\}$ be a sequence of pairwise independent and identically distributed random variables with finite mean $\mu = \mathbb{E}[X_1]$, then

$$\frac{1}{n}\sum_{k=1}^{n} X_k \longrightarrow \mu \quad \text{almost surely.}$$

Etemadi used an interesting truncation method which demonstrates that full mutual independence is not necessary for almost sure convergence. In particular, his result provides a unified framework that encompasses both Kolmogorov's strong law of large numbers and Khinchin's weak law of large numbers.

Going further, it was shown that even weaker forms of pairwise independence suffice. In 1992, Matuła [51] proved that Kolmogorov's SLLN remains valid if the independence assumption is replaced by pairwise negative quadrant dependence (NQD). Negative quadrant dependence is a dependence structure implying that the variables are negatively correlated in a strong sense; Matuła's result thus significantly extended the scope of SLLN to certain dependent sequences. Another dependence condition of interest is exchangeability (invariant under finite permutations), see [20]. Etemadi and Kaminski's 1996 theorem implies a form of the LLN for exchangeable sequences (the sample average converges almost surely to a degenerate limit, which under mild conditions equals the common mean). In fact, pairwise independence or exchangeability can replace mutual independence in both the weak and strong laws.

Generalizations to non-identically distributed sequences were later obtained, notably by Csörgő, Tandori and Totik in [16], who removed the identical distribution assumption while maintaining pairwise independence as follows.

THEOREM. If the sequence of pairwise independent random variables $\{X_n, n \geq 1\}$ satisfy the conditions

$$\sum_{m=1}^{\infty} \frac{\mathrm{Var}(X_m)}{m^2} < \infty$$

and

$$\frac{1}{n}\sum_{m=1}^{n}\mathbb{E}|X_m - \mathbb{E}X_m| = O(1),$$

then

$$\lim_{n\to\infty}\frac{1}{n}\sum_{m=1}^{n}\big(X_m - \mathbb{E}X_m\big) = 0 \quad \text{almost surely.}$$

Beyond pairwise independence, researchers have considered sequences exhibiting various forms of weak dependence, including mixing sequences (such as strong mixing, weak mixing, ϕ-mixing, and ρ-mixing), martingale difference sequences, Markov chains, associated sequences, and others. A result by Birkhoff [5] in 1931 showed that the Strong Law of Large Numbers (SLLN) holds for stationary ergodic processes; this is essentially the ergodic theorem, which extends the classical law of large numbers to dependent stationary sequences. Strong laws for martingale differences have been extensively studied. For example, Fazekas and Klesov [23] established such results by building on earlier work of Brunk [6] and Prokhorov [61]. This result is commonly referred to as the Brunk–Prokhorov theorem for martingales.

In addition, strong laws of large numbers for dependent sequences satisfying ρ-mixing and ϕ-mixing conditions were obtained by Kuczmaszewska in [42]. She established almost sure convergence under appropriate summability conditions on the mixing coefficients together with suitable moment assumptions.

In the monograph by Pál Révész *The Laws of Large Numbers* [64], the strong law of large numbers (SLLN) was extended beyond the classical independence assumption to certain classes of weakly dependent random variables, including orthogonal sequences and stationary processes. In subsequent decades, substantial progress has been made in establishing SLLNs under increasingly general dependence structures. For instance, SLLNs for several classes of mixing and negatively associated random sequences was generalized by Christofides [13] to the framework of demimartingales. Hungarian researchers have also made notable contributions in this area. Tamás F. Móri [53] investigated strong laws for logarithmically weighted sums and obtained the exact almost sure growth rate when decreasing weights are assigned to the summands. Fazekas and Klesov [23] developed a general approach to SLLNs that does not rely on any specific dependence structure. This methodology was subsequently extended by Tómács and Libor [73] to a probability-based framework, replacing moment conditions by probabilistic ones.

Multi-Indexed Strong Laws. Another generalization concerns the index set of the random variables. The classical LLN deals with one-dimensional indices $n = 1, 2, \ldots$. However, in many applications (e.g. random fields, image processing, statistical physics models), we deal with arrays of random variables indexed by multi-dimensional indices (such as $\mathbb{N}^r$ for $r > 1$). This leads naturally to the study of sums of the form

$$S_{\boldsymbol{n}} = \sum_{\boldsymbol{k}\le\boldsymbol{n}} \xi_{\boldsymbol{k}}, \qquad \boldsymbol{n} = (n_1, \ldots, n_d) \in \mathbb{N}^r,$$

where $\boldsymbol{k} \leq \boldsymbol{n}$ denotes coordinatewise order. The fundamental question is whether a strong law of large numbers holds for

$$\frac{S_{\boldsymbol{n}}}{|\boldsymbol{n}|}, \qquad |\boldsymbol{n}| = \prod_{i=1}^{d} n_i,$$

as $\boldsymbol{n} \to \infty$ in a suitable sense. The answer is yes, under appropriate conditions that generalize Kolmogorov's criterion. R.T. Smythe (1973) was one of the first to establish strong laws for multi-dimensional arrays. By [68], Kolmogorov's SLLN is true for independent identically distributed random variables $\{\xi_{\boldsymbol{n}}, \boldsymbol{n} \in \mathbb{N}^r\}$, if and only if $\mathbb{E}|\xi_{\boldsymbol{n}}|(\log^+ |\xi_{\boldsymbol{n}}|)^{r-1} < \infty$. For pairwise independent identically distributed random variables $\xi_{i,j}$, Etemadi in [19] obtained the following SLLN. Let $S_{m,n} = \sum_{i=1}^{m} \sum_{j=1}^{n} \xi_{i,j}$. Then the condition $\mathbb{E}|\xi_{1,1}| \log^+ |\xi_{1,1}| < \infty$ implies $\lim_{m\to\infty, n\to\infty} \frac{S_{m,n}}{mn} \to \mathbb{E}\xi_{1,1}$ almost surely. Moreover, an analogue of the Hartman–Wintner law of the iterated logarithm for random fields was established in 1973 by Wichura [74].

The multidimensional SLLNs were further developed by Gut [27], who established Marcinkiewicz–Zygmund type strong laws and rate of convergence for multi-indexed random variables.

THEOREM (Theorem 3.2 in [27]). Let $\{\xi_{\boldsymbol{n}}, \boldsymbol{n} \in \mathbb{N}^r\}$ be i.i.d. random variables. Suppose that $\mathbb{E}|\xi|^p(\log^+ |\xi|)^{r-1} < \infty$, $0 < p < 2$ and set $\mathbb{E}\xi = 0$ if $p \geq 2$. Then

$$\frac{S_{\boldsymbol{n}}}{|\boldsymbol{n}|^{1/p}} \to 0 \quad \text{a.s. as} \quad \boldsymbol{n} \to \infty. \tag{0.0.3}$$

Conversely, (0.0.3) implies that $\mathbb{E}|\xi|^p(\log^+ |\xi|)^{r-1} < \infty$.

An interesting feature of multi-indexed settings is that several distinct notions of convergence arise, depending on how the index set tends to infinity. For example, one may require that $\frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \xi_{i,j}$ converges a.s. as $m, n \to \infty$ jointly, or perhaps in an iterated sense (first $m \to \infty$ then $n \to \infty$), sectorial convergence, and other constrained modes. In addition, some works study randomly indexed arrays (e.g. $m = m(n)$ for random function of n) and prove strong laws of large numbers under such indexing schemes.

Sectorial convergence is an interesting notion, arising when indices are restricted to a certain sector. For $0 < \theta < 1$, a sector is defined as;

$$\mathcal{T}_\theta^r = \left\{ \boldsymbol{n} \in \mathbb{N}^r : \quad \theta \leq \frac{n_i}{n_j} \leq \frac{1}{\theta}, \quad i \neq j, \quad i, j = 1, 2, \ldots, r \right\}.$$

It was shown by Gut [28] that when the index set is restricted to a sector, both the laws of large numbers and the law of the iterated logarithm hold under the same moment conditions as in the one-dimensional setting. Consequently, the logarithmic correction required for multi-indexed random variables disappears. In particular, Gut [28] established the following Marcinkiewicz–Zygmund strong law of large numbers for random variables indexed by a sector, from which the Kolmogorov strong law follows immediately.

THEOREM (Theorem 5.1 in [28]). Let $\{\xi_{\boldsymbol{n}}, \boldsymbol{n} \in T_\theta^r\}$ be i.i.d. random variables with $\mathbb{E}|\xi_1|^p < \infty$, $0 < p < 2$, and $\mathbb{E}\xi_1 = 0$ if $1 \le p < 2$. Then

$$|n|^{-1/p} S_\theta(n) \to 0 \quad \text{a.s. as } \boldsymbol{n} \to \infty, \tag{0.0.4}$$

where $S_\theta(n) = \sum_{\boldsymbol{k} \in T_\theta^r, \boldsymbol{k} < \boldsymbol{n}} \xi_{\boldsymbol{k}}$. Conversely, (0.0.4) implies that $\mathbb{E}|\xi_1|^p < \infty$.

A comprehensive treatment can be found in Oleg Klesov's monograph [39]: Limit Theorems for Multi Indexed Sums of Random Variables (2014), which provides a unified approach to LLNs and other limit theorems for random fields. Notably, Klesov gives new proofs of the multi-index Kolmogorov SLLN and Hájek–Rényi inequalities, and discusses strong laws under dependence in multi-index contexts. For example, one result from the book shows that Kolmogorov's SLLN extends to identically distributed random fields without requiring independence along diagonals, as long as appropriate correlation conditions are met.

Overall, the theory of multi-indexed strong laws demonstrates that the classical Kolmogorov SLLN extends naturally to random fields, but with important modifications. The appearance of logarithmic factors, the dependence on index geometry, and the existence of sectorial convergence distinguish the multiparameter case from the one-dimensional setting. These results provide the foundation for strong limit theory in random fields and continue to play an important role in modern probability theory.

Vector-valued and Banach space–valued SLLNs. Another important generalization of the law of large numbers concerns random elements taking values in $\mathbb{R}^d$ or, more generally, in Banach spaces. If$\{X_{\boldsymbol{n}}, \boldsymbol{n} \in \mathbb{N}^d\}$ are i.i.d. random vectors with mean $\mathbb{E}[\mathbf{X}_i] = \boldsymbol{\mu}$, then applying the scalar strong law of large numbers to each coordinate yields

$$\frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i \longrightarrow \boldsymbol{\mu} \quad \text{almost surely.}$$

Equivalently, this follows from standard vector-valued martingale arguments. Hence, in finite dimensions no additional assumptions beyond those of the scalar case are required; see, for example, [43, 44]. For infinite-dimensional Banach space-valued random variables, the situation is more delicate. The validity of a strong law in the norm topology depends on geometric properties of the Banach space. It was also showed that the notions of *type* and *cotype* of a Banach space play a fundamental role in limit theorems in Banach spaces; see [43, 59, 3].

Let $(B, \|\cdot\|)$ be a Banach space, and let (ε_i) be a sequence of independent Rademacher random variables, i.e.

$$\mathbb{P}(\varepsilon_i = -1) = \mathbb{P}(\varepsilon_i = 1) = \frac{1}{2},$$

and

$$\mathbb{E}[\varepsilon_i \varepsilon_m] = 0 \quad \text{for } i \neq m, \qquad \text{and} \qquad \mathrm{Var}(\varepsilon_i) = 1.$$

The notation $\mathbb{E}_\varepsilon$ means expectation with respect to the random variables ε.

We say that B is of *type* p for $p \in [1,2]$ if there exists a finite constant $C \geq 1$ such that

$$\mathbb{E}_\varepsilon \left[\left\| \sum_{i=1}^n \varepsilon_i x_i \right\|^p \right] \leq C^p \left[\sum_{i=1}^n \|x_i\|^p \right]$$

for all finite sequences $\{X_n, 1 \leq k \leq n\} \in B^n$.

The sharpest constant C is called the *type* p *constant* of B and is denoted by $T_p(B)$. From the triangle inequality, every Banach space is of type 1.

We say that B is of *cotype* q for $q \in [2,\infty]$ if there exists a finite constant $C \geq 1$ such that

$$\mathbb{E}_\varepsilon \left[\left\| \sum_{i=1}^n \varepsilon_i x_i \right\|^q \right] \geq \frac{1}{C^q} \left[\sum_{i=1}^n \|x_i\|^q \right], \quad \text{if } 2 \leq q < \infty,$$

respectively,

$$\mathbb{E}_\varepsilon \left[\left\| \sum_{i=1}^n \varepsilon_i x_i \right\| \right] \geq \frac{1}{C} \sup_{1 \leq i \leq n} \|x_i\|, \quad \text{if } q = \infty,$$

for all finite sequences $\{X_n, 1 \leq k \leq n\} \in B^n$. The sharpest constant C is called the *cotype* q *constant* of B and is denoted by $C_q(B)$. A Banach space is of type 2 and cotype 2 if and only if the space is also isomorphic to a Hilbert space.

In Banach spaces with suitable geometric properties, strong laws analogous to the classical Kolmogorov and Etemadi theorems continue to hold. The following result is a vector-valued version of the Marcinkiewicz–Zygmund strong law of large numbers for i.i.d. random variables

In Banach spaces of favorable geometry, strong laws analogous to the classical Kolmogorov and Etemadi theorems remain valid. The following result is a vector-valued version of the Marcinkiewicz–Zygmund strong law of large numbers for i.i.d. random variables. In the next theorem we consider a Banach space B for which there is a countable subset D of the unit ball of the dual space B' such that

$$\|x\| = \sup_{f \in D} |f(x)| \quad \text{for all } x \in B.$$

X is a random variable with values in B if $f(X)$ is measurable for all $f \in D$. When $(X_i)_{i \in \mathbb{N}}$ is a sequence of (independent) random variables in B we set, as usual,

$$S_n = X_1 + \cdots + X_n, \quad n \geq 1.$$

THEOREM (Theorem 7.9 in [43]). Let $0 < p < 2$. Let (X_i) be a sequence of i.i.d. random variables distributed like X with values in Banach space B. Then

$$\frac{S_n}{n^{1/p}} \to 0 \quad \text{almost surely}$$

if and only if

$$\mathbb{E}\|X\|^p < \infty \quad \text{and} \quad \frac{S_n}{n^{1/p}} \to 0 \quad \text{in probability.}$$

The following result investigates the relationship between the type condition and the i.i.d. SLLN of Marcinkiewicz–Zygmund.

THEOREM (Theorem 4.1 in [1]). Let B be a separable Banach space, $1 \leq p < 2$. The following conditions are equivalent.

1. B is of Rademacher type p.

2. For every independent identically distributed sequence $(X_j, j \geq 1)$ of B-valued random variables with $\mathbb{E}\|X_1\|^p < \infty$ and $\mathbb{E}X_1 = 0$, one has

$$\frac{S_n}{n^{1/p}} \to 0 \quad \text{a.s.}$$

where $S_n = \sum_{j=1}^n X_j$.

In summary, the law of large numbers has been generalized to cover dependent sequences, multi-indexed collections, and random elements in general spaces. A generalization typically requires finding an analogy of Kolmogorov's inequality or an appropriate maximal inequality, and verifying a convergence of a series or a similar summability condition. Many of these advances were driven by specific applications for instance, statistical physics motivated SLLNs for lattice-indexed random fields, and functional analysis motivated SLLNs in Banach spaces. The unifying theme is that the stabilization of averages is a pervasive phenomenon, but its exact conditions can vary widely depending on the context.

Convergence Rate Refinements to the SLLN

The classical SLLN is an existence theorem: it ensures that $\frac{S_n}{n} \to \mu$ almost surely, but it does not quantify how fast this convergence happens. A natural question is whether we can describe the rate of convergence in the strong law, i.e. find sequences $a_n \to \infty$ such that $a_n(\frac{S_n}{n} - \mu)$ converges to a non-degenerate limit (or 0) almost surely. Results that provide such quantitative refinements are often called quantitative strong laws or complete convergence theorems.

One of the first breakthroughs in this direction was given in [32] by Hsu and Robbins (1947). They showed that if $\{X_n, n \geq 1\}$ are i.i.d. with mean 0, and finite variance, then

$$\sum_{n=1}^{\infty} \mathbb{P}\left(|S_n| > n\varepsilon\right) < \infty$$

for every $\varepsilon > 0$. By the Borel–Cantelli lemma, this implies $\frac{1}{n}\sum_{i=1}^n X_i \to \mu$ almost surely. Paul Erdős [18] soon complemented this by showing the converse for the i.i.d. case: if $\sum_{n=1}^{\infty} \mathbb{P}(|S_n| > n\varepsilon) < \infty$, then necessarily $\mathbb{E}[X_1^2] < \infty$, that is, X_1 has finite variance.

Thus, the condition of Hsu–Robbins is not only sufficient but essentially necessary for an i.i.d. sequence to satisfy a quantitative form of SLLN. Thus, Hsu–Robbins–Erdős Strong Law is the following.

THEOREM. Let $\{X_n, n \geq 1\}$ be the sequence of independent and identically distributed random variables, and define

$$S_n = \sum_{k=1}^{n} X_k, \quad n \geq 1.$$

If $\mathbb{E}[X_1] = 0$ and $\mathbb{E}[X_1^2] < \infty$, then for every $\varepsilon > 0$,

$$\sum_{n=1}^{\infty} \mathbb{P}\left(|S_n| > n\varepsilon\right) < \infty.$$

Conversely, if the sum is finite for some $\varepsilon > 0$, then $\mathbb{E}[X_1] = 0$ and $\mathbb{E}[X_1^2] < \infty$, and the sum is finite for all $\varepsilon > 0$.

These results spurred a line of research into the exact rates of almost sure convergence. A notable milestone was the Baum–Katz theorem in [4], which generalizes the Hsu–Robbins theorem to higher moments. Baum and Katz proved that if $\mathbb{E}[|X_1|^p] < \infty$ and $\mathbb{E}(X_k) = \mu$ for some $p > 1$, $r > 1$ and $\frac{1}{2} < \frac{r}{p} \leq 1$, then

$$\sum_{n=1}^{\infty} n^{r-2}\, \mathbb{P}\Big(|S_n - n\mu| > \varepsilon n^{\frac{r}{p}}\Big) < \infty$$

for all $\varepsilon > 0$. Taking $r = p > 1$, we have

$$\sum_{n=1}^{\infty} n^{p-2}\, \mathbb{P}(|S_n - n\mu| > \varepsilon n) < \infty.$$

The case $p = 2$ recovers the Hsu–Robbins theorem.

One recent development of great interest is the use of proof mining in probability. In 2025, M. Neri [54] applied proof mining (a technique from mathematical logic that extracts quantitative bounds from non constructive proofs) to the strong law. Neri revisited classical proofs of the SLLN and was able to derive explicit rates of convergence that were previously only known to exist abstractly. In his paper, he combined the general method of Csörgő, Tandori and Totik in [16] with proof mining to get effective bounds on how large n needs to be for $\frac{S_n}{n}$ to be within ε of 0 (almost surely). Such results belong to the realm of effective probability theory and are quite novel, they connect the qualitative aspect of almost sure convergence with explicit quantitative modulus of convergence functions.

It is also worth noting specific results like the Baum–Katz–Spitzer theorem, which gives precise asymptotics in the SLLN for distributions in the domain of attraction of stable laws, and various results by Gut and Steinebach on complete convergence for arrays. For arrays (double-index), one can get rates of convergence along different directions in the index plane.

The above mentioned directions can be combined. E.g., in [21] Baum-Katz-type theorems were obtained for Banach space valued random variables with multidimensional indices.

To summarize this section: over the years, researchers have significantly sharpened the law of large numbers by determining how quickly the convergence occurs and under what finer conditions. The sequence S_n/n not only converges, but often does so with a specific speed that can be characterized by moment conditions. Key takeaways include: (i) the necessity of finite second moment for summable probability tails (Hsu–Robbins and Erdős' result), (ii) the spectrum of Baum–Katz theorems tying L^p moments to $n^{1/p}$ normalization (and generalizations thereof), (iii) the influence of heavy tails (where normalization by $n^{1/p}$ with $p < 1$ comes into play, as in Marcinkiewicz–Zygmund and in stable laws), and (iv) the extension of these quantitative laws to dependent and structured scenarios (mixing, association, multi-indexed cases) through clever uses of maximal inequalities. Many of these results will be relevant in the discussion of Chapter 2 of this dissertation, which is devoted to quantitative strong laws and will leverage some of the discussed techniques.

Conditional Strong Laws of Large Numbers

Traditional laws of large numbers are unconditional in the sense that they describe convergence with probability 1, without qualifications. In recent years, however, there has been growing interest in conditional versions of the strong law. An early version of the conditional law of large numbers appears in the work of Rényi [65]. A conditional SLLN typically asserts that the strong law holds in a filtered probability space, relative to some sub-σ-algebra that carries additional information. In other words, we study sequences of random variables X_n along with a sub-σ-algebra $\mathcal{F}$ (often $\mathcal{F}$ is the σ-algebra of "conditioning events" or some background information) and aim to prove that

$$\frac{1}{n}\sum_{i=1}^{n} X_i \longrightarrow \mu \quad \text{a.s.}$$

and that this convergence remains true under conditional probability given $\mathcal{F}$. Equivalently, we want

$$\mathbb{P}\left(\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i = L \,\middle|\, \mathcal{F}\right) = 1 \quad \text{a.s.}$$

for some limit L (often $L = \mathbb{E}[X_1|\mathcal{F}]$ or similar).

One motivation for conditional laws is in non-identically distributed or dynamically changing environments. For instance, consider a sequence of random variables X_n where at each time n, some side information $\mathcal{F}$ is revealed (such as a σ-algebra generated by another process or some underlying random parameter). We might be interested in laws of large numbers conditional on that side information. If $\mathcal{F}$ is trivial (contains only events of probability 0 or 1), the conditional law reduces to the usual law. If $\mathcal{F}$ is larger, a conditional SLLN can inform us that given the information in $\mathcal{F}$, the average of the $X_n's$ converges almost surely.

This is stronger than the unconditional convergence because it implies a kind of robustness to additional knowledge.

The formal study of conditional SLLNs was initiated by the work of Wu, Jiang, and others in the 2000s, but foundational results appeared in the mid-2000s. For example, Majerek, Nowak and Zięba in [46] considered a notion of $\mathcal{F}$-independence, where X_n are independent relative to a sub-σ algebra $\mathcal{F}$. They proved a conditional strong law for $\mathcal{F}$-independent sequences: if X_n are $\mathcal{F}$-independent (meaning roughly that X_n are independent and each X_n is $\mathcal{F}$-measurable or something akin to that) and $\mathbb{E}[|X_n|] < \infty$, then S_n/n converges $\mathcal{F}$-almost surely to 0 (assuming zero conditional means). Their proofs relied on establishing a conditional Kolmogorov inequality within the probability space augmented by $\mathcal{F}$. In essence, they showed that conditioning on $\mathcal{F}$ one can still control the maximal deviations of partial sums.

Following this, Prakasa Rao [62] published a comprehensive study on conditional independence, conditional mixing and conditional association. In that work, various strong limit theorems are proved under conditional versions of classical dependence assumptions. For example, Prakasa Rao defined what it means for a sequence to be conditionally ϕ-mixing or conditionally associated relative to a sub-σ-algebra $\mathcal{F}$, and then proved SLLNs assuming those conditions. These results are quite technical; a simplified description is: if (X_n) is conditionally independent given $\mathcal{F}$ (meaning that, roughly, $\mathbb{P}(X_{i_1} \in A_1, \ldots, X_{i_k} \in A_k \mid \mathcal{F}) = \prod_{j=1}^{k} \mathbb{P}(X_{i_j} \in A_j \mid \mathcal{F})$ a.s. for any finite set of indices), and $\mathbb{E}[X_n \mid \mathcal{F}] = 0$ a.s., then a conditional SLLN holds (i.e. $S_n/n \to 0$ a.s). He obtained analogous results for sequences that are conditionally associated or conditionally weakly dependent. One needs to assume appropriate moment conditions to handle the conditioning.

The key technical tools in conditional SLLNs are conditional forms of Kolmogorov's and Hájek–Rényi's inequalities. A conditional Kolmogorov inequality might state:

$$\mathbb{P}\left(\max_{1\le k\le n} |S_k| \ge \varepsilon \,\middle|\, \mathcal{F} \right) \le \frac{1}{\varepsilon^2} \sum_{i=1}^{n} \mathbb{E}\big(X_i^2 \mid \mathcal{F} \big) \quad \text{a.s.}$$

From it, one can derive a conditional Hájek–Rényi inequality of the form

$$\mathbb{P}\left(\max_{m\le k\le n} \left| \frac{\sum_{j=1}^{k} X_j}{b_k} \right| \ge \varepsilon \,\middle|\, \mathcal{F} \right) \le \varepsilon^{-2} \left(\sum_{j=m+1}^{n} \frac{\mathbb{E}(X_j^2 \mid \mathcal{F})}{b_j^2} + \frac{1}{b_m^2} \sum_{j=1}^{m} \mathbb{E}(X_j^2 \mid \mathcal{F}) \right) \quad \text{a.s.}$$

for appropriate b_j and $m < n$. Versions of the conditional Hájek–Rényi inequality for conditionally associated random variables were established in [10]. Subsequently, Hye Young Seo and Jong-il Baek [66] extended these results to conditionally negatively associated sequences. In both settings, the corresponding conditional Hájek–Rényi inequalities were employed to derive conditional strong laws of large numbers (SLLNs).

The interest in conditional SLLNs is partly theoretical (extending the classical limit theory to a conditional framework aligned with the concept of regular conditional probabilities) and partly practical. In practice, one often has a situation where some background information is present and we want to update our law of large numbers conclusions in the presence of that information. For example, in adaptive algorithms or stochastic approximation, at each step there is some σ-field representing past information, and one might want an "almost sure convergence given the past" result.

To illustrate concretely: suppose $X_n = Y_n \cdot Z$ where Z is some random variable independent of the sequence Y_n. Here one might take $\mathcal{F} = \sigma(Z)$, the sigma-field generated by Z. A conditional SLLN in this context might say $\frac{1}{n}\sum_{i=1}^{n} Y_i Z \to \mathbb{E}[Y_1|Z]$ almost surely, and indeed for each fixed Z, almost surely in Y. In other words, conditionally on Z, the average behaves as if Z were a constant. Results of this flavor can be derived from the general conditional SLLN theory.

The literature on conditional SLLNs is still developing. The works by Majerek et al. and Prakasa Rao cited above laid the groundwork. Subsequent papers have often focused on specific dependence structures. The significance of the conditional Kolmogorov and Hájek–Rényi inequalities cannot be overstated they are the "enabling lemmas" that make conditional SLLNs possible. Once one has a conditional maximal inequality, one can mimic Kolmogorov's proof in the conditional setting: use the maximal inequality to show $\sum_n \mathbb{P}(\max_{k\le n}|S_k| > \varepsilon \mid \mathcal{F}) < \infty$ almost surely, which implies (by the conditional Borel–Cantelli lemma) that $\max_{k\le n}|S_k| < \varepsilon$ eventually (almost surely, and in fact $\mathcal{F}$-conditionally almost surely). Such arguments are carefully executed in the references above.

In conclusion, conditional SLLNs extend the law of large numbers to a conditional probability setting. They require new techniques but ultimately parallel the unconditional case: one proves conditional versions of fundamental inequalities, then establishes almost sure convergence given the conditioning information. The results by Majerek–Nowak–Zięba and Prakasa Rao in the mid 2000s are key milestones. The study of conditional laws is not only mathematically rich but also aligns with modern probabilistic modeling where layered or hierarchical randomness is present.

Strong Laws of Large Numbers on Non-Additive Framework

In classical probability theory, uncertainty is quantified by an additive probability measure $\mathbb{P}$ on a sigma-algebra $\mathcal{F}$, and expectations are taken with respect to this single reference measure. However, in many situations one faces ambiguity or model uncertainty for example, incomplete knowledge of the true distribution or deliberate modeling of multiple prior distributions. In such cases it is natural to consider non-additive measures or capacities, which generalize probabilities by dropping countable additivity while retaining monotonicity. The properties of non-additive expectations and probabilities are discussed in details in Chapters 3 and 4.

The formal study of non-additive measures began with Choquet's seminal 1954 paper *Theory of Capacities* [11]. Choquet introduced capacities of various orders and laid the measure-theoretic foundation for integrating with respect to a capacity (the Choquet integral). This work provided a rigorous analytic backbone for what would much later be called imprecise probabilities. Choquet's integral allows one to define expectations even when probabilities are not precisely known or additive, by essentially averaging with respect to the upper probability $\hat{\mathbb{V}}$ or lower probability v. One notable property is that for any random variable X, the upper Choquet integral with respect to $\hat{\mathbb{V}}$ is always at least as large as the lower Choquet integral with respect to v, and they coincide if and only if X has no ambiguity in expectation.

While Choquet's work provided the mathematical foundation, the idea of using capacities in statistical inference was pioneered by Peter Huber and Volker Strassen [36] in 1973. In their influential paper "Minimax Tests and the Neyman–Pearson Lemma for Capacities," Huber and Strassen introduced capacities to model robust hypothesis testing problems. They considered uncertainty classes of probability distributions; for example, all distributions within a certain "distance" of a nominal model and showed that the classical Neyman–Pearson theory of optimal tests can be extended to this multiple-prior setting by replacing probabilities with capacities. The influence of Huber–Strassen is also evident in later limit theorems for non-additive probabilities.

A breakthrough came with Massimo Marinacci's work in the late 1990s. Marinacci [48] proved several limit laws for non-additive probabilities, including a form of the strong law for i.i.d. sequences under a totally monotone capacity (also called a belief function or necessity measure in some contexts). He showed that under regularity conditions (including a continuity condition on the capacity), the $\liminf$ and $\limsup$ of the sample averages lie between the lower and upper Choquet expected values of X_1 with capacity 1 (i.e. quasi-surely). This result was a natural extension of Kolmogorov's SLLN which is recovered as the special case $v = \hat{\mathbb{V}} = \mathbb{P}$ (a single probability measure so that $C_v[X] = C_{\hat{\mathbb{V}}}[X] = \mathbb{E}[X]$). Marinacci's conditions included continuity of the capacity (roughly, avoiding pathological "jumps" for decreasing sequences of sets) and the requirement that the capacity be totally monotone (a property ensuring it arises from a nested family of additive measures, common in the theory of belief functions). Subsequent refinements by Maccheroni and Marinacci [45] in 2005 relaxed some assumptions. They proved a SLLN for capacities on Polish spaces under either continuity of the capacity or continuity of the bounded i.i.d. random variables.

A major modern development in probability under uncertainty came with Shige Peng's work on sublinear expectations in the 2000s. Peng introduced a systematic framework for nonlinear expectation spaces $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$. He introduced the G expectation as a nonlinear expectation that generates a time-consistent nonlinear Brownian motion (the G-Brownian motion). Peng's seminal papers and monographs [56, 58, 57], laid down a rigorous theory of stochastic integration, martingales, and even a nonlinear Feynman–Kac formula under this sublinear expectation.

Peng's work has sparked a surge of research into limit theorems under sublinear expectations. Many researchers in the last decade have extended these results, often aiming to weaken assumptions or derive analogues for dependent variables. A series of papers by Chen, Zhang, Hu, Li, etc. between 2010–2016 developed maximal inequalities, three-series theorems, and other probabilistic tools in the sublinear context . These culminated in more general SLLNs that require neither identical distributions nor the full Peng independence. Notably, Huang and Wu [35] proved strong laws for general random variables in a sublinear expectation space without assuming any independence.

Chen [8] proved a strong law under an upper probability without needing the capacity to be continuous. Instead, Chen assumed the random variables are quasi-continuous and that the upper probability arises from a weakly compact set of measures, leveraging an approach inspired by Huber–Strassen's results. Following that, Chen, Wu, and Li [9] established a strong law of large numbers for general non-additive probabilities (upper expectations) using Peng's notion of independence. An interesting contemporary result closely related to this dissertation is the work of Zhang, Tang, and Xiong [78], who established a conditional strong law of large numbers under G expectations.

Overview of This Dissertation's Results

Having surveyed the broad landscape of laws of large numbers from classical theorems to generalizations, convergence rates, conditional versions, and non-additive framework, we now turn to the specific contributions of this dissertation. Here we provide a motivating overview of these results and how they relate to the existing work discussed above.

Chapter 1 of this dissertation (based on the paper by Fazekas and Masasila [24]) develops a general approach to conditional strong laws of large numbers. The motivation for this work stems from the observation that many earlier conditional SLLNs were proved on a case-by-case basis, often under specific conditional independence or mixing assumptions (as in Majerek et al. 2005 and Prakasa Rao 2009). There wasn't a single unified theorem that one could apply to derive multiple conditional SLLNs as corollaries in contrast to the unconditional case, where Kolmogorov's SLLN and its proof techniques cover a wide variety of scenarios. Our goal in Chapter 1 was to fill this gap by providing a general, inequality-based framework for obtaining conditional SLLNs.

The main result of Chapter 1 can be summarized (somewhat informally) as follows: if one can establish a conditional Kolmogorov-type inequality for a sequence of random variables relative to a sub-σ-algebra $\mathcal{F}$, then automatically one obtains a conditional Hájek–Rényi-type inequality and hence a conditional SLLN . In other words, the implication "conditional Kolmogorov inequality $\implies$ conditional Hájek–Rényi $\implies$ conditional SLLN" is proved in great generality. We treat both probability inequalities (moment-free) and L^p-moment inequalities in the conditional setting. This framework is powerful: it allows us to recover as special cases the results of Majerek–Nowak–Zięba (for $\mathcal{F}$-independent sequences) and Prakasa Rao (for conditionally independent, mixing, or associated sequences) by simply verifying their conditions ensure a conditional Kolmogorov inequality. Moreover, we derive new results for cases that were not explicitly tackled before– for instance, we prove a conditional SLLN for conditionally negatively associated sequences, which to our knowledge had not been addressed in prior literature, by using a conditional version of the Seo Baek (2012) inequality.

One highlight is that our method in Chapter 1 does not require any specific structure of dependence: as long as you can supply a suitable maximal inequality (conditional Kolmogorov), you get the SLLN. This is analogous to the approach of Fazekas and Klesov (2001) in the unconditional case, who showed that a Hájek–Rényi type inequality for moments yields a host of strong laws for dependent sequences . Here, we elevate that philosophy to the conditional realm. The contribution is therefore a unification and generalization: it unifies proofs of known theorems and generalizes them to new conditional contexts. We hope this will serve as a "black box" tool for future researchers investigating conditional analogues of other limit theorems.

In Chapter 2 (based on Fazekas and Masasila [25]), the focus shifts to quantitative strong laws for double-indexed random variables, particularly establishing rates of convergence. We consider random fields $X_{n,m} : n, m \in \mathbb{N}$ and study the convergence of their double averages $S_{n,m}/(nm)$, where $S_{n,m} = \sum_{i=1}^{n} \sum_{j=1}^{m} X_{i,j}$. The novelty in Chapter 2 is twofold: (i) we obtain explicit convergence rates in the strong law for such random fields, and (ii) we introduce a method involving subsequences and logical analysis (inspired by proof mining) to transfer convergence rates from one sequence to another.

A key result of Chapter 2 is the following type of theorem: if along some carefully chosen subsequence (n_k, m_k) one can establish an almost sure convergence with a certain rate, then that rate actually applies to the full array indexed by (n, m). We develop a quantitative theory of SLLNs for random variables with double indices. The chapter introduces a geometric description of convergence rates by parametrized families of curves in $\mathbb{R}^2_+$, and shows that controlling convergence along suitable subsequences implies convergence along the full lattice. Using this framework, explicit probability bounds of the form

$$\mathbb{P}\left(\sup_{m \geq n} \left|\frac{S_m}{|m|}\right| > \varepsilon\right) \leq C\,\frac{1}{|n|\,\varepsilon^p}$$

are obtained for pairwise independent and quasi-uncorrelated random fields. This provides one of the first systematic treatments of quantitative SLLNs for double-index arrays under weak dependence conditions.

The methodology in Chapter 2 combines classical probability with a form of logical analysis: we carefully analyze the proofs of convergence (often using Borel–Cantelli lemma and exponential bounds) to extract how the tail events decay in n, m. By doing so, we contribute to the growing area of quantitative analysis of probabilistic theorems, which is quite new. Practically, having a rate of convergence is valuable. For example, if one uses a double-indexed random field to model an image or a data table, knowing that the average converges is good, but knowing it converges at (say) a rate of Λ a.s. gives a handle on finite-sample behavior. Our results in Chapter 2 are among the first of their kind for random fields.

The results in Chapter 3 (based on Masasila and Fazekas [50]) build upon the rich history outlined above and push it a step further by developing conditional and quantitative forms of the strong law under very general nonlinear expectations. In particular, we adopt an axiomatic approach to conditional sub-additive expectations and capacities (Section 3.2 of the dissertation) and establish SLLNs without requiring any independence assumptions. This approach was inspired by the general approach introduced in Chapter 1. Here we extend those methods to the non-additive realm. For example, in Section 3 we prove conditional Hájek–Rényi-type inequalities via conditional Kolmogorov type inequalities for sub-additive expectations and capacities, which then serve as the main tool to derive almost sure convergence. This significantly generalizes earlier results (which almost universally assumed some form of independence and identical distribution under conditional sublinear expectation).

The results in Chapter 4 (based on [49]) further extend the theory of strong laws of large numbers to the framework of sublinear expectation spaces, with particular emphasis on φ-sub-Gaussian random variables. While Chapter 3 develops general strong laws under conditional sub-additive expectations and capacities using axiomatic and maximal inequality methods, Chapter 4 focuses on convergence results under exponential-type tail control in nonlinear expectation environments. Specifically, we investigate φ-sub-Gaussian random variables under sublinear expectations and establish strong laws of large numbers under independence assumptions adapted to the sublinear framework. By exploiting the structural properties of capacities and sublinear expectations, as well as the concentration behavior inherent in sub-Gaussian random variables, we derive sufficient conditions ensuring quasi-sure convergence of normalized partial sums. The main results demonstrate that sub-Gaussian-type tail decay provides a natural and powerful alternative to classical moment conditions in guaranteeing strong convergence under nonlinear expectations. These findings complement the general axiomatic framework developed in Chapter 3 and show that strong laws remain valid in sublinear expectation spaces under structurally meaningful distributional assumptions. Consequently, this chapter strengthens the connection between nonlinear expectation theory, capacity-based probability, and modern concentration-based methods in probability theory.

Motivation and Significance of the Research

The research presented in this dissertation is motivated by the need to extend the classical strong law of large numbers (SLLN) to modern stochastic environments in which data are often dependent, multi-dimensional, and influenced by partial or evolving information. While the classical SLLN provides a powerful asymptotic guarantee, many contemporary applications such as network data analysis, spatial statistics, adaptive algorithms, and large scale simulations require results that remain valid under conditioning, weak dependence, complex indexing structures, and non-additive models. Moreover, practical applications frequently demand not only convergence itself, but also quantitative information about the rate at which convergence occurs. This dissertation aims to close these gaps.

The significance of the research can be summarized in four main aspects:

1. Unification of conditional strong laws: Earlier conditional SLLNs were typically established under specific assumptions, such as conditional independence or conditional mixing, and required separate proofs for each dependence structure. By developing a general inequality-based approach, this dissertation shows that once a suitable conditional Kolmogorov-type maximal inequality is available, a conditional Hájek–Rényi in-

equality and a conditional SLLN follow automatically. This provides a flexible "black-box" principle that simplifies proofs and unifies many known conditional results within a single framework.

2. Quantitative theory for multi-indexed random fields: Although strong laws for random fields are well studied, explicit convergence rates for multi-indexed arrays under weak dependence remain limited. This work introduces a geometric and subsequence-based approach that allows convergence rates to be transferred from carefully chosen subsequences to the entire lattice. As a result, explicit probability bounds and quantitative SLLNs are obtained for pairwise independent and quasi-uncorrelated random fields. These results contribute to a systematic quantitative theory for multi-indexed strong laws.

3. Methodological innovation through logical analysis: The dissertation incorporates ideas inspired by proof mining and logical analysis to extract quantitative information from convergence proofs that are traditionally qualitative. This interdisciplinary methodology bridges probability theory and mathematical logic, opening new perspectives for obtaining effective bounds in limit theorems.

4. Extension of strong laws to nonlinear and uncertainty-based frameworks: Beyond classical additive probability models, this dissertation extends strong laws of large numbers to conditional sub-additive and sublinear expectation spaces. Chapters 3 and 4 develop strong laws formulated in terms of capacities and quasi-sure convergence, allowing limit theorems to remain valid under ambiguity and model uncertainty. In particular, the work establishes conditional strong laws without classical independence assumptions and proves convergence results for φ-sub-Gaussian random variables under sublinear expectations using exponential-type tail control. These results broaden the applicability of strong laws to modern uncertainty-sensitive stochastic models arising in risk theory, robust statistics, and nonlinear probability.

Relevance for applications and future research. Quantitative strong laws are essential in statistics, machine learning, and simulation-based optimization, where convergence speed determines reliability and computational efficiency. Conditional strong laws are equally important in stochastic processes, filtering theory, and models with hierarchical or partial information. By extending SLLNs to conditional and nonlinear expectation frameworks, this work provides tools that are potentially applicable to risk theory, robust statistics, and ambiguity-sensitive models.

In summary, this dissertation strengthens the theoretical foundations of the strong law of large numbers by unifying conditional principles, extending quantitative convergence theory to random fields, and introducing new methodological tools. These contributions not only deepen theoretical understanding but also enhance the applicability of strong laws in complex stochastic systems.

Chapter 1

A General Approach to Conditional Strong Laws of Large Numbers

This chapter develops a general method for establishing conditional strong laws of large numbers (SLLNs) by means of conditional expectation and probability inequalities. It is shown that a conditional Kolmogorov type inequality implies a conditional Hájek-Rényi type inequality and this implies strong law of large numbers. Both probability and moment inequalities are considered. Some applications are offered in the last section.

1.1 Introduction

In this chapter, we study conditional strong laws of large numbers for arbitrary random variables. So, let $\{X_n, n \geq 1\}$ be a sequence of random variables defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$. The partial sums of random variables are denoted as $S_n = \sum_{i=1}^{n} X_i$ for $n \geq 1$ and $S_0 = 0$.
For several decades, numerous findings, modifications and applications concerning the strong law of large numbers have been studied. In [23], Fazekas and Klesov presented a general approach to establish the strong law of large numbers for sequences of random variables. Significantly, their method does not impose any restriction on the underlying dependence structure of random variables but it needs a Kolmogorov type inequality. The main aim of this chapter is to obtain the conditional version of the results of Fazekas and Klesov [23]. In our proofs we use the ideas given in [23].

In the last two decades several papers were devoted to conditional versions of well-known theorems of probability theory. In [46] Majerek, Nowak and Zieba studied the conditional strong law of large numbers for $\mathcal{F}$-independent random variables where $\mathcal{F}$ is a σ-subalgebra of $\mathcal{A}$. Their main results were obtained via conditional Kolmogorov's inequality. Prakasa Rao [62] beside conditional indepedence, studied also conditional mixing and conditional association.

In this chapter we shall show, that a Kolmogorov type inequality implies a Hájek-Rényi type inequality and this implies a strong law of large numbers. This approach can be used both for conditional probabilities and for conditional expectations. In the last section of this chapter, we present several applications of the main result. Using our approach we offer alternative proofs to the following theorems: the conditional strong law of large numbers for $\mathcal{F}$-independent random variables (Theorem 3.5 in [46]), a general version of the conditional strong law of large numbers (Theorem 6 in [62]) and strong law of large numbers for conditionally negatively associated random variables (Theorem 3.1 (b) in [66]).

To prove convergence in our major results, we shall use the well-known theorem of Abel and Dini for real-valued non-random sequences.

Proposition 1.1.1 (The Abel-Dini theorem). Let $b_1, b_2, \dots$ be positive real numbers. If $\sum_{k=1}^{\infty} b_k$ converges, then with $T_n = \sum_{k=n}^{\infty} b_k$ as the n^{th} tail sum, then $\sum_{n=1}^{\infty} \frac{b_n}{T_n^{1+\alpha}}$ converges if and only if $\alpha < 0$. For the proof, see [40].

1.2 Conditional strong law of large numbers via Hájek-Rényi inequality for expectations

In this section, we show that a conditional Kolmogorov type inequality implies a conditional Hájek-Rényi type inequality and this implies a strong law of large numbers without assuming further weak dependence conditions. We note that all inequalities and conditions in the theorems of this section hold almost surely.

Theorem 1.2.1 (Fazekas & Masasila [24])**.** *Let $\{X_k, 1 \le k \le n\}$ be a sequence of random variables, let $S_k = X_1 + \dots + X_k$. Let $\mathcal{F}$ be a σ-subalgebra, $\alpha_1, \dots, \alpha_n$ be nonnegative $\mathcal{F}$-measurable random variables, $r > 0$ real number. Assume that the general conditional Kolmogorov's type inequality is true, that is*

$$\mathbb{E}\left(\left[\max_{1\le l\le m} |S_l|\right]^r |\mathcal{F}\right) \le \sum_{l=1}^{m} \alpha_l \qquad \text{for all} \quad 1 \le m \le n. \tag{1.2.1}$$

Then the conditional Hájek-Rényi inequality is true, that is

$$\mathbb{E}\left(\left[\max_{1\le l\le n} \left|\frac{S_l}{\beta_l}\right|\right]^r |\mathcal{F}\right) \le 4\sum_{l=1}^{n} \frac{\alpha_l}{\beta_l^r} \tag{1.2.2}$$

for $\mathcal{F}$-measurable random variables $\beta_1 \le \beta_2 \le \dots \le \beta_n$ with $\beta_1 \ge \beta_0$, where β_0 is a positive constant.

Proof. We can assume that $\beta_1 \ge 1$ during the proof. Let $c = 2^{\frac{1}{r}}$. Let $A_i = \{k : c^i \le \beta_k < c^{i+1}\}, \quad i = 0, 1, 2, \dots$ Then A_i is $\mathcal{F}$-measurable, because β_k is $\mathcal{F}$-measurable. A_i is the set

of subscripts k for which $c^i \leq \beta_k < c^{i+1}$. Let $i(n)$ be the index of the last non-empty A_i. Then $i(n)$ is an $\mathcal{F}$-measurable random variable (possibly having infinity value). Let $k(i)$ be the maximal index in A_i. More precisely,
$k(i) = \max\{k : k \in A_i\}$, if A_i is nonempty, but $k(i) = k(i-1)$ if A_i is empty ($k(-1) = 0$ by definition). Let

$$\delta_l = \sum_{j=k(l-1)+1}^{k(l)} \alpha_j, \quad l = 0, 1, 2, \ldots, \quad \text{be the sum of } \alpha' s \quad \text{in} \quad A_l.$$

Then $k(i)$ and δ are $\mathcal{F}$-measurable, $k(i) \leq n$. Then using (1.2.1), we obtain

$$\begin{aligned}
\mathbb{E}\left(\left[\max_{1\leq l\leq n} \frac{|S_l|}{\beta_l}\right]^r |\mathcal{F}\right) &\leq \sum_{i=0}^{i(n)} \mathbb{E}\left(\left[\max_{l\in A_i} \frac{|S_l|}{\beta_l}\right]^r |\mathcal{F}\right) \\
&\leq \sum_{i=0}^{i(n)} c^{-ir}\mathbb{E}\left(\left[\max_{l\in A_i} |S_l|\right]^r |\mathcal{F}\right) \\
&\leq \sum_{i=0}^{i(n)} c^{-ir}\mathbb{E}\left(\left[\max_{k\leq k(i)} |S_k|\right]^r |\mathcal{F}\right) \\
&\leq \sum_{i=0}^{i(n)} c^{-ir} \sum_{k=1}^{k(i)} \alpha_k = \sum_{i=0}^{i(n)} c^{-ir} \sum_{l=0}^{i} \delta_l \\
&= \sum_{l=0}^{i(n)} \delta_l \sum_{i=l}^{i(n)} c^{-ir} \leq \sum_{l=0}^{i(n)} \delta_l \sum_{i=l}^{\infty} c^{-ir} \\
&= \frac{1}{1-c^{-r}} \sum_{l=0}^{i(n)} c^{-lr}\delta_l \\
&= \frac{1}{1-c^{-r}} \sum_{l=0}^{i(n)} c^{-lr} \sum_{k=k(l-1)+1}^{k(l)} \alpha_k \\
&\leq \frac{1}{1-c^{-r}} \sum_{l=0}^{i(n)} c^{-lr} \sum_{k=k(l-1)+1}^{k(l)} \alpha_k \frac{c^{lr+r}}{\beta_k^r} \\
&= \frac{c^r}{1-c^{-r}} \sum_{l=0}^{i(n)} \sum_{k=k(l-1)+1}^{k(l)} \frac{\alpha_k}{\beta_k^r} = 4\sum_{k=1}^{n} \frac{\alpha_k}{\beta_k^r}.
\end{aligned}$$

During the proof we applied that in A_i we have $1 < \frac{c^{l+1}}{\beta_k}$. We also mention that we applied (1.2.1) for random number of terms, i.e., instead of m we applied it for $k(i)$. One can show that (1.2.1) is true for this index, as $k(i)$ is $\mathcal{F}$-measurable and $k(i) \leq n$. □

Theorem 1.2.2 (Fazekas & Masasila [24]). *Let $\{X_n, n \geq 1\}$ be a sequence of random vari-*

ables, $S_n = X_1 + \cdots + X_n$ for any n. Let $b_0 \le b_1 \le b_2 \le \ldots$ be $\mathcal{F}$-measurable random variables with $b_n \longrightarrow \infty$ a.s., where b_0 is a positive constant. Let $\alpha_1, \alpha_2, \ldots$ be nonnegative $\mathcal{F}$-measurable random variables. Let $r > 0$ be a fixed number. Assume that for any $n \ge 1$

$$\mathbb{E}\left(\left[\max_{1\le l\le n} |S_l|\right]^r |\mathcal{F}\right) \le \sum_{l=1}^{n} \alpha_l. \tag{1.2.3}$$

If $\sum_{l=1}^{\infty} \frac{\alpha_l}{b_l^r} < \infty$ *a.s., then*

$$\lim_{n\longrightarrow\infty} \frac{S_n}{b_n} = 0 \qquad a.s. \tag{1.2.4}$$

Proof. We can assume that $\alpha_n > 0$ for all n a.s. To see it let a non random $\alpha'_n > 0$, for any n and $\sum_n \alpha'_l < \infty$. Then instead of α_n we can consider $\max(\alpha_n, \alpha'_n)$.
Assume that $\alpha_n \ge \alpha'_n > 0$ and α'_n is non random for any n . Let

$$v_n = \sum_{k=n}^{\infty} \frac{\alpha_k}{b_k^r}, \qquad \beta_n = \max_{1\le k\le n} b_k v_k^{\frac{1}{2r}}.$$

Then the sequence β_n is increasing, $\beta_1 > \beta_0 > 0$ where β_0 is non random. Then, because of the assumption $\sum_{l=1}^{\infty} \frac{\alpha_l}{b_l^r} < \infty$ a.s., we have

$$0 < v_n < \infty \quad \text{for all} \quad n \quad \text{a.s.}, \; v_n \longrightarrow 0 \quad \text{a.s.},$$

and v_n is a decreasing sequence. Then, using the Abel-Dini theorem,

$$\sum_{n=1}^{\infty} \frac{\alpha_n}{b_n^r v_n^{\frac{1}{2}}} < \infty \quad \text{a.s.}$$

Therefore we have $0 < \beta_0 \le \beta_1 \le \beta_2 \le \ldots,\quad \beta_0$ is non random,

$$\sum_{k=1}^{\infty} \frac{\alpha_k}{\beta_k^r} < \infty,$$

$$\lim_{k\longrightarrow\infty} \frac{\beta_k}{b_k} = 0 \qquad \text{a.s.}$$

Then our previous theorem implies

$$\mathbb{E}\left(\max_{1\le l\le n} \left|\frac{S_l}{\beta_l}\right|^r |\mathcal{F}\right) \le 4\sum_{l=1}^{n} \frac{\alpha_l}{\beta_l^r} \quad \text{for all} \quad n.$$

So, by the monotone convergence theorem,

$$\mathbb{E}\left(\sup_{1\le l\le\infty} \left|\frac{S_l}{\beta_l}\right|^r |\mathcal{F}\right) \le 4\sum_{l=1}^{\infty} \frac{\alpha_l}{\beta_l^r} < \infty \quad \text{a.s.}$$

So

$$\sup_{1\le l\le\infty}\left|\frac{S_l}{\beta_l}\right|^r<\infty \quad \text{a.s.}$$

Therefore

$$0\le\left|\frac{S_l}{b_l}\right|=\left|\frac{S_l}{\beta_l}\right|\frac{\beta_l}{b_l}\le\left(\sup_{1\le l\le\infty}\left|\frac{S_l}{\beta_l}\right|\right)\frac{\beta_l}{b_l}\longrightarrow 0 \quad \text{a.s. as} \quad l\longrightarrow\infty.$$

□

1.3 Conditional strong law of large numbers via Hájek-Rényi inequality for probabilities

Here we offer the same approach as in the previous section, but we use conditional probabilities instead of conditional expectations. As before, the inequalities and conditions in the theorems of this section hold almost surely.

Theorem 1.3.1 (Fazekas & Masasila [24]). *Let $\{X_k, 1\le k\le n\}$ be a sequence of random variables, $S_k=X_1+\cdots+X_n$. Let $\mathcal{F}$ be a σ-subalgebra. Let r be a positive real number. Let $\beta_1\le\beta_2\le\cdots\le\beta_n$ be $\mathcal{F}$-measurable, $\alpha_1,\ldots,\alpha_n$ nonnegative $\mathcal{F}$-measurable random variables. Assume that $\beta_1\ge\beta_0>0$, where β_0 is non random. If*

$$\mathbb{P}\left(\max_{1\le l\le m}|S_l|\ge\varepsilon|\mathcal{F}\right)\le\frac{1}{\varepsilon^r}\sum_{l=1}^m\alpha_l \qquad \textit{for all} \quad 1\le m\le n \tag{1.3.1}$$

and for all $\varepsilon>0$, then

$$\mathbb{P}\left(\max_{1\le l\le n}\left|\frac{S_l}{\beta_l}\right|\ge\varepsilon|\mathcal{F}\right)\le\frac{4}{\varepsilon^r}\sum_{k=1}^n\frac{\alpha_k}{\beta_k^r} \tag{1.3.2}$$

for all $\varepsilon>0$.

Proof. Using the same notation as in the proof of Theorem 1.2.1, we have

$$\begin{aligned}\mathbb{P}\left(\max_{1\le l\le n}\frac{|S_l|}{\beta_l}\ge\varepsilon|\mathcal{F}\right)&\le\sum_{i=0}^{i(n)}\mathbb{P}\left(\max_{l\in A_i}\frac{|S_l|}{\beta_l}\ge\varepsilon|\mathcal{F}\right)\\&\le\sum_{i=0}^{i(n)}\mathbb{P}\left(\max_{l\in A_i}\frac{|S_l|}{c^i}\ge\varepsilon|\mathcal{F}\right)\\&\le\sum_{i=0}^{i(n)}\mathbb{P}\left(\max_{k\le k(i)}\frac{|S_k|}{c^i}\ge\varepsilon|\mathcal{F}\right)\\&\le\sum_{i=0}^{i(n)}(\varepsilon c^i)^{-r}\sum_{k=1}^{k(i)}\alpha_k=\sum_{i=0}^{i(n)}(\varepsilon c^i)^{-r}\sum_{l=0}^{i}\delta_l\end{aligned}$$

$$\begin{aligned}
&= \sum_{l=0}^{i(n)} \delta_l \sum_{i=l}^{i(n)} (\varepsilon c^i)^{-r} \leq \sum_{l=0}^{i(n)} \delta_l \sum_{i=l}^{\infty} (\varepsilon c^i)^{-r} \\
&= \varepsilon^{-r} \frac{1}{1-c^{-r}} \sum_{l=0}^{i(n)} c^{-lr} \delta_l \\
&= \varepsilon^{-r} \frac{1}{1-c^{-r}} \sum_{l=0}^{i(n)} c^{-lr} \sum_{k=k(l-1)+1}^{k(l)} \alpha_k \\
&\leq \varepsilon^{-r} \frac{1}{1-c^{-r}} \sum_{l=0}^{i(n)} c^{-lr} \sum_{k=k(l-1)+1}^{k(l)} \alpha_k \frac{c^{lr+r}}{\beta_k^r} \\
&= \varepsilon^{-r} \frac{c^r}{1-c^{-r}} \sum_{l=0}^{i(n)} \sum_{k=k(l-1)+1}^{k(l)} \frac{\alpha_k}{\beta_k^r} \\
&= 4\varepsilon^{-r} \sum_{k=1}^{n} \frac{\alpha_k}{\beta_k^r}.
\end{aligned}$$

□

Theorem 1.3.2 (Fazekas & Masasila [24]). *Let $\{X_n, n \geq 1\}$ be a sequence of random variables, $S_k = X_1 + \cdots + X_k$. Let $\mathcal{F}$ be a σ-subalgebra. Let $b_0 \leq b_1 \leq b_2 \dots$ be $\mathcal{F}$-measurable random variables with $b_n \longrightarrow \infty$ a.s., where b_0 is positive constant. Let $\alpha_1, \alpha_2, \dots$ be non-negative $\mathcal{F}$-measurable random variables. Let $r > 0$ be a fixed number. Assume that for any $n \geq 1$*

$$\mathbb{P}\left(\max_{1 \leq l \leq n} |S_l| \geq \varepsilon |\mathcal{F}\right) \leq \frac{1}{\varepsilon^r} \sum_{l=1}^{n} \alpha_l \qquad \text{for all} \quad \varepsilon > 0. \tag{1.3.3}$$

If $\sum_{l=1}^{\infty} \frac{\alpha_l}{b_l^r} < \infty$ *a.s., then*

$$\lim_{n \longrightarrow \infty} \frac{S_n}{b_n} = 0 \qquad a.s. \tag{1.3.4}$$

Proof. Assume that $\alpha_n \geq \alpha_n' > 0$ where α_n' is non random for any n. Let

$$v_n = \sum_{k=n}^{\infty} \frac{\alpha_k}{b_k^r}, \qquad \beta_n = \max_{1 \leq k \leq n} b_k v_k^{\frac{1}{2r}}.$$

Then, because of the assumption $\sum_{l=1}^{\infty} \frac{\alpha_l}{b_l^r} < \infty$ a.s., we have

$$0 < v_n < \infty \quad \text{for all} \quad n \geq 1 \quad \text{a.s. and } v_n \longrightarrow 0 \quad \text{a.s.}$$

Moreover, the Abel-Dini's theorem implies

$$\sum_{n=1}^{\infty} \frac{\alpha_n}{b_n^r v_n^{\frac{1}{2}}} < \infty \quad \text{a.s.}$$

Therefore $\beta_1, \beta_2, ...$ is an increasing sequence, $\beta_1 \geq \beta_0 > 0$, where β_0 is non random,

$$\sum_{k=1}^{\infty} \frac{\alpha_k}{\beta_k^r} < \infty,$$

$$\lim_{k \longrightarrow \infty} \frac{\beta_k}{b_k} = 0 \qquad \text{a.s.}$$

Then our previous theorem implies

$$\mathbb{P}\left(\max_{1 \leq l \leq n} \frac{|S_l|}{\beta_l} \geq \varepsilon | \mathcal{F} \right) \leq \frac{4}{\varepsilon^r} \sum_{l=1}^{n} \frac{\alpha_l}{\beta_l^r} \qquad \text{for all} \quad n \quad \text{and} \quad \varepsilon > 0.$$

So, by the monotone convergence theorem,

$$\mathbb{P}\left(\sup_{1 \leq l < \infty} \frac{|S_l|}{\beta_l} \geq \varepsilon | \mathcal{F} \right) \leq \frac{4}{\varepsilon^r} \sum_{l=1}^{\infty} \frac{\alpha_l}{\beta_l^r}.$$

Let $\varepsilon \longrightarrow \infty$. We have

$$\sup_{1 \leq l < \infty} \frac{|S_l|}{\beta_l} < \infty \quad \text{a.s.}$$

Now

$$0 \leq \left| \frac{S_l}{b_l} \right| = \left| \frac{S_l}{\beta_l} \right| \frac{\beta_l}{b_l} \leq \left(\sup_{1 \leq l < \infty} \frac{|S_l|}{\beta_l} \right) \frac{\beta_l}{b_l} \longrightarrow 0 \quad \text{a.s., as} \quad l \longrightarrow \infty$$

because $\frac{\beta_l}{b_l} \longrightarrow 0$ a.s. Therefore,

$$\lim_{n \longrightarrow \infty} \frac{S_n}{b_n} = 0 \qquad \text{a.s.}$$

□

Remark 1.3.3. It is important to emphasize that the results presented in Section 1.2 are closely related to those in Section 1.3. In particular, by means of the conditional Markov inequality, one can obtain (1.3.1) from (1.2.1) and (1.3.2) from (1.2.2). However, the converse implication does not hold in general. In what follows, we state the conditional Markov inequality and provide a counterexample demonstrating the failure of the converse.

Lemma 1.3.4 (Conditional Markov inequality). Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, let $\mathcal{F} \subseteq$

$\mathcal{A}$ be a sub–σ–algebra, and let $X \geq 0$ be an integrable random variable. Then, for every $t > 0$,

$$\mathbb{P}(X \geq t \mid \mathcal{F}) \leq \frac{\mathbb{E}[X \mid \mathcal{F}]}{t} \qquad \text{a.s.}$$

Example 1.3.5 (Failure of the converse). Let $\mathcal{F} = \{\emptyset, \Omega\}$ be the trivial σ–algebra, and let X be a random variable with Pareto distribution of parameter $\alpha = 1$ and $k = 1$, i.e., with density

$$f(x) = \frac{1}{x^2}, \qquad x \geq 1.$$

Then, for every $t \geq 1$,

$$\mathbb{P}(X \geq t) = \int_t^\infty \frac{1}{x^2}\, dx = \frac{1}{t}.$$

Since $\mathcal{F}$ is trivial, conditional probabilities and expectations coincide with their unconditional counterparts. Hence,

$$\mathbb{P}(X \geq t \mid \mathcal{F}) = \frac{1}{t}, \qquad t \geq 1,$$

so that X satisfies the conditional tail bound

$$\mathbb{P}(X \geq t \mid \mathcal{F}) \leq \frac{C}{t}, \qquad \text{for some constant} \quad C > 1.$$

On the other hand, the conditional expectation is infinite:

$$\mathbb{E}[X \mid \mathcal{F}] = \mathbb{E}[X] = \int_1^\infty x \cdot \frac{1}{x^2}\, dx = \int_1^\infty \frac{1}{x}\, dx = \infty.$$

Therefore, a bound of the form

$$\mathbb{P}(X \geq t \mid \mathcal{F}) \leq \frac{C}{t}$$

does not imply a corresponding bound

$$\mathbb{E}[X \mid \mathcal{F}] \leq C.$$

1.4 Applications

It is well known that when the trivial σ-algebra $\mathcal{F} = \{\emptyset, \Omega\}$ is considered, conditional expectations and probabilities reduce to their unconditional (non-random) counterparts. Consequently, the main results of this chapter coincide with those obtained in [23]. Therefore, this can be viewed as a direct application of the main results developed in this chapter.

Next, we consider conditional Kolmogorov's strong law of large numbers for $\mathcal{F}$-independent random variables. In [46] conditionally independent random variables were studied and

Kolmogorov-type strong laws of large numbers were obtained. In this subsection, we prove Theorem 3.5 in [46] using our general approach. Let $\sigma^2_{\mathcal{F}}(X) = \mathbb{E}\left\{(X - (\mathbb{E}X|\mathcal{F}))^2 |\mathcal{F}\right\}$ denote the conditional variance of X.

Theorem 1.4.1 (Majerek, Nowak & Zięba [46]). *Let $\{X_n, n \geq 1\}$ be a sequence of $\mathcal{F}$-independent random variables such that $\sum_{k=1}^{\infty} \frac{\sigma^2_{\mathcal{F}} X_k}{k^2} < \infty$ a.s. Let $S_n = X_1 + \cdots + X_n, n = 1, 2, \ldots$ Then*

$$\lim_{n \longrightarrow \infty} \frac{S_n - \mathbb{E}(S_n|\mathcal{F})}{n} = 0 \qquad a.s. \tag{1.4.1}$$

Proof. For $\mathcal{F}$-independent random variables the Kolmogorov inequality presented in [46] is

$$\mathbb{P}\left(\max_{1 \leq k \leq n} |S_k - \mathbb{E}(S_k|\mathcal{F})| \geq \varepsilon|\mathcal{F}\right) \leq \sum_{k=1}^{n} \frac{1}{\varepsilon^2} \sigma^2_{\mathcal{F}} X_k. \tag{1.4.2}$$

Then (1.4.2) is the condition (1.3.1) in the Theorem 1.3.1 for $r = 2$. As $\sum_{k=1}^{\infty} \frac{\sigma^2_{\mathcal{F}} X_k}{k^2} < \infty$ a.s., so we can apply Theorem 1.3.2. Therefore

$$\lim_{n \longrightarrow \infty} \frac{S_n - \mathbb{E}(S_n|\mathcal{F})}{n} = 0 \qquad \text{a.s.}$$

□

Remark 1.4.2. By using Theorem 1.3.1, we can obtain the Hájek-Rényi type inequality for conditionally independent random variables as

$$\mathbb{P}\left(\max_{1 \leq k \leq n} \left|\frac{S_k - \mathbb{E}(S_k|\mathcal{F})}{k}\right| \geq \varepsilon|\mathcal{F}\right) \leq \frac{4}{\varepsilon^2} \sum_{k=1}^{n} \frac{\sigma^2_{\mathcal{F}} X_k}{k^2}.$$

Prakasa Rao in [62] obtained a general version of the conditional strong law of large numbers proved in [46]. We apply our Theorem 1.3.2 to prove the following theorem (Theorem 6 in [62]).

Theorem 1.4.3 (Prakasa Rao [62]). *If $\{X_n, n \geq 1\}$ is a sequence of $\mathcal{F}$-independent random variables such that*

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}\left(|X_n - \mathbb{E}(X_n|\mathcal{F})|^{2r} |\mathcal{F}\right)}{n^{r+1}} < \infty \quad a.s., \tag{1.4.3}$$

for some $r \geq 1$, then

$$\frac{S_n - \mathbb{E}(S_n|\mathcal{F})}{n} \longrightarrow 0 \quad a.s. \ as \quad n \longrightarrow \infty. \tag{1.4.4}$$

Proof. By the Kolmogorov inequality in Theorem 4 of [62] for $r \geq 1$, and by inequality (5.1) of [62]

$$\begin{aligned}\mathbb{P}\left(\max_{1\le k\le n}|S_k-\mathbb{E}(S_k|\mathcal{F})|\ge\varepsilon|\mathcal{F}\right)&\le\frac{1}{\varepsilon^{2r}}\mathbb{E}\left(|S_n-\mathbb{E}(S_n|\mathcal{F})|^r\ge\varepsilon|\mathcal{F}\right)\\&\le\frac{1}{\varepsilon^{2r}}n^{r-1}\sum_{k=1}^{n}\mathbb{E}\left(|X_k-\mathbb{E}(X_k|\mathcal{F})|^{2r}\ge\varepsilon|\mathcal{F}\right)\\&=\frac{1}{\varepsilon^{2r}}\Lambda_n,\end{aligned}$$

where, $\Lambda_n=n^{r-1}\sum_{k=1}^{n}\mathbb{E}\left(|X_k-\mathbb{E}(X_k|\mathcal{F})|^{2r}\ge\varepsilon|\mathcal{F}\right)$. We want to represent Λ_n as $\Lambda_n=\alpha_1+\cdots+\alpha_n$. Let $A_k=\mathbb{E}\left(|X_k-\mathbb{E}(X_k|\mathcal{F})|^{2r}\ge\varepsilon|\mathcal{F}\right)$. Then

$$\begin{aligned}\alpha_n=\Lambda_n-\Lambda_{n-1}&=n^{r-1}\sum_{k=1}^{n}A_k-(n-1)^{r-1}\sum_{k=1}^{n-1}A_k\\&=n^{r-1}A_n+[n^{r-1}-(n-1)^{r-1}]\sum_{k=1}^{n-1}A_k.\end{aligned}$$

We have to show, that $\sum_{n=1}^{\infty}\frac{\alpha_n}{n^{2r}}<\infty$.

$$\sum_{n=1}^{\infty}\frac{\alpha_n}{n^{2r}}=\sum_{n=1}^{\infty}\frac{n^{r-1}}{n^{2r}}A_n+\sum_{n=1}^{\infty}\frac{n^{r-1}-(n-1)^{r-1}}{n^{2r}}\sum_{k=1}^{n-1}A_k.$$

Changing the order of the summation in the second term, we obtain

$$\begin{aligned}\sum_{k=1}^{\infty}A_k\sum_{n=k+1}^{\infty}\frac{n^{r-1}-(n-1)^{r-1}}{n^{2r}}&\le\sum_{k=1}^{\infty}A_k\sum_{n=k+1}^{\infty}\frac{Cn^{r-2}}{n^{2r}}\\&=C\sum_{k=1}^{\infty}A_k\sum_{n=k+1}^{\infty}n^{-r-2}\\&\le C\sum_{k=1}^{\infty}A_k\int_k^{\infty}x^{-r-2}dx\\&\le C\sum_{k=1}^{\infty}A_kk^{-r-1},\end{aligned}$$

where we used the mean value theorem and approximation with integral. So

$$\sum_{n=1}^{\infty}\frac{\alpha_n}{n^{2r}}\le C\sum_{n=1}^{\infty}\frac{A_n}{n^{r+1}}<\infty$$

using condition (1.4.3). So

$$\mathbb{P}\left(\max_{1\leq k\leq n}|S_k-\mathbb{E}(S_k|\mathcal{F})|\geq\varepsilon|\mathcal{F}\right)\leq\frac{1}{\varepsilon^{2r}}C\sum_{k=1}^{n}\alpha_k$$

where $\sum_{n=1}^{\infty}\frac{\alpha_n}{n^{2r}}<\infty$. So our Theorem 1.3.2 implies that

$$\frac{S_n-\mathbb{E}(S_n|\mathcal{F})}{n}\longrightarrow 0\quad\text{a.s. as}\quad n\longrightarrow\infty.$$

□

Now we show that our approach gives a quick proof of Theorem 3.1 (b) of [66].

Theorem 1.4.4 (Seo & Baek [66]). *Let b_n be an increasing sequence of positive real numbers, $b_n\longrightarrow\infty$. Let $\{X_n,n\geq 1\}$ be a sequence of conditionally centered $\mathcal{F}$-negatively associated random variables, $1\leq r\leq 2$. Assume that $\sum_{n=1}^{\infty}\frac{\mathbb{E}(|X_n|^r|\mathcal{F})}{b_n^r}<\infty$ a.s. Then*

$$\frac{1}{b_n}\sum_{k=1}^{n}X_k\longrightarrow 0\quad\textit{a.s. as}\quad n\longrightarrow\infty. \tag{1.4.5}$$

Proof. For our random variables the following Kolmogorov type inequality is true.

$$\mathbb{E}\left(\max_{1\leq k\leq n}\left|\sum_{i=1}^{k}X_i\right|^r|\mathcal{F}\right)\leq C\sum_{i=1}^{k}\mathbb{E}\left(|X_i|^r|\mathcal{F}\right)\quad\text{a.s.,}$$

see Lemma 2.1 of [66]. Then our Theorem 1.2.2 gives the result without any further calculation. □

Chapter 2

Quantitative Strong Laws of Large Numbers for Random Variables with Double Indices

The main objective of this chapter is to obtain an explicit rate of convergence in the strong law of large numbers for double-indexed families of random variables, using proof-mining techniques as employed by Neri in [54]. We show that a rate of convergence along a suitable subsequence yields a rate of convergence for the entire sequence. This result is then applied to derive rates of convergence in the strong law of large numbers for double-indexed families of random variables that are pairwise independent, quasi-uncorrelated, and negatively dependent.

2.1 Introduction

The famous results concerning the rate of convergence in the SLLN were proved by Hsu, Robbins, Erdős, Spitzer, Baum, and Katz. A particular case of the results is the following theorem. Let $\{\xi_n, n \geq 1\}$ be independent identically distributed random variables, $S_n = \sum_{i=1}^{n} \xi_i$, $r > 1$, and assume that $\mathbb{E}|\xi_1|^r < \infty$, $\mathbb{E}\xi_1 = 0$. Then

$$\sum_{n=1}^{\infty} n^{r-2} \mathbb{P}\left(\sup_{k \geq n} |S_k/k| > \varepsilon\right) < \infty \quad \text{for all} \quad \varepsilon > 0,$$

see, e.g. [29].

In [54], Neri also studied the rate of convergence in the SLLN. Theorem 1.4 of [54] is the following.

Proposition 2.1.1 (Neri [54]). Suppose $\{\xi_n, n \geq 1\}$ is a sequence of pairwise independent random variables satisfying $\mathbb{E}(\xi_k) = 0$, $\mathbb{E}(|\xi_k|) \leq \tau$ and $\mathrm{Var}(\xi_k) \leq \sigma^2$, for all k and some

$\tau > 0$, $\sigma > 0$. Let $S_n = \sum_{i=1}^{n} \xi_i$. There exists a universal constant $\kappa \leq 1536$ such that for all $0 < \varepsilon \leq \tau$,

$$\mathbb{P}\left(\sup_{m \geq n} |S_m/m| > \varepsilon\right) \leq \frac{\kappa\sigma^2\tau}{n\varepsilon^3}.$$

In the proof of the above theorem, some methods of [16] were applied.

In this chapter, we shall use the ideas of [54] to obtain rate of convergence results for the double indices version of the SLLN. It is known that for the multi-index SLLN we need a stronger moment assumption than for the usual single index case. By [68], Kolmogorov's SLLN is true for independent identically distributed random variables $\xi_{\boldsymbol{n}}$, $\boldsymbol{n} \in \mathbb{N}^r$, if and only if $\mathbb{E}|\xi_{\boldsymbol{n}}|(\log^+ |\xi_{\boldsymbol{n}}|)^{r-1} < \infty$. Here and in what follows $\mathbb{N}$ denotes the set of positive integers.

For pairwise independent identically distributed random variables $\xi_{i,j}$, Etemadi in [19] obtained the following SLLN. Let $S_{m,n} = \sum_{i=1}^{m} \sum_{j=1}^{n} \xi_{i,j}$. Then condition $\mathbb{E}|\xi_{1,1}| \log^+ |\xi_{1,1}| < \infty$ implies $\lim_{m\to\infty, n\to\infty} \frac{S_{m,n}}{mn} \to \mathbb{E}\xi_{1,1}$ almost surely. The main aim of our this chapter is to find the rate of convergence result in this SLLN using ideas of [54].

A related approach based on subsequences was developed by Matuła and Seweryn in [52], where it was shown that, for random fields, almost sure convergence along Etemadi-type subsequences implies convergence for the entire index set. The present work extends this idea to a quantitative setting, establishing that rates of convergence along suitable subsequences yield rates for the full double-indexed sequence.

In Section 2.2 of chapter, we give a possible description of the rate of convergence in the case of two-dimensional indices. In Section 2.3, we show how we can apply certain subsequences to obtain the rate of convergence in the law of large numbers. In Section 2.4, we find the rate of convergence in the strong law of large numbers for pairwise independent random variables with double indices. In Section 2.5, we extend the results of the previous section to quasi uncorrelated random variables.

2.2 Basic definitions

First, we give a description of the convergence rate for sequences of random variables with double indices. In our point of view, it can be defined by a parametrized family of curves in $\mathbb{R}_+^2 = [0, \infty) \times [0, \infty)$.

For short, we shall denote the function and its graph by the same letter. Let $a, b > 0$ be fixed. Let R_{ab}^1 be the set of the graphs of all functions f_{ab}^1 with the following properties: $f_{ab}^1 : [a, \infty) \to [0, b]$ is a non-increasing continuous function with $f_{ab}^1(a) = b$. We denote by R_{ab}^2 the set of those graphs f_{ab}^2 which are reflections of graphs $f_{ab}^1 \in R_{ab}^1$ to the line $f(x) = x$, $x \in \mathbb{R}$.

We define R_{ab} as the set of all curves $\boldsymbol{f}_{ab} = f_{ab}^1 \cup f_{ba}^2$, with $f_{ab}^1 \in R_{ab}^1$ and $f_{ba}^2 \in R_{ba}^2$. As the point (a, b) belongs both to f_{ab}^1 and f_{ba}^2, so $\boldsymbol{f}_{ab}$ is a continuous curve in $\mathbb{R}_+^2$. Now, let

$$R = \bigcup_{a,b>0} R_{ab}.$$

We see that any $\boldsymbol{f} \in R$ is a continuous curve in $\mathbb{R}^2_+$.

Any $\boldsymbol{f} \in R$ divides $\mathbb{R}^2_+$ into two disjoint parts: $A_0^{\boldsymbol{f}}$ and $A_1^{\boldsymbol{f}}$ so that $A_0^{\boldsymbol{f}}$ contains the origin $\mathbf{0} = (0, 0)$. For $\boldsymbol{p} \in \mathbb{R}^2_+$, we shall write that $\boldsymbol{f} \leq \boldsymbol{p}$, if $\boldsymbol{p} \in A_1^{\boldsymbol{f}}$. For $\boldsymbol{p} \in \boldsymbol{f}$, we shall accept that $\boldsymbol{f} \leq \boldsymbol{p}$. If $\boldsymbol{f} \leq \boldsymbol{p}$ but $\boldsymbol{p} \notin \boldsymbol{f}$, then we shall write that $\boldsymbol{f} < \boldsymbol{p}$. For $\boldsymbol{f}, \boldsymbol{g} \in R$, we shall write that $\boldsymbol{f} \leq \boldsymbol{g}$ if $\boldsymbol{f} \leq \boldsymbol{p}$ for any point $\boldsymbol{p} \in \boldsymbol{g}$. To characterize the rate of convergence, we shall use the following type of parametrized families. Let $(\Gamma, \leq)$ be a (non-empty) partially ordered set. We assume that for any $\gamma \in \Gamma$ there exists an element of R which we denote by $\boldsymbol{f}_\gamma$. We assume that $\boldsymbol{f}_\gamma \leq \boldsymbol{f}_\nu$ if $\gamma \leq \nu, \ \gamma, \nu \in \Gamma$. In the next sections, we shall see how a parametrized family of curves $\{\boldsymbol{f}_\gamma \ : \ \gamma \in \Gamma\}$ can be used to describe the rate of convergence of sequences with double indices. For short, a family $\{\boldsymbol{f}_\gamma \ : \ \gamma \in \Gamma\}$ will be called a rate.

Example 2.2.1. In this example, we show how to build up a curve $\boldsymbol{f}_{ab}$ (see figure (2.3)) using f_{ab}^1 (see figure (2.1)) and f_{ba}^1 (see figure (2.2)). Then we visualize the points (i, j) with $(i, j) \geq \boldsymbol{f}_{ab}$ on figure (2.4).

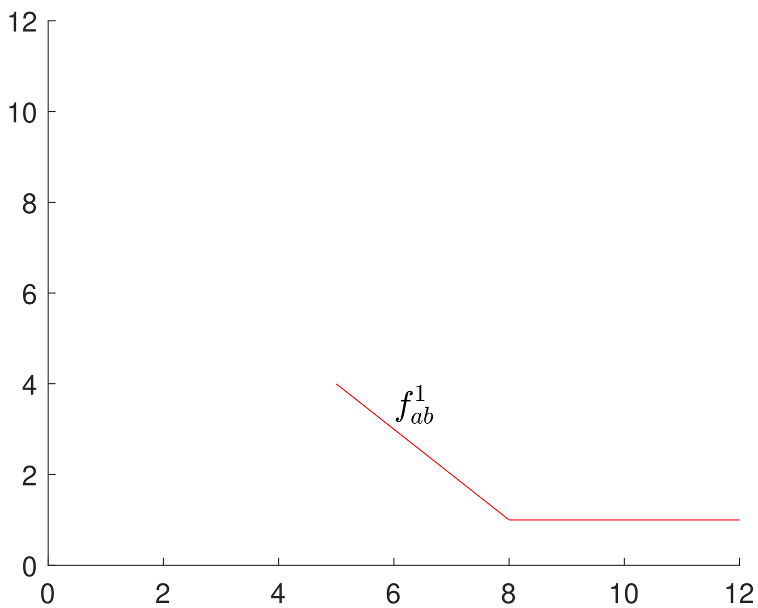

Figure 2.1: Function f_{ab}^1

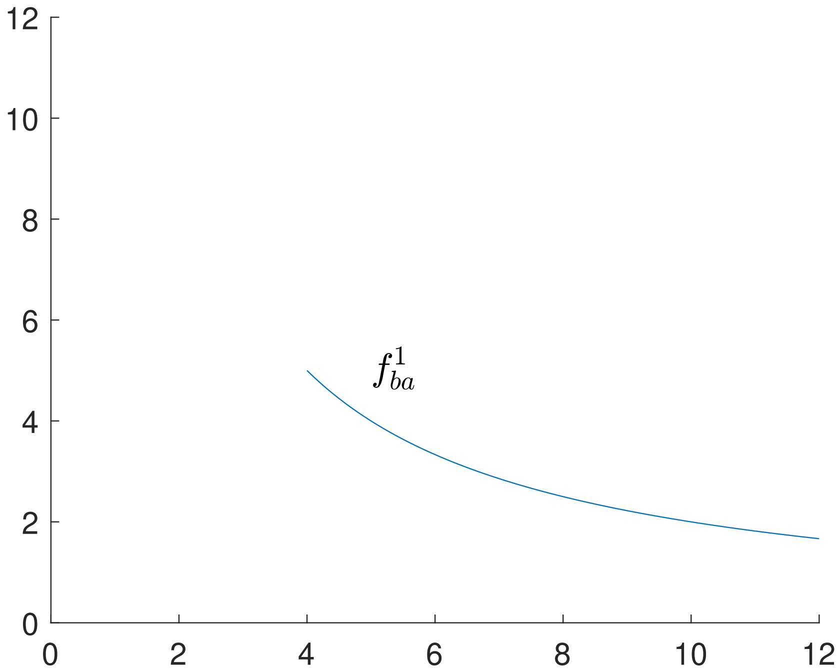

Figure 2.2: Function f_{ba}^1

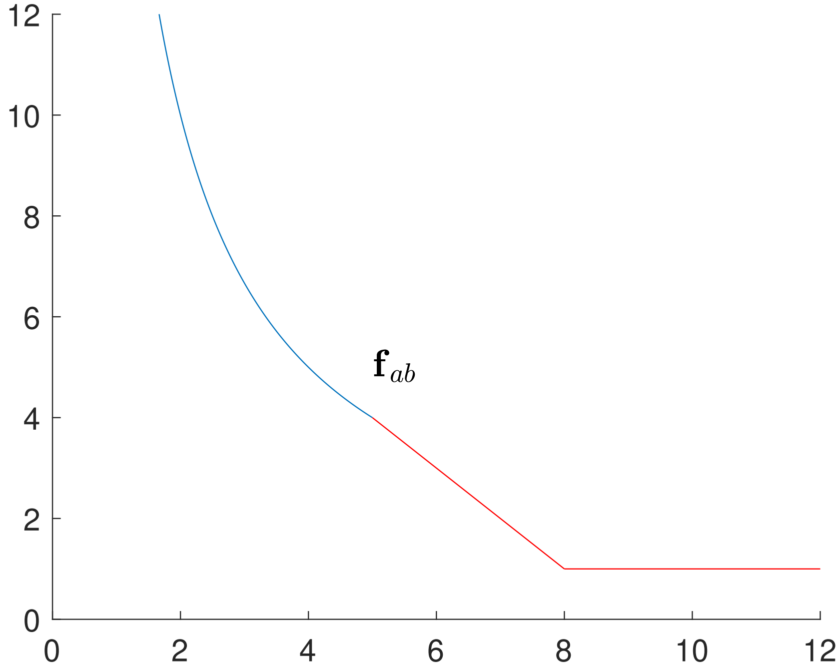

Figure 2.3: The function $\boldsymbol{f}_{ab}$

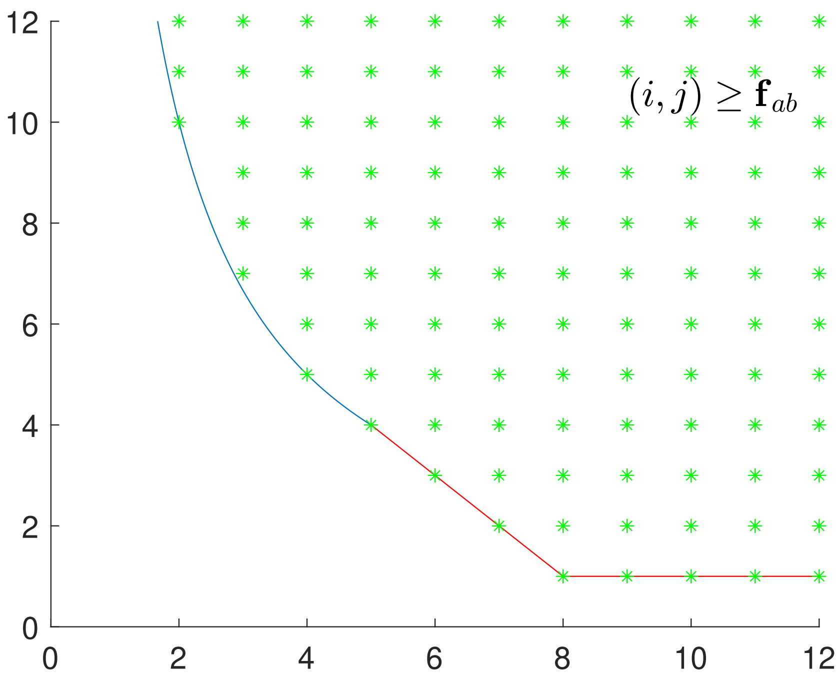

Figure 2.4: The points being greater than $\boldsymbol{f}_{ab}$

In Section 2.3, we shall apply general rates. In Section 2.4 and Section 2.5, we shall use symmetric rate curves.

2.3 General results

We consider sequences of random variables with double indices. $\boldsymbol{n} = (n_1, n_2) \in \mathbb{N}^2$ will denote the indices. Here $\mathbb{N}$ denotes the set of positive integers. We will denote by $\mathbb{N}_0$ the set of non-negative integers. We say that the double sequence of random variables $\eta_{\boldsymbol{n}}, \boldsymbol{n} \in \mathbb{N}^2$, converges almost surely to η, if $\eta_{n_1,n_2}(\omega) \longrightarrow \eta(\omega)$ as $n_1, n_2 \longrightarrow \infty$ for $\omega \in A$, $\mathbb{P}(A) = 1$. We say that the double sequence of random variables $\eta_{\boldsymbol{n}}, \boldsymbol{n} \in \mathbb{N}^2$, converges almost surely strongly to η, if $\eta_{n_1,n_2}(\omega) \longrightarrow \eta(\omega)$ as $\max\{n_1, n_2\} \longrightarrow \infty$ for $\omega \in A$, $\mathbb{P}(A) = 1$.

Let $\alpha > 1$ and let $\boldsymbol{p} = (p_1, p_2) \in \mathbb{N}_0^2$ be fixed. Define the following set of indices

$$C_{\alpha,\boldsymbol{p}} = \{\boldsymbol{n} : \boldsymbol{n} \in \mathbb{N}^2, \alpha^{p_1} \leq n_1 < \alpha^{p_1+1}, \alpha^{p_2} \leq n_2 < \alpha^{p_2+1}\}. \tag{2.3.1}$$

These sets are (possibly empty) rectangles of integer lattice points.

Let $\boldsymbol{k}^-(\boldsymbol{p}) = \min C_{\alpha,\boldsymbol{p}}$ and $\boldsymbol{k}^+(\boldsymbol{p}) = \max C_{\alpha,\boldsymbol{p}}$, if $C_{\alpha,\boldsymbol{p}} \neq \emptyset$, where min and max is defined coordinate-wise. Then $\boldsymbol{k}^-(\boldsymbol{p}) \le \boldsymbol{n} \le \boldsymbol{k}^+(\boldsymbol{p})$ for any $\boldsymbol{n} \in C_{\alpha,\boldsymbol{p}}$ ($\le$ is defined coordinate-wise). When $C_{\alpha,\boldsymbol{p}} = \emptyset$, let $\boldsymbol{k}^-(\boldsymbol{p}) = \boldsymbol{k}^+(\boldsymbol{p}) = \boldsymbol{0} = (0,0)$. We shall use notation $\boldsymbol{k}^\pm(\boldsymbol{p})$ if a relation is true both for $\boldsymbol{k}^-(\boldsymbol{p})$ and $\boldsymbol{k}^+(\boldsymbol{p})$.

Let $|\boldsymbol{n}| = n_1 \cdot n_2$ if $\boldsymbol{n} = (n_1, n_2)$. We shall use the notation $\boldsymbol{1} = (1,1) \in \mathbb{N}^2$.

Proposition 2.3.1 (Fazekas & Masasila [25]). Let $\{\xi_{\boldsymbol{n}}, \boldsymbol{n} \in \mathbb{N}^2\}$, be non-negative random variables, and let $S_{\boldsymbol{n}} = \sum_{\boldsymbol{k}\le\boldsymbol{n}} \xi_{\boldsymbol{k}}$, $\mathbb{E}\xi_{\boldsymbol{n}} = \mu$ for all $\boldsymbol{n} \in \mathbb{N}^2$. If for each $\alpha > 1$,

$$\sum_{\boldsymbol{p}\in\mathbb{N}_0^2, \boldsymbol{k}^\pm(\boldsymbol{p})\neq\boldsymbol{0}} \mathbb{P}\left(\left| \frac{S_{\boldsymbol{k}^\pm(\boldsymbol{p})}}{|\boldsymbol{k}^\pm(\boldsymbol{p})|} - \mu \right| > \varepsilon \right) < \infty, \quad \text{for all} \quad \varepsilon > 0,$$

then

$$\frac{S_{\boldsymbol{n}}}{|\boldsymbol{n}|} \to \mu \quad \text{a.s. strongly as} \quad \boldsymbol{n} \to \infty. \tag{2.3.2}$$

Proof. By the Borel-Cantelli lemma, only finitely many of the events $\left\{ \left| \frac{S_{\boldsymbol{k}^\pm(\boldsymbol{p})}}{|\boldsymbol{k}^\pm(\boldsymbol{p})|} - \mu \right| > \varepsilon \right\}$ occur almost surely. So

$$\frac{S_{\boldsymbol{k}^\pm(\boldsymbol{p})}}{|\boldsymbol{k}^\pm(\boldsymbol{p})|} \to \mu \quad \text{almost surely as} \quad \max\{p_1, p_2\} \to \infty. \tag{2.3.3}$$

Let $\boldsymbol{m} \in C_{\alpha,\boldsymbol{p}}$. Then

$$\frac{S_{\boldsymbol{m}}}{|\boldsymbol{m}|} - \mu \ge \frac{S_{\boldsymbol{k}^-(\boldsymbol{p})}}{|\boldsymbol{m}|} - \mu \ge \frac{S_{\boldsymbol{k}^-(\boldsymbol{p})}}{\alpha^2 |\boldsymbol{k}^-(\boldsymbol{p})|} - \mu \tag{2.3.4}$$

$$= \left(\frac{S_{\boldsymbol{k}^-(\boldsymbol{p})}}{|\boldsymbol{k}^-(\boldsymbol{p})|} - \mu \right) \frac{1}{\alpha^2} + \mu \left(\frac{1}{\alpha^2} - 1 \right). \tag{2.3.5}$$

Here, the first inequality follows from the fact that $\{S_{\boldsymbol{n}}\}$ is monotone (since $\{\xi_{\boldsymbol{n}}\}$ is non-negative) and $\boldsymbol{k}^-(\boldsymbol{p}) \le \boldsymbol{m}$. The second inequality is due to the fact that $|\boldsymbol{m}| \le \alpha^2 |\boldsymbol{k}^-(\boldsymbol{p})|$ (since $\boldsymbol{m} \in C_{\alpha,\boldsymbol{p}}$, so by definition $\boldsymbol{m} < \alpha^{\boldsymbol{p}+\boldsymbol{1}}$ and $\boldsymbol{k}^-(\boldsymbol{p}) \in C_{\alpha,\boldsymbol{p}}$, so, $\alpha^{\boldsymbol{p}} \le \boldsymbol{k}^-(\boldsymbol{p})$).

Using similar arguments, we have,

$$\frac{S_{\boldsymbol{m}}}{|\boldsymbol{m}|} - \mu \le \frac{S_{\boldsymbol{k}^+(\boldsymbol{p})}}{|\boldsymbol{m}|} - \mu \le \left(\frac{S_{\boldsymbol{k}^+(\boldsymbol{p})}}{|\boldsymbol{k}^+(\boldsymbol{p})|} - \mu \right) \alpha^2 + \mu(\alpha^2 - 1). \tag{2.3.6}$$

Now, using (2.3.3) and taking $\alpha \to 1$ in inequalities (2.3.4)-(2.3.6), we obtain (2.3.2). □

Proposition 2.3.2 (Fazekas & Masasila [25]). Let $\{\xi_{\boldsymbol{n}}, \boldsymbol{n} \in \mathbb{N}^2\}$, be non-negative random variables, and let $S_{\boldsymbol{n}} = \sum_{\boldsymbol{k}\le\boldsymbol{n}} \xi_{\boldsymbol{k}}$, $\mathbb{E}\xi_{\boldsymbol{n}} = \mu$ for all $\boldsymbol{n}$. Assume that for any $\varepsilon > 0$ and any $\alpha > 1$

$$\sum_{\boldsymbol{p}\ge\boldsymbol{l}, \boldsymbol{k}^\pm(\boldsymbol{p})\neq\boldsymbol{0}} \mathbb{P}\left(\left| \frac{S_{\boldsymbol{k}^\pm(\boldsymbol{p})}}{|\boldsymbol{k}^\pm(\boldsymbol{p})|} - \mu \right| > \varepsilon \right) \le \lambda \quad \text{if} \quad \boldsymbol{l} \ge \boldsymbol{\Lambda}_{\varepsilon,\alpha}(\lambda), \tag{2.3.7}$$

where $\{\mathbf{\Lambda}_{\varepsilon,\alpha}(\lambda) : \lambda > 0\}$ is a rate. Then

$$\mathbb{P}\left(\sup_{\boldsymbol{m}\geq\boldsymbol{n}}\left|\frac{S_{\boldsymbol{m}}}{|\boldsymbol{m}|}-\mu\right|>\varepsilon\right)\leq\lambda \quad \text{if} \quad \boldsymbol{n}\geq\mathbf{\Phi}_{\varepsilon,\mathbf{\Lambda},\alpha}(\lambda),$$

where $\mathbf{\Phi}_{\varepsilon,\mathbf{\Lambda},\alpha}(\lambda) = \alpha^{\lceil \mathbf{\Lambda}_{\frac{\varepsilon}{2\alpha^2},\alpha}(\lambda)\rceil}$ is the rate of convergence of $\frac{S_{\boldsymbol{m}}}{|\boldsymbol{m}|} \to \mu$ a.s., and $\alpha^2 = \frac{\varepsilon}{2\mu}+1$. Here $\lceil x\rceil$ denotes the smallest integer being not smaller than x and $\alpha^{\mathbf{\Gamma}} = (\alpha^{\Gamma_1}, \alpha^{\Gamma_2})$ if $\mathbf{\Gamma} = (\Gamma_1, \Gamma_2)$.

Proof. Let $\boldsymbol{m} \geq \boldsymbol{n}$ and assume that

$$\left|\frac{S_{\boldsymbol{m}}}{|\boldsymbol{m}|}-\mu\right|>\varepsilon. \tag{2.3.8}$$

Then $\alpha^{\boldsymbol{p}} \leq \boldsymbol{m} < \alpha^{\boldsymbol{p}+\mathbf{1}}$ for some $\boldsymbol{p} = (p_1, p_2) \in \mathbb{N}_0^2$. So, by (2.3.4) and (2.3.6), either

$$\frac{S_{\boldsymbol{k}^+(\boldsymbol{p})}}{|\boldsymbol{k}^+(\boldsymbol{p})|}\alpha^2-\mu\alpha^2+\mu(\alpha^2-1)>\varepsilon$$

or

$$\left(\frac{S_{\boldsymbol{k}^-(\boldsymbol{p})}}{|\boldsymbol{k}^-(\boldsymbol{p})|}-\mu\right)\frac{1}{\alpha^2}+\mu\left(\frac{1}{\alpha^2}-1\right)<-\varepsilon.$$

That is either

$$\frac{S_{\boldsymbol{k}^+(\boldsymbol{p})}}{|\boldsymbol{k}^+(\boldsymbol{p})|}-\mu>\frac{\varepsilon-\mu(\alpha^2-1)}{\alpha^2} \tag{2.3.9}$$

or

$$\frac{S_{\boldsymbol{k}^-(\boldsymbol{p})}}{|\boldsymbol{k}^-(\boldsymbol{p})|}-\mu<-\left(\varepsilon-\mu\left(1-\frac{1}{\alpha^2}\right)\right)\alpha^2. \tag{2.3.10}$$

Now, let $\mu(\alpha^2-1) = \frac{\varepsilon}{2}$, that is $\alpha^2 = \frac{\varepsilon}{2\mu}+1$. Then (2.3.9) is equivalent to

$$\frac{S_{\boldsymbol{k}^+(\boldsymbol{p})}}{|\boldsymbol{k}^+(\boldsymbol{p})|}-\mu>\frac{\varepsilon}{2\alpha^2}. \tag{2.3.11}$$

Moreover, (2.3.10) implies that

$$\frac{S_{\boldsymbol{k}^-(\boldsymbol{p})}}{|\boldsymbol{k}^-(\boldsymbol{p})|}-\mu<-\left(\varepsilon-\mu\left(1-\frac{1}{\alpha^2}\right)\right)\alpha^2 \tag{2.3.12}$$

$$=-\varepsilon\alpha^2+\frac{\varepsilon}{2}\leq\frac{-\varepsilon}{2\alpha^2} \tag{2.3.13}$$

as $\alpha > 1$. So (2.3.8) implies that either

$$\left|\frac{S_{\boldsymbol{k}^+(\boldsymbol{p})}}{|\boldsymbol{k}^+(\boldsymbol{p})|} - \mu\right| > \frac{\varepsilon}{2\alpha^2}$$

or

$$\left|\frac{S_{\boldsymbol{k}^-(\boldsymbol{p})}}{|\boldsymbol{k}^-(\boldsymbol{p})|} - \mu\right| > \frac{\varepsilon}{2\alpha^2}.$$

By (2.3.7), the total probability of these events is smaller than λ if $\boldsymbol{l} \geq \boldsymbol{\Lambda}_{\frac{\varepsilon}{2\alpha^2},\alpha}(\lambda)$, that is if $\boldsymbol{n} \geq \alpha^{\lceil \boldsymbol{\Lambda}_{\frac{\varepsilon}{2\alpha^2},\alpha}(\lambda)\rceil} = \boldsymbol{\Phi}_{\varepsilon,\boldsymbol{\Lambda},\alpha}(\lambda)$. □

2.4 Pairwise independent random variables

Lemma 2.4.1 (Fazekas & Masasila [25]). Let $\{\xi_{\boldsymbol{k}}, \boldsymbol{k} \in \mathbb{N}^2\}$, be pairwise independent random variables with $\mathbb{E}\xi_{\boldsymbol{k}} = \mu$ and $\mathrm{Var}\xi_{\boldsymbol{k}} \leq \sigma^2$ for all $\boldsymbol{k} \in \mathbb{N}^2$, $S_{\boldsymbol{n}} = \sum_{\boldsymbol{k}\leq\boldsymbol{n}} \xi_{\boldsymbol{k}}$. Let $\alpha > 1$. Then

$$\sum_{\boldsymbol{n}\geq\boldsymbol{p},\boldsymbol{k}^\pm(\boldsymbol{n})\neq\boldsymbol{0}} \mathbb{P}\left(\left|\frac{S_{\boldsymbol{k}^\pm(\boldsymbol{n})}}{|\boldsymbol{k}^\pm(\boldsymbol{n})|} - \mu\right| > \varepsilon\right) \leq \lambda, \tag{2.4.1}$$

if $p_1 + p_2 \geq \log_\alpha(\frac{\sigma^2\alpha^2}{\varepsilon^2\lambda(\alpha-1)^2}) = \rho_{\varepsilon,\alpha}(\lambda)$, that is the rate of convergence is given by $\boldsymbol{\Lambda}_{\varepsilon,\alpha}(\lambda)$, where $\boldsymbol{\Lambda}_{\varepsilon,\alpha}(\lambda)$ is determined by the curve $x + y = \rho_{\varepsilon,\alpha}(\lambda)$.

Proof.

$$\begin{aligned}
\sum_{\boldsymbol{n}\geq\boldsymbol{p},\boldsymbol{k}^\pm(\boldsymbol{n})\neq\boldsymbol{0}} \mathbb{P}\left(\left|\frac{S_{\boldsymbol{k}^\pm(\boldsymbol{n})}}{|\boldsymbol{k}^\pm(\boldsymbol{n})|} - \mu\right| > \varepsilon\right) &\leq \frac{1}{\varepsilon^2} \sum_{\boldsymbol{n}\geq\boldsymbol{p},\boldsymbol{k}^\pm(\boldsymbol{n})\neq\boldsymbol{0}} \frac{\mathrm{Var}(S_{\boldsymbol{k}^\pm(\boldsymbol{n})})}{|\boldsymbol{k}^\pm(\boldsymbol{n})|^2} \\
&= \frac{1}{\varepsilon^2} \sum_{\boldsymbol{n}\geq\boldsymbol{p},\boldsymbol{k}^\pm(\boldsymbol{n})\neq\boldsymbol{0}} \frac{\sum_{\boldsymbol{k}\leq\boldsymbol{k}^\pm(\boldsymbol{n})} \mathrm{Var}(\xi_{\boldsymbol{k}})}{|\boldsymbol{k}^\pm(\boldsymbol{n})|^2} \\
&\leq \frac{\sigma^2}{\varepsilon^2} \sum_{\boldsymbol{n}\geq\boldsymbol{p},\boldsymbol{k}^\pm(\boldsymbol{n})\neq\boldsymbol{0}} \frac{|\boldsymbol{k}^\pm(\boldsymbol{n})|}{|\boldsymbol{k}^\pm(\boldsymbol{n})|^2} \\
&\leq \frac{\sigma^2}{\varepsilon^2} \sum_{\boldsymbol{n}\geq\boldsymbol{p},\boldsymbol{k}^\pm(\boldsymbol{n})\neq\boldsymbol{0}} \alpha^{-\boldsymbol{n}} \\
&\leq \frac{\sigma^2}{\varepsilon^2} \sum_{n_1=p_1}^{\infty} \alpha^{-n_1} \sum_{n_2=p_2}^{\infty} \alpha^{-n_2} \\
&= \frac{\sigma^2}{\varepsilon^2} \cdot \frac{\alpha^{-p_1}}{1-\frac{1}{\alpha}} \cdot \frac{\alpha^{-p_2}}{1-\frac{1}{\alpha}} \\
&= \frac{\sigma^2}{\varepsilon^2} \cdot \frac{\alpha^{-p_1+1}\alpha^{-p_2+1}}{(\alpha-1)^2} \leq \lambda,
\end{aligned}$$

if $\alpha^{-(p_1+p_2)} \frac{\sigma^2}{\varepsilon^2} \frac{\alpha^2}{(\alpha-1)^2} \leq \lambda$.

Here, the first inequality follows from Chebyshev's inequality. The second inequality is derived using pairwise independence and the assumption that $\mathrm{Var}\xi_{\boldsymbol{k}} \leq \sigma^2$. The third inequality is satisfied by the condition $(\boldsymbol{k}^{\pm}(\boldsymbol{n}))_i \geq \alpha^{n_i}$ for $i = 1, 2$. The remaining steps are obtained through straightforward calculations with the application of the formula for the sum of an infinite geometric series.

The last inequality is equivalent to

$$\alpha^{p_1+p_2} \geq \frac{\sigma^2\alpha^2}{\lambda\varepsilon^2(\alpha-1)^2}$$

or

$$p_1 + p_2 \geq \log_\alpha\left(\frac{\sigma^2\alpha^2}{\lambda\varepsilon^2(\alpha-1)^2}\right).$$

□

Lemma 2.4.2 (Fazekas & Masasila [25]). Let $\{\xi_{\boldsymbol{k}}, \boldsymbol{k} \in \mathbb{N}^2\}$, be pairwise independent non-negative random variables with $\mathbb{E}\xi_{\boldsymbol{k}} = \mu$ and $\mathrm{Var}\xi_{\boldsymbol{k}} \leq \sigma^2$ for all $\boldsymbol{k} \in \mathbb{N}^2$. Let $S_{\boldsymbol{n}} = \sum_{\boldsymbol{k}\leq\boldsymbol{n}} \xi_{\boldsymbol{k}}$, $\alpha^2 = \frac{\varepsilon}{2\mu} + 1$. Then for all $\varepsilon > 0$, $\lambda > 0$,

$$\mathbb{P}\left(\sup_{\boldsymbol{m}\geq\boldsymbol{n}}\left|\frac{S_{\boldsymbol{m}}}{|\boldsymbol{m}|} - \mu\right| > \varepsilon\right) \leq \lambda, \tag{2.4.2}$$

if $\boldsymbol{n} \geq \boldsymbol{\Phi}_{\varepsilon,\boldsymbol{\Lambda}}(\lambda)$, where $\boldsymbol{\Phi}_{\varepsilon,\boldsymbol{\Lambda}}(\lambda) = \alpha^{\lceil \boldsymbol{\Lambda}_{\frac{\varepsilon}{2\alpha^2},\alpha}(\lambda)\rceil}$ and $\boldsymbol{\Lambda}_{\varepsilon,\alpha}(\lambda)$ is determined by the curve $x + y = \rho_{\varepsilon,\alpha}(\lambda)$ with $\rho_{\varepsilon,\alpha}(\lambda) = \log_\alpha\left(\frac{\sigma^2\alpha^2}{\lambda\varepsilon^2(\alpha-1)^2}\right)$. Inequality (2.4.2) is satisfied if $n_1 \cdot n_2 \geq \frac{4\sigma^2\alpha^8}{\lambda\varepsilon^2(\alpha-1)^2}$.

Proof. By Lemma 2.4.1,

$$\sum_{\boldsymbol{n}\geq\boldsymbol{p},\boldsymbol{k}^{\pm}(\boldsymbol{n})\neq\boldsymbol{0}} \mathbb{P}\left(\left|\frac{S_{\boldsymbol{k}^{\pm}(\boldsymbol{n})}}{|\boldsymbol{k}^{\pm}(\boldsymbol{n})|} - \mu\right| > \varepsilon\right) \leq \lambda,$$

if $p_1 + p_2 \geq \log_\alpha(\frac{\sigma^2\alpha^2}{\varepsilon^2\lambda(\alpha-1)^2})$. Therefore, by Proposition 2.3.2,

$$\mathbb{P}\left(\sup_{\boldsymbol{m}\geq\boldsymbol{n}}\left|\frac{S_{\boldsymbol{m}}}{|\boldsymbol{m}|} - \mu\right| > \varepsilon\right) \leq \lambda \quad \text{if} \quad \boldsymbol{n} \geq \boldsymbol{\Phi}_{\varepsilon,\boldsymbol{\Lambda},\alpha}(\lambda) = \alpha^{\lceil \boldsymbol{\Lambda}_{\frac{\varepsilon}{2\alpha^2},\alpha}(\lambda)\rceil}.$$

We see that the above inequality is satisfied if

$$n_1 \cdot n_2 \geq \alpha^2 \cdot \frac{\sigma^2\alpha^2}{\lambda(\frac{\varepsilon}{2\alpha^2})^2(\alpha-1)^2} = \frac{4\sigma^2\alpha^8}{\lambda\varepsilon^2(\alpha-1)^2}.$$

□

Remark 2.4.3. We visualize the rates in Lemma 2.4.1 and Lemma 2.4.2 using figures. On figure (2.5), we show a rate curve in Lemma 2.4.1. That Curve $\mathbf{\Lambda}_{\varepsilon,\alpha}(\lambda)$ is determined by a part of the x axis, a part of the y axis and the section of the line $x + y = \rho_{\varepsilon,\alpha}(\lambda)$ being in the first quadrant. On figure (2.6), we show a sequence of rate curves in Lemma 2.4.2. These curves are of the shape $\mathbf{\Phi}_{\varepsilon,\mathbf{\Lambda}}(\lambda) = \alpha^{\lceil \mathbf{\Lambda}_{\frac{\varepsilon}{2\alpha^2},\alpha}(\lambda)\rceil}$. In explicit form these curves are given by the formula $x \cdot y = \frac{4\sigma^2\alpha^8}{\lambda\varepsilon^2(\alpha-1)^2}$.

Proposition 2.4.4 (Fazekas & Masasila [25]). Let $\{\xi_{\boldsymbol{n}}, \boldsymbol{n} \in \mathbb{N}^2\}$, be pairwise independent random variables with $\mathbb{E}\xi_{\boldsymbol{n}} = 0$, $\mathbb{E}|\xi_{\boldsymbol{n}}| = \tau > 0$ and $\mathrm{Var}\xi_{\boldsymbol{n}} \le \sigma^2 > 0$ for all $\boldsymbol{n} \in \mathbb{N}^2$. Let $S_{\boldsymbol{n}} = \sum_{\boldsymbol{k}\le\boldsymbol{n}} \xi_{\boldsymbol{k}}$. Then for all $\varepsilon > 0$, $\lambda > 0$,

$$\mathbb{P}\left(\sup_{\boldsymbol{m}\ge\boldsymbol{n}} \frac{|S_{\boldsymbol{m}}|}{|\boldsymbol{m}|} > \varepsilon\right) \le \lambda \quad \text{if} \quad |\boldsymbol{n}| \ge \frac{32\sigma^2\alpha^8}{\lambda\varepsilon^2(\alpha-1)^2} \quad \text{and} \quad \alpha^2 = \frac{\varepsilon}{2\tau} + 1.$$

Proof. For the positive and the negative parts, we have $\mathrm{Var}\xi_{\boldsymbol{n}}^+ + \mathrm{Var}\xi_{\boldsymbol{n}}^- \le \mathbb{E}(\xi_{\boldsymbol{n}}^+)^2 + \mathbb{E}(\xi_{\boldsymbol{n}}^-)^2 = \mathbb{E}\xi_n^2 = \mathrm{Var}\xi_{\boldsymbol{n}} \le \sigma^2$, therefore $\mathrm{Var}\xi_{\boldsymbol{n}}^+ \le \sigma^2$, $\mathrm{Var}\xi_{\boldsymbol{n}}^- \le \sigma^2$. Moreover, $\mathbb{E}\xi_{\boldsymbol{n}}^+ = \mathbb{E}\xi_{\boldsymbol{n}}^- = \frac{\tau}{2} = \mu$. By Lemma 2.4.2,

$$\mathbb{P}\left(\sup_{\boldsymbol{m}\ge\boldsymbol{n}} \left|\frac{S_{\boldsymbol{m}}^{\pm}}{|\boldsymbol{m}|} - \mu\right| > \frac{\varepsilon}{2}\right) \le \frac{\lambda}{2} \quad \text{if} \quad |\boldsymbol{n}| \ge \frac{4\sigma^2\alpha^8}{(\frac{\lambda}{2})(\frac{\varepsilon}{2})^2(\alpha-1)^2},$$

where $\alpha^2 = \frac{\varepsilon}{2} \cdot \frac{1}{2\mu} + 1 = \frac{\varepsilon}{2\tau} + 1$. Then

$$\begin{aligned}\mathbb{P}\left(\sup_{\boldsymbol{m}\ge\boldsymbol{n}} \frac{|S_{\boldsymbol{m}}|}{|\boldsymbol{m}|} > \varepsilon\right)& \\ &\le \mathbb{P}\left(\sup_{\boldsymbol{m}\ge\boldsymbol{n}} \left|\frac{S_{\boldsymbol{m}}^{+}}{|\boldsymbol{m}|} - \mu\right| > \frac{\varepsilon}{2}\right) + \mathbb{P}\left(\sup_{\boldsymbol{m}\ge\boldsymbol{n}} \left|\frac{S_{\boldsymbol{m}}^{-}}{|\boldsymbol{m}|} - \mu\right| > \frac{\varepsilon}{2}\right) \\ &\le \frac{\lambda}{2} + \frac{\lambda}{2} = \lambda,\end{aligned}$$

if $|\boldsymbol{n}| \ge \frac{32\sigma^2\alpha^8}{\lambda\varepsilon^2(\alpha-1)^2}$. □

Lemma 2.4.5 (Fazekas & Masasila [25]). Let $\tau > 0$ and $\alpha^2 = \frac{\varepsilon}{2\tau} + 1$ be the values from Proposition 2.4.4. Then for any $C > 4$, we have

$$\frac{\alpha^8}{(\alpha-1)^2} \le \frac{4\tau^2}{\varepsilon^2} C \tag{2.4.3}$$

for small enough $\varepsilon > 0$. More precisely, for any $C > 4$ there exists an $a \in (0, 1/2)$ for which $\left(\frac{(1-a)^4}{a^5}\right)^2 = C$, and with $b = (1-2a)/a^2$ the inequality (2.4.3) is satisfied for $0 < \varepsilon < 2\tau b$.

Proof. Analysing the function $f(x) = \sqrt{x+1} - 1$, $x \geq 0$, and the straight line ax, where $a \in (0, 1/2)$ is a parameter, we get the following.

$$ax \leq \sqrt{x+1} - 1 \qquad \text{if} \qquad 0 \leq x \leq b,$$

where $b = (1-2a)/a^2$ and a is a fixed value from the interval $(0, 1/2)$.

Now, let $\alpha^2 = \frac{\varepsilon}{2\tau} + 1$ be from Proposition 2.4.4, and apply the above inequality with $x = \frac{\varepsilon}{2\tau}$. Then

$$\frac{\alpha^8}{(\alpha-1)^2} = \frac{\alpha^8}{\left(\sqrt{\frac{\varepsilon}{2\tau}+1} - 1\right)^2} \leq \frac{\alpha^8}{\left(a\frac{\varepsilon}{2\tau}\right)^2} \leq \frac{4\tau^2\alpha^8}{a^2\varepsilon^2}.$$

As now $\frac{\varepsilon}{2\tau} = x \leq b$, we have

$$\frac{\alpha^8}{(\alpha-1)^2} \leq \frac{4\tau^2}{\varepsilon^2}\frac{(b+1)^4}{a^2} = \frac{4\tau^2}{\varepsilon^2}\frac{\left(\frac{1-2a}{a^2}+1\right)^4}{a^2} = \frac{4\tau^2}{\varepsilon^2}C,$$

where $C = \left(\frac{(1-a)^4}{a^5}\right)^2$. Now, we consider the function $g(a) = \frac{(1-a)^4}{a^5}$ for $a \in (0, 1/2)$. Its infimum is 2. So for any given $C > 4$ we can find an appropriate a so that $C = \left(\frac{(1-a)^4}{a^5}\right)^2$. □

Theorem 2.4.6 (Fazekas & Masasila [25]). *Let $\{\xi_{\boldsymbol{n}}, \boldsymbol{n} \in \mathbb{N}^2\}$, be pairwise independent random variables with $\mathbb{E}\xi_{\boldsymbol{n}} = 0$, $\mathbb{E}|\xi_{\boldsymbol{n}}| = \tau > 0$ and $\operatorname{Var}\xi_{\boldsymbol{n}} \leq \sigma^2 > 0$ for all $\boldsymbol{n} \in \mathbb{N}^2$. Let $S_{\boldsymbol{n}} = \sum_{\boldsymbol{k}\leq\boldsymbol{n}} \xi_{\boldsymbol{k}}$. Then*

$$\mathbb{P}\left(\sup_{\boldsymbol{m}\geq\boldsymbol{n}} \frac{|S_{\boldsymbol{m}}|}{|\boldsymbol{m}|} > \varepsilon\right) \leq \frac{K\sigma^2\tau^2}{\varepsilon^4|\boldsymbol{n}|} \tag{2.4.4}$$

if $K > 512$ and $\varepsilon > 0$ is small enough.

Proof. By Proposition 2.4.4, we have

$$\mathbb{P}\left(\sup_{\boldsymbol{m}\geq\boldsymbol{n}} \frac{|S_{\boldsymbol{m}}|}{|\boldsymbol{m}|} > \varepsilon\right) \leq \frac{32\sigma^2\alpha^8}{\varepsilon^2|\boldsymbol{n}|(\alpha-1)^2}.$$

Applying Lemma 2.4.5, we obtain

$$\mathbb{P}\left(\sup_{\boldsymbol{m}\geq\boldsymbol{n}} \frac{|S_{\boldsymbol{m}}|}{|\boldsymbol{m}|} > \varepsilon\right) \leq \frac{32\sigma^2}{\varepsilon^2|\boldsymbol{n}|}\frac{4\tau^2}{\varepsilon^2}C \leq \frac{K\sigma^2\tau^2}{\varepsilon^4|\boldsymbol{n}|}.$$

□

2.5 Quasi uncorrelated random variables

Definition 2.5.1. A sequence $\xi_{\boldsymbol{n}}, \boldsymbol{n} \in \mathbb{N}^2$, of random variables is said to be quasi uncorrelated if each $\xi_{\boldsymbol{n}}, \boldsymbol{n} \in \mathbb{N}^2$, has finite variance and there exists a positive constant c such that

$$\mathrm{Var}(S_{\boldsymbol{n}}) \le c \sum_{\boldsymbol{k} \le \boldsymbol{n}} \mathrm{Var}(\xi_{\boldsymbol{k}}) \tag{2.5.1}$$

for any $\boldsymbol{n} \in \mathbb{N}^2$, see [37].

We can see that the results for pairwise independent random variables can be adapted to the case of quasi uncorrelated random variables. So we present them without proofs. Next proposition is a version of Proposition 2.4.4 for quasi uncorrelated random variables.

Proposition 2.5.2 (Fazekas & Masasila [25]). Let $\{\xi_{\boldsymbol{n}}, \boldsymbol{n} \in \mathbb{N}^2\}$, be quasi uncorrelated random variables with $\mathbb{E}\xi_{\boldsymbol{n}} = 0$, $\mathbb{E}|\xi_{\boldsymbol{n}}| = \tau > 0$ and $\mathrm{Var}\xi_{\boldsymbol{n}} \le \sigma^2 > 0$ for all $\boldsymbol{n} \in \mathbb{N}^2$. Let $S_{\boldsymbol{n}} = \sum_{\boldsymbol{k} \le \boldsymbol{n}} \xi_{\boldsymbol{k}}$. Then for all $\varepsilon > 0$, $\lambda > 0$,

$$\mathbb{P}\left(\sup_{\boldsymbol{m} \ge \boldsymbol{n}} \frac{|S_{\boldsymbol{m}}|}{|\boldsymbol{m}|} > \varepsilon\right) < \lambda \quad \text{if} \quad |\boldsymbol{n}| \ge \frac{c 32 \sigma^2 \alpha^8}{\lambda \varepsilon^2 (\alpha - 1)^2} \quad \text{and} \quad \alpha^2 = \frac{\varepsilon}{2\tau} + 1,$$

where c is from (2.5.1).

Now, we turn to the quasi uncorrelated version of Theorem 2.4.6.

Theorem 2.5.3. *Let* $\{\xi_{\boldsymbol{n}}, \boldsymbol{n} \in \mathbb{N}^2\}$, *be quasi uncorrelated random variables with* $\mathbb{E}\xi_{\boldsymbol{n}} = 0$, $\mathbb{E}|\xi_{\boldsymbol{n}}| = \tau > 0$ *and* $\mathrm{Var}\xi_{\boldsymbol{n}} \le \sigma^2 > 0$ *for all* $\boldsymbol{n} \in \mathbb{N}^2$. *Let* $S_{\boldsymbol{n}} = \sum_{\boldsymbol{k} \le \boldsymbol{n}} \xi_{\boldsymbol{k}}$. *Then with* c *from* (2.5.1), *we have*

$$\mathbb{P}\left(\sup_{\boldsymbol{m} \ge \boldsymbol{n}} \frac{|S_{\boldsymbol{m}}|}{|\boldsymbol{m}|} > \varepsilon\right) \le \frac{K c \sigma^2 \tau^2}{\varepsilon^4 |\boldsymbol{n}|} \tag{2.5.2}$$

if $K > 512$ *and* $\varepsilon > 0$ *is small enough.*

Now, we turn to negatively dependent random variables. Random variables ξ and η are called negatively dependent (ND) if

$$\mathbb{P}(\xi \le x, \eta \le y) \le \mathbb{P}(\xi \le x)\mathbb{P}(\eta \le y)$$

for all real numbers x and y. A set of random variables is said to be pairwise ND if every pair of random variables in the set is ND.

Remark 2.5.4. For pairwise ND random variables Proposition 2.5.2 and Theorem 2.5.3 are true with $c = 1$. For the proof, we just remark that in a pairwise ND set, for two different random variables ξ and η, we have $\mathrm{Cov}(\xi, \eta) \le 0$, see e.g. [72]. So an ND set of random variables is quasi uncorrelated with $c = 1$.

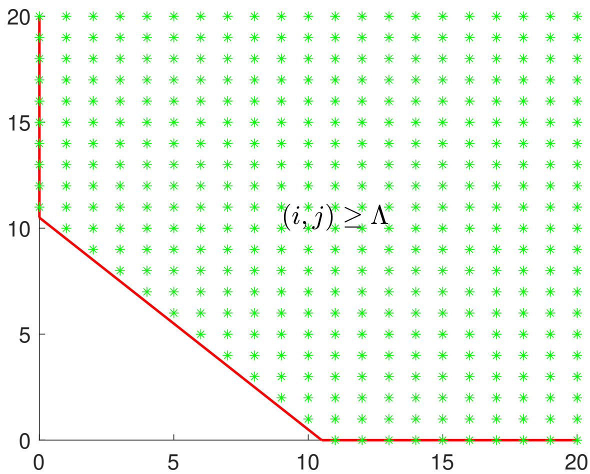

Figure 2.5: A curve Λ in Lemma 2.4.1

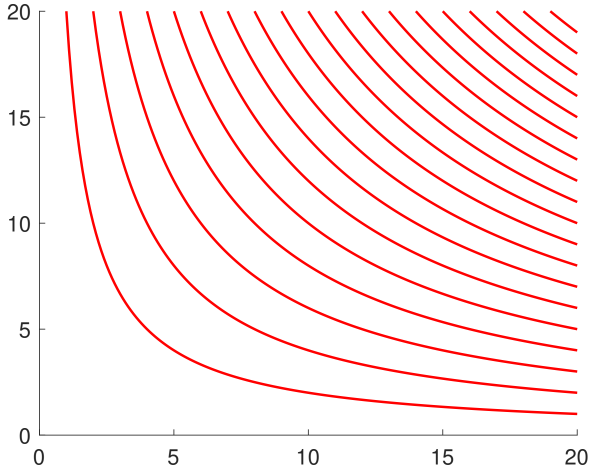

Figure 2.6: Curves in Lemma 2.4.2

Chapter 3

Strong Laws of Large Numbers for General Random Variables under Conditional Sub-additive Expectation and Capacity

In this chapter, we study strong laws of large numbers in a non-linear framework based on conditional sub-additive expectations and conditional sub-additive capacities. Using an axiomatic approach to conditional sub-additive expectation, we establish a conditional Hájek-Rényi-type maximal inequality assuming a general conditional Kolmogorov-type maximal inequality but without imposing any weak dependence structure on the underlying sequence. As a consequence, we derive a general conditional strong law of large numbers. Finally, we introduce a notion of conditional negative dependence under sub-additive expectations and obtain the corresponding conditional Kolmogorov-type maximal inequality, leading to a conditional strong law of large numbers for conditionally negatively dependent random variables.

3.1 Introduction

The Strong Laws of Large Numbers (SLLN) is a fundamental result in probability theory that guarantees that sample averages converge to the expected value under appropriate assumptions. However, in many complex applications, the standard assumptions of additivity of probability measures and linearity of expectations may fail. To address such situations, alternative probabilistic frameworks have been developed, particularly those involving sub-additive probabilities (capacities) and sub-additive expectations.

One of the most famous results in probability theory is the Kolmogorov's SLLN. The original proof of this SLLN is based on the well-known Kolmogorov's maximal inequality.

Another essential tool for proving the SLLN is the Hájek-Rényi inequality (see [30]). Both the Kolmogorov and the Hájek-Rényi inequalities have numerous extensions to non-independent random variables.

Fazekas and Klesov in [23] presented a general approach to SLLNs for sequences of possibly non-independent random variables. They proved, that a Kolmogorov-type inequality implies a Hájek-Rényi-type inequality, which in turn implies an SLLN directly; see also [22]. Their method imposes no restriction on the underlying dependence structure of the random variables. Therefore, the approach of [23] was applied and extended by several authors, see, e.g., [34].

Conditional versions of numerous SLLNs were obtained, see, e.g., [46]. The conditional versions of some results of [23] were obtained in [24]. Another way to extend the scope of the usual SLLNs is to obtain their appropriate version for sub-additive probabilities and for sub-additive expectations. Huang and Wu in [35] employed the method of [23] to obtain a general SLLN of the form (3.5.2) in sub-additive expectation spaces that does not require independence of the random variables. A novel approach is to obtain SLLNs for conditional sub-additive probabilities and conditional sub-additive expectations; see [78].

In this chapter, we want to extend some results of [23] and [22] to conditional sub-additive expectations and conditional sub-additive probabilities. To this end, we shall find some plausible axioms of conditional sub-additive expectations and conditional sub-additive probabilities that guarantee a certain general SLLN. We aim to understand which assumptions are necessary to prove the SLLN and which are superfluous.

In the setting of sub-additive expectation and sub-additive probability, the Kolmogorov-type SLLN takes a generalized form: it asserts that every cluster point of the sequence of empirical averages lies between the lower and upper expectations, with lower capacity equal to 1. Formally,

$$v\Big(\mathcal{E}[X_1] \le \liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k \le \limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k \le \hat{\mathbb{E}}[X_1]\Big) = 1. \tag{3.1.1}$$

Here and in what follows $\hat{\mathbb{E}}[X]$ is the upper expectation (sub-linear expectation) of X, $\mathcal{E}[X] = -\hat{\mathbb{E}}[-X]$ is the lower expectation, $\hat{\mathbb{V}}(A)$ is the upper probability (upper capacity) of A, $v(A) = 1 - \hat{\mathbb{V}}(A^c)$ is the lower probability (lower capacity), where A^c is the complementary set of A, see [11], [8], [58] and Section 3.2 of this paper. For the precise details of the SLLN in (3.1.1), see [45], [8], and [9].

Chen [8] summarized key properties and lemmas concerning upper and lower probabilities (capacities), providing a basis for limit theorems under non-additive measures. Then, in [8], an SLLN was proved for independent random variables having uniformly bounded second moments. Subsequently, Chen, Wu, and Li [9] established a strong law of large numbers for independent random variables having uniformly bounded $1 + \alpha$ moments ($\alpha > 0$) under non-additive probabilities. More recently, Zhang, Tang, and Xiong [78] extended the results of [9] to the G-expectation framework by proving conditional SLLNs, using the theory of conditional G-expectations stated in Hu and Peng [33]. A main result of [78] is of the shape

$$\hat{\mathbb{V}}\left(\limsup_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}X_i=\hat{\mathbb{E}}[X_1|\mathcal{F}]\right)=1, \tag{3.1.2}$$

where the random variables X_i are conditionally independent and identically distributed having finite $1+\alpha$ moments ($\alpha>0$). The result of [78] can be viewed as a conditional version of the non-additive SLLN obtained in [8] and [9].

A well-established constructive framework for sub-additive probability and expectation, and their conditional versions, is given by the theory of G-Brownian motion. This theory was introduced and developed by Peng and co-authors; see, e.g., [56] and [58]. This framework has drawn considerable interest from the research community, leading to numerous contributions that have advanced this field. Zhang, Tang, and Xiong used this framework in [78] to obtain conditional versions of the non-additive SLLN. Another starting point for conditional non-linear expectation and probability is the use of several probability measures on the same space. Then the upper probability and upper expectation are defined as the supremum of probabilities and expectations, respectively. To define conditional upper probability and upper expectation, one can apply the essential supremum. In this paper, we do not use these two approaches directly. Instead, we summarize specific properties and establish them as axioms. Then, based on the axioms of the sub-additive probability and expectation, and their conditional versions, we prove our results. We emphasize that the usual probability and expectation are additive, which is a significant difference between the classical and sub-additive frameworks.

This chapter is organized as follows. Section 3.2 introduces some basic concepts about sub-additive probabilities and expectations. Building on the ideas of several previous papers, we introduce the axioms we need in the subsequent sections. In Section 3.3, we fix our axioms of conditional sub-additive probabilities and expectations. In Section 3.4, we prove the fundamental inequalities. We show that a Kolmogorov-type inequality implies a Hájek-Rényi-type inequality in the conditional non-linear setting. Here, we follow the ideas of [23] and [22], see also [24] and [35]. In Section 3.5, we prove our SLLN in the conditional non-linear setting. It shows that a Kolmogorov-type maximal inequality implies the SLLN directly. Section 3.6 gives an application to conditional negatively dependent random variables.

3.2 Sub-additive probability and expectation

In this section, we present the fundamental concepts and results of sub-additive expectation spaces, which provide the framework for conditional sub-additive expectations. Let Ω be a non-empty set and let $\mathcal{A}$ be a σ-algebra of subsets of Ω. Ω is the sample space, $\mathcal{A}$ is the family of events. Let $\hat{\mathbb{V}}$ be a real-valued function on $\mathcal{A}$ which we call sub-additive probability (upper probability or capacity).

Here, we list the properties of the sub-additive probability.

Definition 3.2.1. $\hat{\mathbb{V}}$ is called a sub-additive probability if it satisfies the following properties.

1. $\hat{\mathbb{V}}(\Omega) = 1$, $\hat{\mathbb{V}}(\emptyset) = 0$.

2. (Monotonicity.) If $A \subset B$, then $\hat{\mathbb{V}}(A) \leq \hat{\mathbb{V}}(B)$ for any $A, B \in \mathcal{A}$.

3. (Sub-additivity.) If $A, B \in \mathcal{A}$, then

$$\hat{\mathbb{V}}(A \cup B) \leq \hat{\mathbb{V}}(A) + \hat{\mathbb{V}}(B).$$

4. (Lower-continuity.) If $A_n \uparrow A$, $A_n \in \mathcal{A}$, then $\hat{\mathbb{V}}(A_n) \uparrow \hat{\mathbb{V}}(A)$.

We mention that sub-additivity and lower-continuity imply the following σ-sub-additivity: For any sequence $\{A_n\}_{n=1}^{\infty} \subset \mathcal{A}$,

$$\hat{\mathbb{V}}\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \hat{\mathbb{V}}(A_n).$$

An event A will be called a quasi-sure (for short q.s.) event if $\hat{\mathbb{V}}(A^c) = 0$, where A^c is the complementary set of A.

Example 3.2.2. A well-known method for obtaining a sub-additive probability is the following (see, e.g., Choquet [11]). Let $\mathcal{P}$ be a family of usual probabilities on $(\Omega, \mathcal{A})$. Let

$$\hat{\mathbb{V}}(A) = \sup_{Q \in \mathcal{P}} Q(A), \quad \forall A \in \mathcal{A}. \tag{3.2.1}$$

Then $\hat{\mathbb{V}}$ satisfies the properties listed in Definition 3.2.1.

Now, we turn to the notion of sub-additive (sometimes called sub-linear) expectation $\hat{\mathbb{E}}$. We shall apply the usual notion of a random variable. So an extended real-valued function X on Ω will be called a random variable if $X^{-1}(A) \in \mathcal{A}$ for any Borel set A. We assume, that there exists a subset $\mathcal{H}$ of the random variables and there exists an extended real-valued function $\hat{\mathbb{E}}[X]$ of $X \in \mathcal{H}$. We also assume that any non-negative random variable belongs to $\mathcal{H}$. We call attention to the fact that, in Definition 3.2.3, the sum $\hat{\mathbb{E}}[X] + \hat{\mathbb{E}}[Y]$ is undefined if it were $\hat{\mathbb{E}}[X] + \hat{\mathbb{E}}[Y] = \infty + (-\infty)$.

Definition 3.2.3. An extended real valued function $\hat{\mathbb{E}}[X]$ of $X \in \mathcal{H}$ is called a sub-additive expectation if it satisfies the following properties.

1. Monotonicity: If $X \leq Y$, then $\hat{\mathbb{E}}[X] \leq \hat{\mathbb{E}}[Y]$.

2. Constant preserving: $\hat{\mathbb{E}}[c] = c$, for any $c \in \mathbb{R}$.

3. Sub-additivity: $\hat{\mathbb{E}}[X + Y] \leq \hat{\mathbb{E}}[X] + \hat{\mathbb{E}}[Y]$.

4. Positive homogeneity: $\hat{\mathbb{E}}[\lambda X] = \lambda \hat{\mathbb{E}}[X]$ for any constant $\lambda \geq 0$.

5. Monotone convergence: If $X_n \uparrow X$, and $X_1 \geq 0$, then $\hat{\mathbb{E}}[X_n] \uparrow \hat{\mathbb{E}}[X]$.

$(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ is called a sub-additive expectation space in contrast with a probability space. Given $X = (X_1, X_2, \ldots, X_n)$, we say $X \in \mathcal{H}^n$ if $X_i \in \mathcal{H}$ for all $i = 1, \ldots, n$.

If a sub-additive probability $\hat{\mathbb{V}}$ is given in advance on $(\Omega, \mathcal{A})$, then we assume that $\hat{\mathbb{E}}[X] = \hat{\mathbb{E}}[Y]$ if $X = Y$ q.s.

If the sub-additive expectation $\hat{\mathbb{E}}[\,.\,]$ is given in advance, then we can introduce a sub-additive probability $\hat{\mathbb{V}}$ by $\hat{\mathbb{V}}(A) = \hat{\mathbb{E}}[\mathbb{I}_A]$, for any event A, where $\mathbb{I}_A$ denotes the indicator of A. Then this $\hat{\mathbb{V}}$ satisfies the properties given in Definition 3.2.1.

Example 3.2.4. A well-known method to obtain a sub-additive expectation is the following. Let $\mathcal{P}$ be a family of probabilities on $(\Omega, \mathcal{A})$. Let

$$\hat{\mathbb{E}}[X] = \sup_{Q \in \mathcal{P}} \mathbb{E}_Q[X], \tag{3.2.2}$$

where $\mathbb{E}_Q$ is the usual expectation corresponding to $Q \in \mathcal{P}$ such that

$$\mathbb{E}_Q[\mathbb{I}_A] = Q(A), \quad \forall A \in \mathcal{A}.$$

A simple calculation shows that the properties listed in Definition 3.2.3 are satisfied. If $\hat{\mathbb{E}}[X]$ is defined by (3.2.2), then relations among random variables in Definition 3.2.3 are considered as quasi-sure relations with respect to $\hat{\mathbb{V}}$ defined by (3.2.1).

Remark 3.2.5. Assume that an abstract sub-additive expectation $\hat{\mathbb{E}}[\,.\,]$ is given, which satisfies the properties listed in Definition 3.2.3. Under certain conditions, $\hat{\mathbb{E}}$ has a representation (3.2.2). For precise formulations of the statement, see [17] and [58].

3.3 Conditional sub-additive probability and expectation

In a classical probability space $(\Omega, \mathcal{A}, Q)$, the conditional expectation $\mathbb{E}_Q[X|\mathcal{F}]$ of a random variable X with respect to a sub-σ-algebra $\mathcal{F} \subset \mathcal{A}$ is itself an $\mathcal{F}$-measurable random variable which is simpler than X. The existence and uniqueness (up to Q-a.s. equivalence) of conditional expectation are guaranteed by the Radon-Nikodym theorem.

Here, we shall use the following abstract concept of sub-additive conditional expectation. Let $(\Omega, \mathcal{A})$ be a measurable space, and let $\mathcal{F} \subset \mathcal{A}$ be a sub-σ-algebra. We shall assume that there exists a subset $\mathcal{H}$ of the random variables on $(\Omega, \mathcal{A})$ and for any $X \in \mathcal{H}$ there exists an extended real-valued $\mathcal{F}$-measurable random variable $\hat{\mathbb{E}}[X|\mathcal{F}]$ so that the axioms in Definition 3.3.1 are satisfied. We call attention to the fact that, in Definition 3.3.1, we exclude the case of $\infty + (-\infty)$. We assume that any constant random variable belongs to $\mathcal{H}$.

Definition 3.3.1. $\hat{\mathbb{E}}[\cdot|\mathcal{F}]$ is called a sub-additive conditional expectation operator if $\hat{\mathbb{E}}[X|\mathcal{F}]$ is an $\mathcal{F}$-measurable random variable for any $X \in \mathcal{H}$, which satisfies the following properties.

1. *Non-triviality:* $-\infty < \hat{\mathbb{E}}[0|\mathcal{F}] < +\infty$;

2. *Monotonicity:* $\hat{\mathbb{E}}[X|\mathcal{F}] \geq \hat{\mathbb{E}}[Y|\mathcal{F}]$ if $X \geq Y$;

3. *Sub-additivity:* $\hat{\mathbb{E}}[X+Y|\mathcal{F}] \le \hat{\mathbb{E}}[X|\mathcal{F}] + \hat{\mathbb{E}}[Y|\mathcal{F}]$;

4. *Measurability:* $\hat{\mathbb{E}}[X+Y|\mathcal{F}] = X + \hat{\mathbb{E}}[Y|\mathcal{F}]$ if X is $\mathcal{F}$-measurable;

5. *Positive homogeneity:* $\hat{\mathbb{E}}[cX|\mathcal{F}] = c\hat{\mathbb{E}}[X|\mathcal{F}]$ if $c \ge 0$ is a constant;

6. *Monotone convergence:* If $X_n \uparrow X$, and $X_1 \ge 0$, then $\hat{\mathbb{E}}[X_n|\mathcal{F}] \uparrow \hat{\mathbb{E}}[X|\mathcal{F}]$.

The non-triviality axiom of Definition 3.3.1 excludes the cases $\hat{\mathbb{E}}[X|\mathcal{F}] \equiv +\infty$ and $\hat{\mathbb{E}}[X|\mathcal{F}] \equiv -\infty$. Then, by positive homogeneity, $\hat{\mathbb{E}}[0|\mathcal{F}] = \hat{\mathbb{E}}[0 \cdot 0|\mathcal{F}] = 0\hat{\mathbb{E}}[0|\mathcal{F}] = 0$. Then, by the measurability axiom, $\hat{\mathbb{E}}[X|\mathcal{F}] = \hat{\mathbb{E}}[X+0|\mathcal{F}] = X + \hat{\mathbb{E}}[0|\mathcal{F}] = X$ if X is $\mathcal{F}$-measurable.

Now, we recall the well-known definition of the essential supremum of a family of random variables, see [26] and [60]. Let $(\Omega, \mathcal{F}, P)$ be a probability space, let $\mathcal{K}$ be an arbitrary set of extended real valued random variables on Ω. Then there exists a countable subset $\mathcal{K}_0$ of $\mathcal{K}$ such that $\xi(\omega) = \sup\{\eta(\omega) : \eta \in \mathcal{K}_0\}$ is an extended real-valued random variable, and

(a) $P(\xi \ge \eta) = 1$ for each $\eta \in \mathcal{K}$,

(b) if τ is another extended real valued random variable for which $P(\tau \ge \eta) = 1$ for each $\eta \in \mathcal{K}$, then $P(\tau \ge \xi) = 1$.

ξ is called the essential supremum of $\mathcal{K}$ and it is denoted by

$$\xi = \operatorname*{ess\,sup}_P \{\eta : \eta \in \mathcal{K}\}.$$

Example 3.3.2. A simple way to obtain a sub-additive conditional expectation is as follows. Let $\mathcal{P}$ be a family of usual probabilities on $(\Omega, \mathcal{A})$. Let $\mathcal{F} \subset \mathcal{A}$ be a sub-σ-algebra and let P be a fixed probability measure on $(\Omega, \mathcal{F})$. Assume that, for every $Q \in \mathcal{P}$, the restriction of Q to the σ-algebra $\mathcal{F}$ and the measure P are mutually absolutely continuous.

Let $\mathcal{H}$ be the set of random variables on $(\Omega, \mathcal{A})$ such that $\mathbb{E}_Q[X|\mathcal{F}]$, i.e., the usual conditional expectation according to Q, which is an extended real valued $\mathcal{F}$-measurable function, exists for any $Q \in \mathcal{P}$. Let

$$\hat{\mathbb{E}}[X|\mathcal{F}] = \operatorname*{ess\,sup}_P \{\mathbb{E}_Q[X|\mathcal{F}] : Q \in \mathcal{P}\}. \tag{3.3.1}$$

A simple calculation shows that the properties listed in Definition 3.3.1 are satisfied. If $\hat{\mathbb{E}}[X|\mathcal{F}]$ is defined by (3.3.1), then relations among random variables in Definition 3.3.1 are considered as quasi-sure relations with respect to $\hat{\mathbb{V}}$ defined by (3.2.1).

Now, we turn to the abstract notion of conditional sub-additive probability (in other words, conditional capacity). Let $(\Omega, \mathcal{A})$ be a measurable space, and let $\mathcal{F} \subset \mathcal{A}$ be a sub-σ-algebra. We shall assume that for any $A \in \mathcal{A}$ there exists a real valued $\mathcal{F}$-measurable random variable $\hat{\mathbb{V}}[A|\mathcal{F}]$ so that the axioms of Definition 3.3.3 are satisfied.

Definition 3.3.3. $\hat{\mathbb{V}}[\,\cdot\,|\mathcal{F}]$ is called a sub-additive conditional probability operator if $\hat{\mathbb{V}}[A|\mathcal{F}]$ is an $\mathcal{F}$-measurable random variable for any $A \in \mathcal{A}$, which satisfies the following properties.

1. *Normalized:* $\hat{\mathbb{V}}[\emptyset|\mathcal{F}] = 0$, $\hat{\mathbb{V}}[\Omega|\mathcal{F}] = 1$;
2. *Monotonicity:* $\hat{\mathbb{V}}[A|\mathcal{F}] \le \hat{\mathbb{V}}[B|\mathcal{F}]$ if $A \subseteq B$;
3. *Sub-additivity:* $\hat{\mathbb{V}}[A \cup B|\mathcal{F}] \le \hat{\mathbb{V}}[A|\mathcal{F}] + \hat{\mathbb{V}}[B|\mathcal{F}]$;
4. *Lower continuity:* $\hat{\mathbb{V}}[A_n|\mathcal{F}] \uparrow \hat{\mathbb{V}}[A|\mathcal{F}]$ if $A_n \uparrow A$.

Example 3.3.4. If the conditional sub-additive expectation $\hat{\mathbb{E}}[\,\cdot\,|\mathcal{F}]$ is defined according to Definition 3.3.1 and the indicator $\mathbb{I}_A$ belongs to $\mathcal{H}$ for any event from $\mathcal{A}$, then $\hat{\mathbb{V}}[A|\mathcal{F}] = \hat{\mathbb{E}}[\mathbb{I}_A|\mathcal{F}]$ defines a sub-additive conditional probability satisfying the properties in Definition 3.3.3.

Now, we consider the case of Example 3.3.2.

Example 3.3.5. Let $\mathcal{P}$ be a family of usual probabilities on $(\Omega, \mathcal{A})$. Let $\mathcal{F} \subset \mathcal{A}$ be a sub-σ-algebra and let P be a fixed probability measure on $(\Omega, \mathcal{F})$. Assume that, for every $Q \in \mathcal{P}$, the restriction of Q to the σ-algebra $\mathcal{F}$ and the measure P are mutually absolutely continuous.

For any $Q \in \mathcal{P}$, let $Q[A|\mathcal{F}]$, be the usual conditional probability given $\mathcal{F}$ according to the usual probability Q. $Q[A|\mathcal{F}]$ is a real-valued $\mathcal{F}$-measurable random variable for any $Q \in \mathcal{P}$. Let

$$\hat{\mathbb{V}}[A|\mathcal{F}] = \operatorname*{ess\,sup}_{P}\{Q[A|\mathcal{F}] \;:\; Q \in \mathcal{P}\}. \tag{3.3.2}$$

Of course, $\hat{\mathbb{V}}[A|\mathcal{F}] = \hat{\mathbb{E}}[\mathbb{I}_A|\mathcal{F}]$ if $\hat{\mathbb{E}}[\,\cdot\,|\mathcal{F}]$ is defined by (3.3.1). A simple calculation shows that the properties listed in Definition 3.3.3 are satisfied. If $\hat{\mathbb{V}}[A|\mathcal{F}]$ is defined by (3.3.2), then relations among random variables in Definition 3.3.3 are considered as quasi-sure relations with respect to $\hat{\mathbb{V}}$ defined by (3.2.1): $\hat{\mathbb{V}}(A) = \sup_{Q\in\mathcal{P}} Q(A)$.

We see that in the situation of this example, each of our four operators is defined: $\hat{\mathbb{V}}(\,\cdot\,)$ is defined by equation (3.2.1), $\hat{\mathbb{V}}[\,\cdot\,|\mathcal{F}]$ is defined by (3.3.2), $\hat{\mathbb{E}}[\,\cdot\,]$ is defined by (3.2.2), and $\hat{\mathbb{E}}[\,\cdot\,|\mathcal{F}]$ is defined by (3.3.1).

We shall need information concerning the joint behaviour of $\hat{\mathbb{V}}$ and $\hat{\mathbb{E}}[\,\cdot\,|\mathcal{F}]$. So we introduce a further plausible axiom.

Definition 3.3.6. Let $(\Omega, \mathcal{A})$ be a measurable space, and let $\mathcal{F} \subset \mathcal{A}$ be a sub-σ-algebra. Let $\hat{\mathbb{V}}$ be a sub-additive probability on $\mathcal{F}$ satisfying the axioms given in Definition 3.2.1. Let $\hat{\mathbb{E}}[\,\cdot\,|\mathcal{F}]$ be a sub-additive conditional expectation satisfying the $\hat{\mathbb{V}}$ quasi-sure versions of the axioms in Definition 3.3.1. (It means that in Definition 3.3.1, any relation among random variables is understood as a $\hat{\mathbb{V}}$ quasi-sure relation.) Then the finiteness axiom is the following:

1. *Finiteness:* If $\hat{\mathbb{E}}[X|\mathcal{F}] < \infty$ $\hat{\mathbb{V}}$-quasi-surely, then $X < \infty$ $\hat{\mathbb{V}}$-quasi-surely.

The above Finiteness axiom excludes certain non-conventional cases as the example below shows.

Example 3.3.7. Let B_1 and B_2 be non-empty disjoint sets, $\Omega = B_1 \cup B_2$, $\mathcal{A} = \{\Omega, \emptyset, B_1, B_2\}$, $\mathcal{F} = \{\Omega, \emptyset\}$. Let $\hat{\mathbb{V}}$ be a usual probability on $\mathcal{A}$ with $\hat{\mathbb{V}}(B_1) > 0$, $\hat{\mathbb{V}}(B_2) > 0$. For a random variable $X = b_1 \mathbb{I}_{B_1} + b_2 \mathbb{I}_{B_2}$, where b_1 and b_2 are fixed numbers (possibly $\pm\infty$) define $\hat{\mathbb{E}}[X|\mathcal{F}] = b_2$. Then the axioms in Definition 3.3.1 are satisfied, but the Finiteness axiom in Definition 3.3.6 fails with $X = \infty \cdot \mathbb{I}_{B_1} + 0 \cdot \mathbb{I}_{B_2}$ because $\hat{\mathbb{V}}(B_1) > 0$.

For the proof of some SLLNs, we shall need an upper continuity property. However, it may fail even in the case of a non-conditional sub-additive probability.

Example 3.3.8. Let $\Omega = [-1, 1]$, let $\mathcal{A}$ be the family of its Borel sets, and let Q_n be the uniform distribution on $[-1/n, +1/n]$. Let $\hat{\mathbb{V}}(A) = \sup_n Q_n(A)$ for any Borel set of Ω. Then for any $-1 < a < 0$ and $0 < b < 1$, we have $\hat{\mathbb{V}}([a, b]) = 1$, but $\hat{\mathbb{V}}(\{0\}) = 0$. Now, let $A_\infty = \{0\}$ and $A_n = [-1/n, +1/n]$. Therefore, $A_n \downarrow A_\infty$ but $\hat{\mathbb{V}}(A_n) = 1$ so it does not converge to $\hat{\mathbb{V}}(A_\infty) = 0$. We remark that, in this case, the family of probability measures Q_n, $n = 1, 2, \ldots$ is not compact.

However, in the case of $A_n \downarrow A_\infty$ and $\hat{\mathbb{V}}(A_n) \downarrow 0$, we have $\hat{\mathbb{V}}(A_\infty) = 0$. It is a consequence of the non-negativity and the monotonicity axioms: $0 \le \hat{\mathbb{V}}(A_\infty) \le \hat{\mathbb{V}}(A_n) \downarrow 0$. So in this particular case, $\hat{\mathbb{V}}(A_n) \downarrow \hat{\mathbb{V}}(A_\infty)$. We shall need a similar upper continuity for the conditional sub-additive probability. It turns out that the following plausible recursivity axiom does the job.

Let us introduce the recursivity axiom concerning the joint behaviour of $\hat{\mathbb{V}}$ and $\hat{\mathbb{V}}[\,\cdot\,|\mathcal{F}]$.

Definition 3.3.9. Let $(\Omega, \mathcal{A})$ be a measurable space, and let $\mathcal{F} \subset \mathcal{A}$ be a sub-σ-algebra. Let $\hat{\mathbb{V}}$ be a sub-additive probability on $\mathcal{F}$ satisfying the axioms given in Definition 3.2.1. Let $\hat{\mathbb{V}}[\,\cdot\,|\mathcal{F}]$ be a sub-additive conditional probability satisfying the $\hat{\mathbb{V}}$ quasi-sure versions of the axioms in Definition 3.3.3. (It means that in Definition 3.3.3, any relation among random variables are understood as a $\hat{\mathbb{V}}$ quasi-sure relation.) Then the recursivity axiom is the following:

1. *Recursivity:* If $\hat{\mathbb{V}}[A|\mathcal{F}] = 0$ $\hat{\mathbb{V}}$-quasi-surely, then $\hat{\mathbb{V}}(A) = 0$.

The Recursivity axiom excludes unconventional cases. For example, when both $\hat{\mathbb{V}}[A|\mathcal{F}] = P_1(A)$ and $\hat{\mathbb{V}}(A) = P_2(A)$ are usual probabilities and P_2 is not absolutely continuous with respect to P_1.

Remark 3.3.10. The recursivity axiom implies the following upper continuity property. If $\hat{\mathbb{V}}[A_n|\mathcal{F}] \to 0$ $\hat{\mathbb{V}}$-quasi-surely, and $A_n \downarrow A_\infty$, then $\hat{\mathbb{V}}(A_\infty) = 0$. To show it, we apply the monotonicity

$$0 \le \hat{\mathbb{V}}[A_\infty|\mathcal{F}] \le \hat{\mathbb{V}}[A_n|\mathcal{F}] \to 0 \quad \hat{\mathbb{V}}\text{-quasi-surely},$$

so $\hat{\mathbb{V}}[A_\infty|\mathcal{F}] = 0$ $\hat{\mathbb{V}}$-quasi-surely, so the recursivity axiom gives $\hat{\mathbb{V}}(A_\infty) = 0$.

The finiteness and the recursivity axioms are satisfied in the following cases.

Example 3.3.11. Consider the setting of Example 3.3.5. That is, let $\mathcal{P}$ be a family of usual probabilities on $(\Omega, \mathcal{A})$. Let $\mathcal{F} \subset \mathcal{A}$ be a sub-σ-algebra and let P be a fixed probability measure on $(\Omega, \mathcal{F})$. Assume that, for every $Q \in \mathcal{P}$, the restriction of Q to the σ-algebra $\mathcal{F}$ and the measure P are mutually absolutely continuous. Let $\hat{\mathbb{V}}(A) = \sup_{Q\in\mathcal{P}} Q(A)$ be the sub-additive probability.

1. Let $\hat{\mathbb{E}}[X|\mathcal{F}] = \operatorname{ess\,sup}_P\{E_Q[X|\mathcal{F}] : Q \in \mathcal{P}\}$ be the sub-additive conditional expectation. Then the finiteness axiom from Definition 3.3.6 is satisfied.

2. Let $\hat{\mathbb{V}}[\,A\,|\mathcal{F}] = \operatorname{ess\,sup}_P\{Q[\,A\,|\mathcal{F}] : Q \in \mathcal{P}\}$ be the sub-additive conditional probability. Then the recursivity axiom from Definition 3.3.9 is satisfied.

Remark 3.3.12. A well-known example of sub-additive (conditional) probability and expectation is served by the theory of G-Brownian motion elaborated by Peng and his co-authors, see [56] and [58]. Within this theory, the sub-additive expectation and conditional expectation are known as the G-expectation and the conditional G-expectation, respectively. The properties we want to apply are known for those notions. A representation like (3.3.1) is also known for the conditional G-expectation, see [69].

Remark 3.3.13. In [15], Cohen applied an axiomatic approach to sub-additive expectation and the corresponding sub-additive conditional expectation and looked for a (3.3.1)-like representation. Under the so-called Hahn-property, he proved that $\hat{\mathbb{E}}[X|\mathcal{F}]$ is equal to the generalized essential supremum of certain usual conditional expectations.

3.4 Hájek-Rényi-type maximal inequalities for conditional sub-additive expectations and capacities

In this section, we follow the approach presented in [23]. First, we prove that the conditional Kolmogorov inequality for sub-additive expectation implies the conditional Hájek-Rényi inequality for sub-additive expectation. In the following theorem, we consider the setting of Definition 3.3.1. We assume that all random variables studied belong to the space $\mathcal{H}$. Let $\{X_i\}_{i\geq 1}$ denote a sequence of random variables in the space $\mathcal{H}$. Let the partial sums of the random variables be $S_k = \sum_{i=1}^{k} X_i$, for all $k \in \mathbb{N}$, and let $S_0 = 0$.

Theorem 3.4.1 (Masasila & Fazekas [50])**.** *Let $\{X_k, 1 \leq k \leq n\}$ be random variables belonging to the space $\mathcal{H}$. Assume that the conditional expectation operator $\hat{\mathbb{E}}[\,\cdot\,|\mathcal{F}]$ on space $\mathcal{H}$ satisfies the monotonicity, sub-additivity and positive homogeneity axioms of Definition 3.3.1. Let $\alpha_1, \dots, \alpha_n$ be non-negative $\mathcal{F}$-measurable random variables, and $r > 0$ be real number. Assume that the general conditional Kolmogorov-type inequality is true, that is,*

$$\hat{\mathbb{E}}\left[\left(\max_{1\leq l\leq m} |S_l|\right)^r \,\middle|\mathcal{F}\right] \leq \sum_{l=1}^{m} \alpha_l \qquad \text{for all} \quad 1 \leq m \leq n. \tag{3.4.1}$$

Then the conditional Hájek-Rényi inequality is true, that is,

$$\hat{\mathbb{E}}\left[\left(\max_{1\le l\le n}\left|\frac{S_l}{\beta_l}\right|\right)^r|\mathcal{F}\right] \le 4\sum_{l=1}^{n}\frac{\alpha_l}{\beta_l^r} \tag{3.4.2}$$

for $\mathcal{F}$-measurable random variables $\beta_1 \le \beta_2 \le \ldots \le \beta_n$ with $\beta_1 \ge \beta_0$, where β_0 is a positive constant.

Proof. Multiplying both sides of inequality (3.4.2) by β_0^r, we see that we can assume $\beta_1 \ge 1$ during the proof. Let $c = 2^{\frac{1}{r}}$. Let K_i be the set of subscripts k for which $c^i \le \beta_k < c^{i+1}$, i.e. $K_i = \{k : c^i \le \beta_k < c^{i+1}\}, \quad i = 0, 1, 2, \ldots$. Then, K_i is $\mathcal{F}$-measurable because β_k is $\mathcal{F}$-measurable. Let $i(n)$ be the index of the last non-empty K_i. Then $i(n)$ is an $\mathcal{F}$-measurable random variable (possibly having value ∞). Let $k(i)$ be the maximal index in K_i. More precisely, $k(i) = \max\{k : k \in K_i\}$, if A_i is non-empty, but $k(i) = k(i-1)$ if K_i is empty ($k(-1) = 0$ by definition).

Let

$$\delta_l = \sum_{j=k(l-1)+1}^{k(l)} \alpha_j, \quad l = 0, 1, 2, \ldots$$

be the sum of the values α_j in K_l. Then $k(i)$ and δ_l are $\mathcal{F}$-measurable, $k(i) \le n$. Then, by using monotonicity, sub-additivity, and positive homogeneity of $\hat{\mathbb{E}}[\cdot|\mathcal{F}]$, and inequality (3.4.1), we obtain the following sequence of inequalities.

$$\begin{aligned}
\hat{\mathbb{E}}\left[\left(\max_{1\le l\le n}\frac{|S_l|}{\beta_l}\right)^r|\mathcal{F}\right] &\le \hat{\mathbb{E}}\left[\sum_{i=0}^{i(n)}\left(\max_{l\in K_i}\frac{|S_l|}{\beta_l}\right)^r|\mathcal{F}\right] && \text{(by monotonicity)}\\
&\le \sum_{i=0}^{i(n)}\hat{\mathbb{E}}\left[\left(\max_{l\in K_i}\frac{|S_l|}{\beta_l}\right)^r|\mathcal{F}\right] && \text{(by sub-additivity)}\\
&\le \sum_{i=0}^{i(n)}c^{-ir}\hat{\mathbb{E}}\left[\left(\max_{l\in K_i}|S_l|\right)^r|\mathcal{F}\right] && \text{(by positive homogeneity)}\\
&\le \sum_{i=0}^{i(n)}c^{-ir}\hat{\mathbb{E}}\left[\left(\max_{k\le k(i)}|S_k|\right)^r|\mathcal{F}\right] && \text{(by monotonicity)}\\
&\le \sum_{i=0}^{i(n)}c^{-ir}\sum_{k=1}^{k(i)}\alpha_k = \sum_{i=0}^{i(n)}c^{-ir}\sum_{l=0}^{i}\delta_l && \text{(by (3.4.1))}\\
&= \sum_{l=0}^{i(n)}\delta_l\sum_{i=l}^{i(n)}c^{-ir} \le \sum_{l=0}^{i(n)}\delta_l\sum_{i=l}^{\infty}c^{-ir}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{1-c^{-r}} \sum_{l=0}^{i(n)} c^{-lr} \delta_l \\
&= \frac{1}{1-c^{-r}} \sum_{l=0}^{i(n)} c^{-lr} \sum_{k=k(l-1)+1}^{k(l)} \alpha_k \\
&\leq \frac{1}{1-c^{-r}} \sum_{l=0}^{i(n)} c^{-lr} \sum_{k=k(l-1)+1}^{k(l)} \alpha_k \frac{c^{lr+r}}{\beta_k^r} \\
&= \frac{c^r}{1-c^{-r}} \sum_{l=0}^{i(n)} \sum_{k=k(l-1)+1}^{k(l)} \frac{\alpha_k}{\beta_k^r} \\
&= 4 \sum_{k=1}^{n} \frac{\alpha_k}{\beta_k^r}.
\end{aligned}
$$

So we obtained inequality (3.4.2). □

Remark 3.4.2. Another version of Theorem 3.4.1 can be obtained as follows. Assume that there is a sub-additive probability $\hat{\mathbb{V}}$ on the space $(\Omega, \mathcal{A})$. Suppose that all assumptions of Theorem 3.4.1 are satisfied $\hat{\mathbb{V}}$-quasi-surely, including the axioms listed in Definition 3.3.1. That is, in the axioms of monotonicity, sub-additivity, and positive homogeneity, all relations among random variables are understood in the $\hat{\mathbb{V}}$-quasi-sure sense. Then inequality (3.4.2) is true $\hat{\mathbb{V}}$-quasi-surely.

Our next theorem shows that the conditional Kolmogorov inequality for sub-additive probability implies the conditional Hájek-Rényi inequality for sub-additive probability. In the following theorem, we consider the setting of Definition 3.3.3.

Theorem 3.4.3 (Masasila & Fazekas [50])**.** *Let $\{X_n, 1 \leq k \leq n\}$ be random variables, $S_k = X_1 + \cdots + X_k$. Let $\hat{\mathbb{V}}[\,\cdot\,|\mathcal{F}]$ be a conditional sub-additive probability satisfying the axioms normalization, monotonicity, and sub-additivity of Definition 3.3.3. Let r be a positive real number. Let $\beta_1 \leq \beta_2 \leq \cdots \leq \beta_n$ be $\mathcal{F}$-measurable, $\alpha_1, \ldots, \alpha_n$ non-negative $\mathcal{F}$-measurable random variables. Assume that $\beta_1 \geq \beta_0 > 0$, where β_0 is non random. If*

$$
\hat{\mathbb{V}}\left[\max_{1\leq l\leq m} |S_l| \geq \varepsilon \middle| \mathcal{F}\right] \leq \frac{1}{\varepsilon^r} \sum_{l=1}^{m} \alpha_l \qquad \text{for all} \quad 1 \leq m \leq n \tag{3.4.3}
$$

and for all $\varepsilon > 0$, then

$$
\hat{\mathbb{V}}\left[\max_{1\leq l\leq n} \left|\frac{S_l}{\beta_l}\right| \geq \varepsilon \middle| \mathcal{F}\right] \leq \frac{4}{\varepsilon^r} \sum_{k=1}^{n} \frac{\alpha_k}{\beta_k^r} \tag{3.4.4}
$$

for all $\varepsilon > 0$.

Proof. We use the same notation as in the proof of Theorem 3.4.1. Then, by sub-additivity and monotonicity of $\hat{\mathbb{V}}[\,\cdot\,|\mathcal{F}]$ and inequality (3.4.3),

$$
\begin{aligned}
\hat{\mathbb{V}}\left[\max_{1\le l\le n}\frac{|S_l|}{\beta_l}\ge\varepsilon\big|\mathcal{F}\right] &\le \sum_{i=0}^{i(n)}\hat{\mathbb{V}}\left[\max_{l\in K_i}\frac{|S_l|}{\beta_l}\ge\varepsilon\big|\mathcal{F}\right] \quad \text{(by sub-additivity)}\\
&\le \sum_{i=0}^{i(n)}\hat{\mathbb{V}}\left[\max_{l\in K_i}\frac{|S_l|}{c^i}\ge\varepsilon\big|\mathcal{F}\right] \quad \text{(by monotonicity)}\\
&\le \sum_{i=0}^{i(n)}\hat{\mathbb{V}}\left[\max_{k\le k(i)}\frac{|S_k|}{c^i}\ge\varepsilon\big|\mathcal{F}\right] \quad \text{(by monotonicity)}\\
&\le \sum_{i=0}^{i(n)}(\varepsilon c^i)^{-r}\sum_{k=1}^{k(i)}\alpha_k = \sum_{i=0}^{i(n)}(\varepsilon c^i)^{-r}\sum_{l=0}^{i}\delta_l \quad \text{(by (3.4.3))}\\
&= \sum_{l=0}^{i(n)}\delta_l\sum_{i=l}^{i(n)}(\varepsilon c^i)^{-r} \le \sum_{l=0}^{i(n)}\delta_l\sum_{i=l}^{\infty}(\varepsilon c^i)^{-r}\\
&= \varepsilon^{-r}\frac{1}{1-c^{-r}}\sum_{l=0}^{i(n)}c^{-lr}\delta_l\\
&= \varepsilon^{-r}\frac{1}{1-c^{-r}}\sum_{l=0}^{i(n)}c^{-lr}\sum_{k=k(l-1)+1}^{k(l)}\alpha_k\\
&\le \varepsilon^{-r}\frac{1}{1-c^{-r}}\sum_{l=0}^{i(n)}c^{-lr}\sum_{k=k(l-1)+1}^{k(l)}\alpha_k\frac{c^{lr+r}}{\beta_k^r}\\
&= \varepsilon^{-r}\frac{c^r}{1-c^{-r}}\sum_{l=0}^{i(n)}\sum_{k=k(l-1)+1}^{k(l)}\frac{\alpha_k}{\beta_k^r}\\
&= 4\varepsilon^{-r}\sum_{k=1}^{n}\frac{\alpha_k}{\beta_k^r}.
\end{aligned}
$$

So we obtained inequality (3.4.4). □

Remark 3.4.4. Another version of Theorem 3.4.3 can be obtained as follows. Assume that there is a sub-additive probability $\hat{\mathbb{V}}$ of the space $(\Omega,\mathcal{A})$. Suppose that all assumptions of Theorem 3.4.3 are satisfied $\hat{\mathbb{V}}$-quasi-surely, including the axioms listed in Definition 3.3.3. That is, in Definition 3.3.3, in the axioms of monotonicity, sub-additivity, and positive homogeneity, all relations among random variables are understood in the $\hat{\mathbb{V}}$-quasi-sure sense. Then inequality (3.4.4) is true $\hat{\mathbb{V}}$ quasi-surely.

3.5 Strong laws of large numbers in terms of conditional sub-additive expectations and capacities

The intuitive background of our next theorem is the well-known Kolmogorov SLLN (see [29], p. 288). If we consider a usual probability space and independent, zero-mean random variables $X_1, X_2, \dots$ and $\alpha_l = C\mathrm{var}(X_l)$, then assumption (3.5.1) is satisfied by the Kolmogorov-Doob inequality (see [29], p. 505), and our Theorem 3.5.1 is the same as the above mentioned Kolmogorov SLLN.

We now prove a general strong law of large numbers under assumptions formulated in terms of conditional sub-additive expectations. Let $(\Omega, \mathcal{A})$ be a measurable space, and assume that there exists a sub-additive probability $\hat{\mathbb{V}}$ on $(\Omega, \mathcal{A})$ satisfying the axioms given in Definition 3.2.1. Throughout the next theorem, a quasi-sure event is understood in the sense of $\hat{\mathbb{V}}$. Let $\mathcal{F}$ be a σ-sub-algebra of $\mathcal{A}$, and let $\hat{\mathbb{E}}[\,\cdot\,|\mathcal{F}]$ denote a sub-additive conditional expectation. In the next theorem all random variables will be defined on $(\Omega, \mathcal{A})$.

Theorem 3.5.1 (Masasila & Fazekas [50]). *Let $\{X_n, n \geq 1\}$ be random variables, $S_n = X_1 + \cdots + X_n$ for any n. Let $b_1, b_2, \dots$ be q.s. finite, $\mathcal{F}$-measurable random variables with $b_0 \leq b_1 \leq b_2 \leq \dots$ q.s., $b_n \to \infty$ q.s., where b_0 is a positive constant. Let $\alpha_1, \alpha_2, \dots$ be non-negative $\mathcal{F}$-measurable random variables. Assume that for the conditional expectation $\hat{\mathbb{E}}[\,\cdot\,|\mathcal{F}]$ the axioms in Definition 3.3.1 are satisfied, where all relations among random variables are understood in the $\hat{\mathbb{V}}$-quasi-sure sense. Assume further that the finiteness axiom in Definition 3.3.6 holds. Let $r > 0$ be a fixed number and suppose that, for any $n \geq 1$*

$$\hat{\mathbb{E}}\left[\left(\max_{1\leq l\leq n} |S_l|\right)^r |\mathcal{F}\right] \leq \sum_{l=1}^{n} \alpha_l \quad q.s. \tag{3.5.1}$$

If $\sum_{l=1}^{\infty} \frac{\alpha_l}{b_l^r} < \infty$ q.s., then

$$\lim_{n\longrightarrow\infty} \frac{S_n}{b_n} = 0 \quad q.s. \tag{3.5.2}$$

and $\frac{S_n}{b_n} = \mathrm{O}\left(\frac{\beta_n}{b_n}\right)$, where β_n is defined by (3.5.3).

Proof. We shall apply the method of [23]. We can assume that $\alpha_n \geq \alpha_n' > 0$ and α_n' is non random for any n. To see it, let a non random $\alpha_n' > 0$, for any n so that $\sum_n \alpha_l' < \infty$. Then, instead of α_n, we can consider $\max\{\alpha_n, \alpha_n'\}$. Let

$$v_n = \sum_{k=n}^{\infty} \frac{\alpha_k}{b_k^r}, \qquad \beta_n = \max_{1\leq k\leq n} b_k v_k^{\frac{1}{2r}}. \tag{3.5.3}$$

Then the sequence β_n is increasing, $\beta_1 > \beta_0 > 0$ where β_0 is non random. Then, because of the assumption $\sum_{l=1}^{\infty} \frac{\alpha_l}{b_l^r} < \infty$ q.s., we have

$$0 < v_n < \infty \quad \text{q.s. for all, } n \quad v_n \to 0 \quad \text{q.s.,}$$

and v_n is a decreasing sequence. Then, using the Abel-Dini theorem,

$$\sum_{n=1}^{\infty} \frac{\alpha_n}{b_n^r v_n^{\frac{1}{2}}} < \infty \quad \text{q.s.}$$

Therefore we have

$$0 < \beta_0 \leq \beta_1 \leq \beta_2 \leq \dots, \quad \beta_0 \quad \text{is non random,}$$

$$\sum_{k=1}^{\infty} \frac{\alpha_k}{\beta_k^r} < \infty \qquad \text{q.s.,}$$

$$\lim_{k \longrightarrow \infty} \frac{\beta_k}{b_k} = 0 \qquad \text{q.s.}$$

Then our Theorem 3.4.1 implies

$$\hat{\mathbb{E}}\left[\max_{1 \leq l \leq n} \left|\frac{S_l}{\beta_l}\right|^r |\mathcal{F}\right] \leq 4 \sum_{l=1}^{n} \frac{\alpha_l}{\beta_l^r} \quad \text{q.s. for all} \quad n.$$

So, by the monotone convergence axiom from Definition 3.3.1,

$$\hat{\mathbb{E}}\left[\sup_{1 \leq l \leq \infty} \left|\frac{S_l}{\beta_l}\right|^r |\mathcal{F}\right] \leq 4 \sum_{l=1}^{\infty} \frac{\alpha_l}{\beta_l^r} < \infty \quad \text{q.s.}$$

So, by the finiteness axiom in Definition 3.3.6,

$$\sup_{1 \leq l \leq \infty} \left|\frac{S_l}{\beta_l}\right|^r < \infty \quad \text{q.s.}$$

Therefore

$$0 \leq \left|\frac{S_n}{b_n}\right| = \left|\frac{S_n}{\beta_n}\right| \frac{\beta_n}{b_n} \leq \left(\sup_{1 \leq l \leq \infty} \left|\frac{S_l}{\beta_l}\right|\right) \frac{\beta_n}{b_n} \longrightarrow 0 \quad \text{q.s. as} \quad n \to \infty.$$

Hence, $\lim_{n \to \infty} \frac{S_n}{b_n} = 0$ q.s and $\frac{S_n}{b_n} = \mathrm{O}\left(\frac{\beta_n}{b_n}\right)$ q.s. □

We now turn to a general strong law of large numbers in which the assumptions are formulated in terms of conditional sub-additive probability. In the following theorem, all random variables are defined on the measurable space $(\Omega, \mathcal{A})$, which is equipped with a sub-additive probability $\hat{\mathbb{V}}$ satisfying the axioms of Definition 3.2.1. As before, quasi-sure events are understood in the sense of $\hat{\mathbb{V}}$. Let $\mathcal{F}$ be a sub-σ-algebra of $\mathcal{A}$ and let $\hat{\mathbb{V}}[\,\cdot\,|\mathcal{F}]$ denote a sub-additive conditional probability.

Theorem 3.5.2 (Masasila & Fazekas [50]). *Let $\{X_n, n \geq 1\}$ be random variables, $S_n =$*

$X_1 + \cdots + X_n$ for any n. Let $b_1, b_2, \ldots$ be q.s. finite, $\mathcal{F}$-measurable random variables with $b_0 \le b_1 \le b_2 \le \ldots$ q.s., $b_n \to \infty$ q.s., where b_0 is a positive constant. Let $\alpha_1, \alpha_2, \ldots$ be non-negative $\mathcal{F}$-measurable random variables. Assume that the conditional sub-additive probability $\hat{\mathbb{V}}[\cdot|\mathcal{F}]$ satisfies the axioms in Definition 3.3.3, where all relations among random variables are understood in the $\hat{\mathbb{V}}$-quasi-sure sense. Assume further that the recursivity axiom in Definition 3.3.9 holds. Let $r > 0$ be a fixed number and suppose that, for all $n \ge 1$ and all $\varepsilon > 0$,

$$\hat{\mathbb{V}}\left[\max_{1\le l\le n} |S_l| \ge \varepsilon \middle| \mathcal{F}\right] \le \frac{1}{\varepsilon^r}\sum_{l=1}^{n} \alpha_l \quad q.s. \tag{3.5.4}$$

If $\sum_{l=1}^{\infty} \frac{\alpha_l}{b_l^r} < \infty$ q.s., then

$$\lim_{n\to\infty} \frac{S_n}{b_n} = 0 \quad q.s. \tag{3.5.5}$$

with the convergence rate $\frac{S_n}{b_n} = \mathrm{O}\left(\frac{\beta_n}{b_n}\right)$ q.s.

Proof. As in the proof of Theorem 3.5.1, we can assume that $\alpha_n \ge \alpha_n' > 0$, where α_n' is non random for any n. Let

$$v_n = \sum_{k=n}^{\infty} \frac{\alpha_k}{b_k^r}, \qquad \beta_n = \max_{1\le k\le n} b_k v_k^{\frac{1}{2r}}.$$

Then, because of the assumption $\sum_{l=1}^{\infty} \frac{\alpha_l}{b_l^r} < \infty$ q.s., we have

$$0 < v_n < \infty \quad \text{for all } n \ge 1 \quad \text{q.s. and } v_n \to 0 \quad \text{q.s.}$$

Moreover, the Abel-Dini theorem implies

$$\sum_{n=1}^{\infty} \frac{\alpha_n}{b_n^r v_n^{\frac{1}{2}}} < \infty \quad \text{q.s.}$$

Therefore $\beta_1, \beta_2, \ldots$ is an increasing sequence, $\beta_1 \ge \beta_0 > 0$, where β_0 is non random,

$$\sum_{k=1}^{\infty} \frac{\alpha_k}{\beta_k^r} < \infty, \qquad \text{q.s.,}$$

$$\lim_{k\longrightarrow\infty} \frac{\beta_k}{b_k} = 0 \qquad \text{q.s.}$$

Then our Theorem 3.4.3 implies

$$\hat{\mathbb{V}}\left[\max_{1\le l\le n} \frac{|S_l|}{\beta_l} \ge \varepsilon \middle| \mathcal{F}\right] \le \frac{4}{\varepsilon^r}\sum_{l=1}^{n} \frac{\alpha_l}{\beta_l^r} \quad \text{q.s. for all} \quad n \quad \text{and} \quad \varepsilon > 0.$$

So, by the lower continuity axiom,

$$\hat{\mathbb{V}}\left[\sup_{1\le l<\infty}\frac{|S_l|}{\beta_l}\ge\varepsilon\big|\mathcal{F}\right]\le\frac{4}{\varepsilon^r}\sum_{l=1}^{\infty}\frac{\alpha_l}{\beta_l^r}\quad\text{q.s.}$$

Let $\varepsilon\to\infty$. Then by the recursivity axiom, we have

$$\sup_{1\le l<\infty}\frac{|S_l|}{\beta_l}<\infty\quad\text{q.s.}$$

Now

$$0\le\left|\frac{S_n}{b_n}\right|=\left|\frac{S_n}{\beta_n}\right|\frac{\beta_n}{b_n}\le\left(\sup_{1\le l<\infty}\frac{|S_l|}{\beta_l}\right)\frac{\beta_n}{b_n}\to 0\quad\text{q.s. as}\quad n\to\infty$$

because $\frac{\beta_n}{b_n}\to 0$ q.s. Therefore,

$$\lim_{n\to\infty}\frac{S_n}{b_n}=0\qquad\text{q.s.}$$

So we obtained (3.5.5). □

3.6 Application to conditionally negatively dependent random variables

In this section, we assume the following setting. All random variables are defined on a measurable space $(\Omega,\mathcal{A})$ equipped with a sub-additive probability $\hat{\mathbb{V}}$ satisfying the axioms given in Definition 3.2.1. A quasi-sure event will be understood as a $\hat{\mathbb{V}}$ quasi-sure (q.s.) event. Let $\mathcal{F}$ be a σ-sub-algebra of $\mathcal{A}$. Let $\hat{\mathbb{E}}[\cdot|\mathcal{F}]$ denote a sub-additive conditional expectation. Assume that in Definition 3.3.1, any relation among random variables is a quasi-sure relation. Moreover, in this section, we understand any relation among random variables as a quasi-sure relation.

Negative dependence plays an important role in probability theory, as it constitutes a dependence structure that is strictly weaker than independence. In the classical framework of additive probabilities and linear expectations, various aspects of negatively dependent random variables have been extensively studied; see, for example, [51], [55], and [75].

Within the framework of sub-additive expectations, Zhang LiXin [79] introduced a notion of negative dependence and established useful maximal inequalities under this setting. These inequalities were subsequently applied by Huang [35] to derive strong laws of large numbers for negatively dependent random variables under sub-additive expectations.

In this section, we introduce a definition of negative dependence for random variables under conditional sub-additive expectations, which naturally extends the concept proposed in [79]. By adapting the approach of [79], we establish corresponding maximal inequalities in this conditional framework. Finally, combining these inequalities with our general results developed earlier, we derive a strong law of large numbers for conditionally negatively dependent random variables under sub-additive expectations. Throughout this section, we follow the notation of [79].

First, we recall the notion of a local Lipschitz function, see, e.g., [79]. $C_{l,\mathrm{Lip}}(\mathbb{R}^n)$ denotes the linear space of real-valued functions φ satisfying

$$|\varphi(x) - \varphi(y)| \le C(1 + \|x\|^m + \|y\|^m)\|x - y\|, \quad \text{for all} \quad x, y \in \mathbb{R}^n,$$

for some $C > 0$, and $m \in \mathbb{N}$ depending on φ. Here $\|x\|$ denotes the Euclidean norm of $x \in \mathbb{R}^n$.

Definition 3.6.1 (Conditional negative dependence under sub-additive expectation)**.** Let X be an m-dimensional random vector and Y be an n-dimensional random vector.

We say that Y is *negatively dependent on X under the conditional sub-additive expectation* $\hat{\mathbb{E}}[\,\cdot\,|\mathcal{F}]$ if, for every pair of test functions $\psi_1 \in C_{l,\mathrm{Lip}}(\mathbb{R}^m)$ and $\psi_2 \in C_{l,\mathrm{Lip}}(\mathbb{R}^n)$, we have

$$\hat{\mathbb{E}}\left[\,\psi_1(X)\psi_2(Y)\,|\,\mathcal{F}\right] \le \hat{\mathbb{E}}\left[\,\psi_1(X)\,|\,\mathcal{F}\right]\hat{\mathbb{E}}\left[\,\psi_2(Y)\,|\,\mathcal{F}\right],$$

whenever $\psi_1(X) \ge 0$, $\hat{\mathbb{E}}[\,\psi_2(Y)|\mathcal{F}] \ge 0$, and $\hat{\mathbb{E}}[|\psi_1(X)\psi_2(Y)|\,|\mathcal{F}] < \infty$, $\hat{\mathbb{E}}[|\psi_1(X)|\,|\mathcal{F}] < \infty$, $\hat{\mathbb{E}}[|\psi_2(Y)|\,|\mathcal{F}] < \infty$, and either both ψ_1 and ψ_2 are coordinate-wise non-decreasing or both are coordinate-wise non-increasing.

We give a simple example for conditional negative dependence under sub-additive expectation. We apply the idea used in the non-conditional case, see [36], [79].

Example 3.6.2. Let $P_1, P_2, \dots$ be usual probabilities on the space $(\Omega, \mathcal{A})$. Let X, Y, Z be random variables, let $\mathcal{F}$ be the sub-σ-algebra generated by Z. Let $\hat{\mathbb{E}}[\,\cdot\,|\mathcal{F}] = \sup_i \mathbb{E}_i[\,\cdot\,|\mathcal{F}]$, where $\mathbb{E}_i[\,\cdot\,|\mathcal{F}]$ is the usual conditional expectation under the probability P_i. Assume that X and Y are conditionally independent given Z under each P_i. If the joint density functions exist, it means $f_i(x, y|z) = f_i(x|z)f_i(y|z)$ for any i, where the subscript i refers to P_i (This factorization holds, for example, in the Gaussian case when X and Y are conditionally uncorrelated given Z). This equality gives us

$$\mathbb{E}_i[\psi_1(X)\psi_2(Y)|Z = z] = \mathbb{E}_i[\psi_1(X)|Z = z]\mathbb{E}_i[\psi_2(Y)|Z = z]$$

for any i and $\psi_1, \psi_2 \in C_{l,\mathrm{Lip}}(\mathbb{R}^1)$. We assume also $\psi_1(X) \ge 0$, $\hat{\mathbb{E}}[\,\psi_2(Y)|\mathcal{F}] \ge 0$. From the above equation, we have

$$\mathbb{E}_i[\psi_1(X)\psi_2(Y)|\mathcal{F}] = \mathbb{E}_i[\psi_1(X)|\mathcal{F}]\mathbb{E}_i[\psi_2(Y)|\mathcal{F}].$$

In this equation taking the supremum for all i, and using the assumption $\psi_1(X) \ge 0$, we obtain

$$\hat{\mathbb{E}}[\psi_1(X)\psi_2(Y)|\mathcal{F}] \le \hat{\mathbb{E}}[\psi_1(X)|\mathcal{F}]\hat{\mathbb{E}}[\psi_2(Y)|\mathcal{F}],$$

that is X and Y are conditionally negatively dependent under $\hat{\mathbb{E}}[\,\cdot\,|\mathcal{F}]$.

The following proposition is a straightforward extension of Proposition 2.4 of [35].

Proposition 3.6.3. Let $\{f_i\}_{i=1}^{n_1} \subset C_{l,\text{Lip}}(\mathbb{R}^n)$ and $\{g_i\}_{i=1}^{m_1} \subset C_{l,\text{Lip}}(\mathbb{R}^m)$ be coordinate-wise non-decreasing or coordinate-wise non-increasing functions. If the n-dimensional random vector Y is conditionally negatively dependent on the m-dimensional random vector X under conditional sub-additive expectation, then $(f_1(Y), \ldots, f_{n_1}(Y))$ is conditionally negatively dependent on $(g_1(X), \ldots, g_{m_1}(X))$. In particular, $-Y$ is conditionally negatively dependent on $-X$.

Proof. For any test functions $\psi_1 \in C_{l,\text{Lip}}(\mathbb{R}^{m_1})$ and $\psi_2 \in C_{l,\text{Lip}}(\mathbb{R}^{n_1})$ such that

$$\psi_1(g_1(X), \ldots, g_{m_1}(X)) \geq 0, \quad \hat{\mathbb{E}}[\psi_2(f_1(Y), \ldots, f_{n_1}(Y))|\mathcal{F}] \geq 0,$$

and

$$\hat{\mathbb{E}}\big[\big|\psi_1(g_1(X), \ldots, g_{m_1}(X))\, \psi_2(f_1(Y), \ldots, f_{n_1}(Y))\big|\big|\mathcal{F}\big] < \infty,$$
$$\hat{\mathbb{E}}[|\psi_1(g_1(X), \ldots, g_{m_1}(X))||\mathcal{F}] < \infty,$$
$$\hat{\mathbb{E}}[|\psi_2(f_1(Y), \ldots, f_{n_1}(Y))||\mathcal{F}] < \infty,$$

and either ψ_1, ψ_2 are coordinate-wise non-decreasing or both are coordinate-wise non-increasing, define

$$\tilde{\psi}_1(\cdot) = \psi_1(g_1(\cdot), \ldots, g_{m_1}(\cdot)), \qquad \tilde{\psi}_2(\cdot) = \psi_2(f_1(\cdot), \ldots, f_{n_1}(\cdot)).$$

Then $\tilde{\psi}_1 \in C_{l,\text{Lip}}(\mathbb{R}^m)$, $\tilde{\psi}_2 \in C_{l,\text{Lip}}(\mathbb{R}^n)$, and they are either coordinate-wise non-decreasing or coordinate-wise non-increasing. Meanwhile, $\tilde{\psi}_1(X) \geq 0$, $\hat{\mathbb{E}}[\tilde{\psi}_2(Y)|\mathcal{F}] \geq 0$, and the integrability conditions hold.

By the definition of negative dependence, we have

$$\begin{aligned}
\hat{\mathbb{E}}\,[\psi_1(g_1(X), \ldots, g_{m_1}(X))\, \psi_2(f_1(Y), \ldots, f_{n_1}(Y))|\,\mathcal{F}]& \\
= \hat{\mathbb{E}}[\tilde{\psi}_1(X)\tilde{\psi}_2(Y)|\mathcal{F}]& \\
\leq \hat{\mathbb{E}}[\tilde{\psi}_1(X)|\mathcal{F}]\, \hat{\mathbb{E}}[\tilde{\psi}_2(Y)|\mathcal{F}]& \\
= \hat{\mathbb{E}}[\psi_1(g_1(X), \ldots, g_{m_1}(X))|\mathcal{F}]\, \hat{\mathbb{E}}[\psi_2(f_1(Y), \ldots, f_{n_1}(Y))|\mathcal{F}]&.
\end{aligned}$$

Hence, $(f_1(Y), \ldots, f_{n_1}(Y))$ is negatively dependent on $(g_1(X), \ldots, g_{m_1}(X))$ under $\hat{\mathbb{E}}[\cdot|\mathcal{F}]$. □

Our first result is a Kolmogorov-type maximal inequality (actually it is a Rosenthal inequality). It was presented in [79] and in [35] for usual (i.e. not conditional) sub-additive expectation.

Theorem 3.6.4 (Masasila & Fazekas [50])**.** *Let $\hat{\mathbb{E}}[\cdot|\mathcal{F}]$ satisfy the axioms of monotonicity, sub-additivity, and positive homogeneity. Let $1 \leq p \leq 2$, and let $\{X_k, 1 \leq k \leq n\}$ be a sequence of random variables with $\hat{\mathbb{E}}[X_k|\mathcal{F}] = \hat{\mathbb{E}}[-X_k|\mathcal{F}] = 0$ q.s. for all $k = 1, \ldots, n$. If*

X_k is negatively dependent on $(X_{k+1},\dots,X_n)$ for all $k=1,\dots,n-1$, then

$$\hat{\mathbb{E}}\left[\max_{1\le k\le n}|S_k|^p\big|\mathcal{F}\right]\le 2^{3-p}\sum_{k=1}^{n}\hat{\mathbb{E}}[|X_k|^p|\mathcal{F}] \quad q.s.\ ,$$

where $S_k=\sum_{i=1}^{k}X_i$.

Proof. We follow the lines of [79]. Define,

$$T_k := \max\big\{X_k,\ X_k+X_{k+1},\ \dots,\ X_k+X_{k+1}+\cdots+X_n\big\},\qquad k=1,\dots,n.$$

We see that T_k is an increasing function of $X_k,X_{k+1},\dots,X_n$. Then it is clear that

$$T_k = X_k + T^+_{k+1},\quad \text{where}\quad T^+_{k+1}:=\max\{0,T_{k+1}\}, \tag{3.6.1}$$

and that at the terminal index

$$T_n = X_n. \tag{3.6.2}$$

Moreover, since $T_1=\max_{1\le j\le n}S_j$ with $S_j=X_1+\cdots+X_j$, we have

$$\hat{\mathbb{E}}\left[\big|\max_{1\le j\le n}S_j\big|^p\big|\mathcal{F}\right]=\hat{\mathbb{E}}[|T_1|^p|\mathcal{F}].$$

The following elementary inequality can be found e.g. in [79], one can prove it by Taylor's expansion. For $1\le p\le 2$ and any $x,y\in\mathbb{R}$,

$$|x+y|^p\le 2^{2-p}|x|^p+|y|^p+p\,x\,|y|^{p-1}\operatorname{sgn}(y). \tag{3.6.3}$$

Setting $x=X_k$ and $y=T^+_{k+1}$ in (3.6.3) gives

$$|T_k|^p=|X_k+T^+_{k+1}|^p\le 2^{2-p}|X_k|^p+(T^+_{k+1})^p+p\,X_k(T^+_{k+1})^{p-1}.$$

Taking the conditional sub-additive expectation $\hat{\mathbb{E}}[\,\cdot\,|\mathcal{F}]$ of both sides yields

$$\hat{\mathbb{E}}[|T_k|^p|\mathcal{F}]\le 2^{2-p}\hat{\mathbb{E}}[|X_k|^p|\mathcal{F}]+\hat{\mathbb{E}}[(T^+_{k+1})^p|\mathcal{F}]+p\,\hat{\mathbb{E}}[X_k(T^+_{k+1})^{p-1}|\mathcal{F}]\quad \text{q.s.} \tag{3.6.4}$$

Since X_k is negatively dependent on $(X_{k+1},\dots,X_n)$ and T^+_{k+1} is a coordinate-wise non-decreasing function of these variables, we have by definition of negative dependence

$$\hat{\mathbb{E}}[X_k(T^+_{k+1})^{p-1}|\mathcal{F}]\le\hat{\mathbb{E}}[X_k|\mathcal{F}]\,\hat{\mathbb{E}}[(T^+_{k+1})^{p-1}|\mathcal{F}]\le 0,$$

because $\hat{\mathbb{E}}[X_k|\mathcal{F}]=0$ and $T^+_{k+1}\ge 0$. Consequently, inequality (3.6.4) simplifies to

$$\hat{\mathbb{E}}[|T_k|^p|\mathcal{F}]\le 2^{2-p}\hat{\mathbb{E}}[|X_k|^p|\mathcal{F}]+\hat{\mathbb{E}}[|T_{k+1}|^p|\mathcal{F}]. \tag{3.6.5}$$

Applying (3.6.5) successively for $k = n-1, n-2, \ldots, 1$ and using (3.6.2), we obtain

$$\begin{aligned}\hat{\mathbb{E}}[|T_1|^p|\mathcal{F}] &\le 2^{2-p}\hat{\mathbb{E}}[|X_1|^p|\mathcal{F}] + \hat{\mathbb{E}}[|T_2|^p|\mathcal{F}] \\ &\le 2^{2-p}\big(\hat{\mathbb{E}}[|X_1|^p|\mathcal{F}] + \hat{\mathbb{E}}[|X_2|^p|\mathcal{F}]\big) + \hat{\mathbb{E}}[|T_3|^p|\mathcal{F}] \\ &\;\vdots \\ &\le 2^{2-p}\sum_{k=1}^{n-1}\hat{\mathbb{E}}[|X_k|^p|\mathcal{F}] + \hat{\mathbb{E}}[|T_n|^p|\mathcal{F}] \\ &= 2^{2-p}\sum_{k=1}^{n}\hat{\mathbb{E}}[|X_k|^p|\mathcal{F}],\end{aligned}$$

where the last equality uses $T_n = X_n$. Since $\max_{1\le k\le n} S_k = T_1$, then, for $p \ge 1$, we get

$$\hat{\mathbb{E}}\left[\left|\max_{1\le k\le n} S_k\right|^p \Big|\mathcal{F}\right] = \hat{\mathbb{E}}[|T_1|^p|\mathcal{F}] \le 2^{2-p}\sum_{k=1}^{n}\hat{\mathbb{E}}[|X_k|^p|\mathcal{F}] \quad \text{q.s.}$$

By Proposition 3.6.3, $-X_k$ is negatively dependent on $(-X_{k+1}, \ldots, -X_n)$ for all $k = 1, \ldots, n-1$. Hence,

$$\hat{\mathbb{E}}\left[\left|\max_{1\le k\le n}(-S_k)\right|^p \Big|\mathcal{F}\right] \le 2^{2-p}\sum_{k=1}^{n}\hat{\mathbb{E}}[|X_k|^p|\mathcal{F}].$$

Observe that

$$\max_{1\le k\le n}|S_k|^p = \big(\max_{1\le k\le n}|S_k|\big)^p \le \left|\max_{1\le k\le n} S_k\right|^p + \left|\max_{1\le k\le n}(-S_k)\right|^p.$$

Taking $\hat{\mathbb{E}}[\,\cdot\,|\mathcal{F}]$ on both sides and using its monotonicity and sub-additivity, we get

$$\hat{\mathbb{E}}\left[\max_{1\le k\le n}|S_k|^p\Big|\mathcal{F}\right] \le 2\cdot 2^{2-p}\sum_{k=1}^{n}\hat{\mathbb{E}}[|X_k|^p|\mathcal{F}] = 2^{3-p}\sum_{k=1}^{n}\hat{\mathbb{E}}[|X_k|^p|\mathcal{F}],$$

which completes the proof. □

In the following theorem we assume that $\hat{\mathbb{V}}[\,\cdot\,|\mathcal{F}]$ is the conditional capacity induced by $\hat{\mathbb{E}}[\,\cdot\,|\mathcal{F}]$.

Theorem 3.6.5 (Masasila & Fazekas [50]). *Let $\hat{\mathbb{E}}[\,\cdot\,|\mathcal{F}]$ satisfy the axioms of monotonicity, sub-additivity, and positive homogeneity. Let $\hat{\mathbb{V}}[A|\mathcal{F}] = \hat{\mathbb{E}}[\mathbb{I}_A|\mathcal{F}]$ for any event A. Let $1 \le p \le 2$, and let $\{X_k, 1 \le k \le n\}$ be a sequence of random variables with $\hat{\mathbb{E}}[X_k|\mathcal{F}] = \hat{\mathbb{E}}[-X_k|\mathcal{F}] = 0$ for all $l = 1, \ldots, n$. If X_k is negatively dependent on $(X_{k+1}, \ldots, X_n)$ for all*

$k = 1, \ldots, n-1$, *then, for any* $\varepsilon > 0$,

$$\hat{\mathbb{V}}\left[\max_{1\le k\le n} |S_k| > \varepsilon \big| \mathcal{F}\right] \le 2^{3-p}\,\varepsilon^{-p}\sum_{k=1}^{n}\hat{\mathbb{E}}[|X_k|^p|\mathcal{F}], \quad q.s., \tag{3.6.6}$$

where $S_k = \sum_{i=1}^{k} X_i$.

Proof. By using Chebyshev inequality under conditional sub-additive expectation, we have

$$\hat{\mathbb{V}}\left[\max_{1\le k\le n} |S_k| > \varepsilon \big| \mathcal{F}\right] \le \varepsilon^{-p}\,\hat{\mathbb{E}}\left[\max_{1\le k\le n} |S_k|^p \big| \mathcal{F}\right].$$

Then inequality (3.6.6) follows from Theorem 3.6.4 directly. □

Theorem 3.6.6 (Masasila & Fazekas [50]). *Assume that the conditional expectation operator* $\hat{\mathbb{E}}[\,\cdot\,|\mathcal{F}]$ *satisfies the axioms in Definition 3.3.1, where all relations among random variables are understood in the* $\hat{\mathbb{V}}$*-quasi-sure sense. Assume that the finiteness axiom in Definition 3.3.6 is also satisfied. Let* $1 \le p \le 2$*, and let* $\{X_n, n \ge 1\}$ *be a sequence of random variables with* $\hat{\mathbb{E}}[X_n|\mathcal{F}] = \hat{\mathbb{E}}[-X_n|\mathcal{F}] = 0$ *for all* $n \in \mathbb{N}$*. Let* $b_1, b_2, \ldots$ *be q.s. finite,* $\mathcal{F}$*-measurable random variables with* $b_0 \le b_1 \le b_2 \le \ldots$ *q.s.,* $b_n \to \infty$ *q.s., where* b_0 *is a positive constant. If* X_k *is negatively dependent on* $(X_{k+1}, \ldots, X_{k+n})$ *for all* $n, k \in \mathbb{N}$ *with*

$$\sum_{n=1}^{\infty}\frac{\hat{\mathbb{E}}[|X_n|^p|\mathcal{F}]}{b_n^p} < \infty \quad q.s.,$$

then

$$\lim_{n\to\infty}\frac{S_n}{b_n} = 0 \quad q.s.$$

Proof. Set $\alpha_k = \hat{\mathbb{E}}[|X_k|^p|\mathcal{F}]$ for all $k \in \mathbb{N}$. Then, by Theorem 3.6.4, we have

$$\hat{\mathbb{E}}\left[\max_{1\le k\le n} |S_k|^p \big| \mathcal{F}\right] \le 2^{3-p}\sum_{k=1}^{n}\alpha_k, \quad n \ge 1.$$

Since

$$\sum_{n=1}^{\infty}\frac{\alpha_n}{b_n^p} = \sum_{n=1}^{\infty}\frac{\hat{\mathbb{E}}[|X_n|^p|\mathcal{F}]}{b_n^p} < \infty,$$

we can deduce the theorem from Theorem 3.5.1 directly. □

3.7 Discussion

In the usual theory of probability, a probability space $(\Omega, \mathcal{A}, P)$ is a measure space, and the expectation of a random variable X is its Lebesgue integral: $\mathbb{E}X = \int_\Omega X dP$. The conditional expectation and the conditional probability given a sub-σ-field are random variables that can be obtained via the Radon-Nikodym theorem.

To introduce sub-linear expectation and sub-additive probability, one can start with several usual probability measures and use supremum to define these quantities; see Examples 3.2.2 and 3.2.4. Similarly, to introduce sub-linear conditional expectation and sub-additive conditional probability, one can use the essential supremum of usual conditional expectations and usual conditional probabilities, see Examples 3.3.2 and 3.3.5.

However, in this paper, we followed another approach. We aimed to find small sets of properties of the sub-linear conditional expectation and sub-additive conditional probability and fix them as axioms. Then, during the proofs, we applied only these properties of the sub-linear conditional expectation and the sub-additive conditional probability.

Below, we visualize the main implications in the chapter. For sub-linear conditional expectation, we proved that a Kolmogorov-type maximal inequality implies a Hájek-Rényi-type maximal inequality, and this implies an SLLN:

$$\{\text{Definition 3.3.1}\} \quad \Longrightarrow \quad \{\text{Theorem 3.4.1}\} \quad \Longrightarrow \quad \{\text{Theorem 3.5.1}\}$$

Similarly, for sub-additive conditional probability, we proved that a Kolmogorov-type maximal inequality implies a Hájek-Rényi-type maximal inequality, and this implies an SLLN:

$$\{\text{Definition 3.3.3}\} \quad \Longrightarrow \quad \{\text{Theorem 3.4.3}\} \quad \Longrightarrow \quad \{\text{Theorem 3.5.2}\}$$

Our assumptions were similar to those in the unconditional case, so our conclusions are also similar to that case, see [35].

For conditionally negatively dependent random variables, we obtained a conditional Kolmogorov-type maximal inequality, leading to a conditional strong law of large numbers:

$$\{\text{Theorem 3.6.4}\} \quad \& \quad \{\text{Theorem 3.5.1}\} \quad \Longrightarrow \quad \{\text{Theorem 3.6.6}\}$$

Our SLLNs are extensions of known SLLNs to sub-linear conditional expectation and sub-additive conditional probability frameworks.

Chapter 4

Strong law of large numbers for φ-sub-Gaussian random variables under sub-linear expectation spaces

In this chapter, we introduce the notions of sub-Gaussian and φ-sub-Gaussian random variables in sub-linear expectation spaces. To avoid the problem caused by the existence of two different expectations, i.e., the upper expectation and the lower expectation, we divide the definition of the sub-Gaussian property into an upper part and a lower part. It turns out that this approach fits well to the sub-linear setting; it provides a proper framework for extending general result of Zajkowski [77] to sublinear expectation spaces. Within our framework, we establish a strong law of large numbers for sub-Gaussian sequences. We present an example showing the usefulness of our results.

4.1 Introduction

In the usual probability framework, the strong law of large numbers (SLLN) relies on linearity of expectation and additivity of probability measure. In many modern applications, including risk measures, robust finance, and models with ambiguity, these assumptions may fail. This has motivated the development of *sub-linear expectation spaces*, in which expectations are sub-linear functionals and probabilities (capacities) are sub-additive. In this setting, the SLLN takes another form: it asserts that every cluster point of the sequence of empirical averages lies between the lower and upper expectations, with lower capacity equal to 1. Formally,

$$v\left(\omega \in \Omega : \mathcal{E}(X_1) \le \liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k(\omega) \le \limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k(\omega) \le \hat{\mathbb{E}}(X_1)\right) = 1, \tag{4.1.1}$$

where $\hat{\mathbb{E}}$ denotes the upper (sub-linear) expectation, $\mathcal{E}$ the lower (conjugate) expectation, and v the corresponding lower capacity.

Building on this framework, numerous authors have investigated versions of the strong law of large numbers (SLLN) under various sets of assumptions. Particularly, in [48] and [45], an SLLN was established for bounded and continuous random variables under a totally monotone capacity. Subsequently, Chen, Wu, and Li in [9] proved an SLLN for independent random variables assuming finite $(1+\alpha)^{th}$ moments with respect to upper expectation, while Chen [8] further relaxed the framework by removing the continuity assumption on the upper capacity.

Sub-Gaussian distributions were studied in [38]. Then it was used to prove laws of large numbers (see, e.g., [12], [71] and [7]); see also [2] for related results. Recently, [77] developed a general strong law for φ-sub-Gaussian sequences under linear expectations. The purpose of the present chapter is to extend this framework to sub-linear expectation spaces.

4.2 Basic Properties of Capacities and Sub-linear Expectations

Based on the classical results of [11] and subsequent developments by [36], we summarize the properties of capacities. Let Ω be a non-empty set and let $\mathcal{F}$ be a σ-algebra of subsets of Ω.

Definition 4.2.1. $\hat{\mathbb{V}}$ is called a sub-additive probability (upper probability or upper capacity), if

1. (Normalized.) $\hat{\mathbb{V}}(\Omega) = 1$, $\hat{\mathbb{V}}(\emptyset) = 0$.
2. (Monotone.) If $A, B \in \mathcal{F}$ and $A \subset B$, then $\hat{\mathbb{V}}(A) \le \hat{\mathbb{V}}(B)$.
3. (Sub-additive.) If $A, B \in \mathcal{F}$, then

$$\hat{\mathbb{V}}(A \cup B) \le \hat{\mathbb{V}}(A) + \hat{\mathbb{V}}(B).$$

4. (Lower-continuous.) If $A_n \uparrow A$, $A_n \in \mathcal{F}$, then $\hat{\mathbb{V}}(A_n) \uparrow \hat{\mathbb{V}}(A)$.

Sub-additivity and lower-continuity imply σ-sub-additivity: For any sequence $\{A_n\}_{n=1}^{\infty} \subset \mathcal{F}$,

$$\hat{\mathbb{V}}\left(\bigcup_{n=1}^{\infty} A_n\right) \le \sum_{n=1}^{\infty} \hat{\mathbb{V}}(A_n).$$

An event A is called a quasi-sure (q.s.) event if $\hat{\mathbb{V}}(A^c) = 0$, where A^c denotes the complement of the event A in Ω.

The lower capacity corresponding to $\hat{\mathbb{V}}$ can be expressed in terms of $\hat{\mathbb{V}}$ as $v(A) = 1 - \hat{\mathbb{V}}(A^c)$. The set functions $\hat{\mathbb{V}}$ and v form a pair of conjugate capacities.

Example 4.2.2. The basic method for obtaining a sub-additive probability is as follows. Let $\mathcal{P}$ be a family of usual (that is, σ-additive) probabilities on $(\Omega, \mathcal{F})$. Introduce $\hat{\mathbb{V}}$ as

$$\hat{\mathbb{V}}(A) = \sup_{Q \in \mathcal{P}} Q(A), \quad \forall A \in \mathcal{F}. \tag{4.2.1}$$

It is easy to see that $\hat{\mathbb{V}}$ satisfies the properties listed in Definition 4.2.1. The corresponding lower capacity is

$$v(A) = \inf_{Q \in \mathcal{P}} Q(A), \quad \forall A \in \mathcal{F}.$$

Standard results such as Chebyshev's inequality and the Borel–Cantelli lemma remain valid in this framework (see, [9] and [58]).

Now, we list basic properties of the sub-linear expectation $\hat{\mathbb{E}}$; see [58] for details. We mention that, in certain papers, $\hat{\mathbb{E}}$ is called the upper expectation or sub-additive expectation. As usual, an extended real-valued function X on Ω is called a random variable if $X^{-1}(A) \in \mathcal{F}$ for any Borel set A. We assume, that there exists a subset $\mathcal{H}$ of the random variables and an extended real-valued function $\hat{\mathbb{E}}[X]$ of $X \in \mathcal{H}$ so that the assumptions of Definition 4.2.3 are satisfied. We shall suppose that the non-negative random variables belong to $\mathcal{H}$. In Definition 4.2.3, the case $\hat{\mathbb{E}}[X] + \hat{\mathbb{E}}[Y] = \infty + (-\infty)$ is excluded.

Definition 4.2.3. An extended real-valued function $\hat{\mathbb{E}}[X]$ of $X \in \mathcal{H}$ is called a sub-linear expectation if it satisfies the following properties.

1. Monotone: If $X \leq Y$, then $\hat{\mathbb{E}}[X] \leq \hat{\mathbb{E}}[Y]$.
2. Constant preserving: $\hat{\mathbb{E}}[c] = c$, for any $c \in \mathbb{R}$.
3. Sub-additive: $\hat{\mathbb{E}}[X + Y] \leq \hat{\mathbb{E}}[X] + \hat{\mathbb{E}}[Y]$.
4. Positive homogeneous: $\hat{\mathbb{E}}[\lambda X] = \lambda \hat{\mathbb{E}}[X]$ for any constant $\lambda \geq 0$.
5. Monotone convergence: If $X_n \uparrow X$, and $X_1 \geq 0$, then $\hat{\mathbb{E}}[X_n] \uparrow \hat{\mathbb{E}}[X]$.

$(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ is called a sub-linear expectation space.

If the sub-linear expectation $\hat{\mathbb{E}}[\,.\,]$ is given in advance, then we can introduce a sub-additive probability $\hat{\mathbb{V}}$ by $\hat{\mathbb{V}}(A) = \hat{\mathbb{E}}[\mathbb{I}_A]$, for any event A, where $\mathbb{I}_A$ denotes the indicator of A. Then this $\hat{\mathbb{V}}$ satisfies the properties given in Definition 4.2.1.

Throughout this paper, we assume that an upper probability $\hat{\mathbb{V}}$ is given on $(\Omega, \mathcal{F})$, and a sub-linear expectation is given on $\mathcal{H}$ so that $\hat{\mathbb{V}}(A) = \hat{\mathbb{E}}[\mathbb{I}_A]$. Relations between random variables are considered as quasi-sure relations, e.g., $\hat{\mathbb{E}}[X] = \hat{\mathbb{E}}[Y]$ if $X = Y$ q.s.

Example 4.2.4. The usual method to define a sub-additive expectation is as follows. Let $\mathcal{P}$ be a family of probabilities on $(\Omega, \mathcal{F})$. Let

$$\hat{\mathbb{E}}[X] = \sup_{Q \in \mathcal{P}} \mathbb{E}_Q[X], \tag{4.2.2}$$

where $\mathbb{E}_Q$ is the usual expectation corresponding to $Q \in \mathcal{P}$ such that $\mathbb{E}_Q[\mathbb{I}_A] = Q(A), \quad \forall A \in \mathcal{F}$. Then, properties listed in Definition 4.2.3 are satisfied.

We remark that if an abstract sub-linear expectation $\hat{\mathbb{E}}[\,.\,]$ is given with the properties listed in Definition 4.2.3, and certain additional conditions are satisfied, then $\hat{\mathbb{E}}$ has a representation (4.2.2). For the precise statement, see [17] and [57].

For $X \in \mathcal{H}$, $\hat{\mathbb{E}}(X)$ can be called supermean, whereas $\mathcal{E}(X) = -\hat{\mathbb{E}}(-X)$ is called sub-mean. By the above properties of $\hat{\mathbb{E}}(X)$, we have $\mathcal{E}(X) \le \hat{\mathbb{E}}(X)$. If $\mathcal{E}(X) \ne -\hat{\mathbb{E}}(X)$, then X is said to have mean uncertainty.

4.3 Sub-Gaussian Random Variables under Sub-linear Expectations

For a usual probability space $(\Omega, \mathcal{F}, P)$, a random variable $X : \Omega \to \mathbb{R}$ is called *sub-Gaussian* if there exists a constant $a \in [0, \infty)$ such that

$$\mathbb{E}_P\left[e^{\lambda X}\right] \le \exp\left(\frac{a^2\lambda^2}{2}\right), \qquad \text{for all } \lambda \in \mathbb{R}. \tag{4.3.1}$$

Condition (4.3.1) shows that a random variable is sub-Gaussian if and only if its moment generating function is dominated by that of a centered Gaussian random variable with variance parameter a^2. The term *sub-Gaussian* reflects this Gaussian-type exponential moment control rather than any exact distributional identity. We see that condition (4.3.1) holds only for zero-mean random variables. For sub-linear expectation, we shall replace it with two one-sided conditions. More generally, we shall do it for φ-sub-Gaussian property.

Therefore, we shall introduce the notion of φ-sub-Gaussian random variables for sub-linear expectation spaces. For usual probability spaces, this was studied e.g., in [77]. We call a real continuous even convex function $\varphi(x)$, $x \in \mathbb{R}$, an N-function if the following conditions are satisfied:

(a) $\varphi(0) = 0$ and $\varphi(x)$ is monotone increasing for $x > 0$,

(b) $\lim_{x\to 0} \frac{\varphi(x)}{x} = 0$ and $\lim_{x\to\infty} \frac{\varphi(x)}{x} = \infty$.

An N-function φ is called a quadratic N-function if, in addition, $\varphi(x) = cx^2$ for all $|x| \le x_0$ with $c > 0$ and $x_0 > 0$.

In our results, we shall use the following quadratic N-function, see [77].

Definition 4.3.1. For $p \geq 1$, let

$$\varphi_p(x) = \begin{cases} \dfrac{x^2}{2}, & \text{if } |x| \leq 1, \\ \dfrac{1}{p}|x|^p - \dfrac{1}{p} + \dfrac{1}{2}, & \text{if } |x| > 1. \end{cases}$$

The function φ_p can be considered as a standardization of the function $|x|^p$. We shall see that, for $p = 2$, we have the case of the usual sub-Gaussian random variables.

Definition 4.3.2. Let φ be a quadratic N-function. A random variable ξ is said to be φ-sub-Gaussian under the sub-linear expectation $\hat{\mathbb{E}}$ if there exist constants $a > 0$, $\overline{m}$, and $\underline{m}$, such that

$$\hat{\mathbb{E}}e^{\lambda(\xi - \overline{m})} \leq e^{\varphi(a\lambda)}, \quad \text{for} \quad \lambda > 0, \tag{4.3.2}$$

and

$$\hat{\mathbb{E}}e^{\lambda(\xi - \underline{m})} \leq e^{\varphi(a\lambda)}, \quad \text{for} \quad \lambda < 0. \tag{4.3.3}$$

$a > 0$, $\overline{m}$, and $\underline{m}$ are parameters of the φ-sub-Gaussian variable.

For the upper bound, we shall use convex conjugate. For any real function $\varphi(x)$, $x \in \mathbb{R}$, the function $\varphi^*(y)$, $y \in \mathbb{R}$, is called the convex conjugate of φ, if $\varphi^*(y) = \sup_{x \in \mathbb{R}}\{xy - \varphi(x)\}$.

The following properties are known for the convex conjugate, see e.g. [76]. If φ is a quadratic N-function, then φ^* is also a quadratic N-function. For $p, q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$, we have $\varphi_p^* = \varphi_q$. The convex conjugation is order-reversing, i.e. if $\varphi_1 \geq \varphi_2$, then $\varphi_1^* \leq \varphi_2^*$. The following scaling property holds. For $a > 0$ and $b \neq 0$, let $\psi(x) = a\varphi(bx)$. Then $\psi^*(y) = a\varphi^*\left(\frac{y}{ab}\right)$.

Lemma 4.3.3. Assume that (4.3.2) and (4.3.3) are satisfied. Then for $\varepsilon > 0$ we have

$$\hat{\mathbb{V}}\left(\{\xi - \overline{m} > \varepsilon\} \cup \{\xi - \underline{m} < -\varepsilon\}\right) \leq 2e^{-\varphi^*(\varepsilon/a)}. \tag{4.3.4}$$

Proof. Let $\varepsilon > 0$ and $\lambda > 0$. Then, by the Chebyshev inequality and the φ-sub-Gaussian property, we have

$$\hat{\mathbb{V}}(\xi - \overline{m} > \varepsilon) = \hat{\mathbb{V}}\left(e^{\lambda(\xi - \overline{m})} > e^{\lambda\varepsilon}\right) \leq \frac{1}{e^{\lambda\varepsilon}}\hat{\mathbb{E}}\left(e^{\lambda(\xi - \overline{m})}\right) \tag{4.3.5}$$

$$\leq \frac{1}{e^{\lambda\varepsilon}}e^{\varphi(a\lambda)} = e^{-(\lambda\varepsilon - \varphi(a\lambda))}. \tag{4.3.6}$$

To find the optimal upper bound that is offered by the above inequality, we use convex conjugate. So we obtain

$$\hat{\mathbb{V}}(\xi - \overline{m} > \varepsilon) \leq e^{-\varphi^*(\varepsilon/a)}, \quad \text{for } \varepsilon > 0. \tag{4.3.7}$$

For the other inequality, let $\varepsilon < 0$ and $\lambda < 0$. Then, by the Chebyshev inequality and the

φ-sub-Gaussian property, we have

$$\hat{\mathbb{V}}(\xi - \underline{m} < \varepsilon) \leq \frac{1}{e^{\lambda\varepsilon}} \hat{\mathbb{E}}\left(e^{\lambda(\xi - \underline{m})}\right) \leq e^{-(\lambda\varepsilon - \varphi(a\lambda))}. \tag{4.3.8}$$

From this inequality, the optimal upper bound is

$$\hat{\mathbb{V}}(\xi - \overline{m} < \varepsilon) \leq e^{-\varphi^*(\varepsilon/a)}, \quad \text{for } \varepsilon < 0. \tag{4.3.9}$$

As φ^* is an even function, we obtain the result. □

To obtain the optimal value of the parameter a, we need the following quantity.

Definition 4.3.4. For a φ-sub-Gaussian random variable ξ with fixed $\overline{m}$ and $\underline{m}$ let $\tau_\varphi(\xi)$ be defined as

$$\tau_\varphi(\xi) = \inf\Big\{a \geq 0 \,:\, \hat{\mathbb{E}} e^{\lambda(\xi - \overline{m})} \leq e^{\varphi(a\lambda)}, \text{ for } \lambda > 0,$$
$$\text{and } \hat{\mathbb{E}} e^{\lambda(\xi - \underline{m})} \leq e^{\varphi(a\lambda)}, \text{ for } \lambda < 0\Big\}.$$

4.4 The Main Result

The following theorem is an extension of Theorem 2.1 of [77] to sub-linear expectation.

Theorem 4.4.1 (Masasila & Fazekas [49]). *For some $p > 1$, let $\{Z_n, n \geq 1\}$, be a sequence of φ_p-sub-Gaussian random variables with parameters $\overline{m}$ and $\underline{m}$. Let $\tau_\varphi(Z_n)$ be defined according to Definition 4.3.4. If there exist positive numbers c and α such that for every natural number n, the condition $\tau_{\varphi_p}(Z_n) \leq c n^{-\alpha}$ is satisfied, then*

$$\hat{\mathbb{V}}\left(\left\{\liminf_{n\to\infty} Z_n < \underline{m}\right\} \cup \left\{\limsup_{n\to\infty} Z_n > \overline{m}\right\}\right) = 0 \tag{4.4.1}$$

and

$$v\left(\underline{m} \leq \liminf_{n\to\infty} Z_n \leq \limsup_{n\to\infty} Z_n \leq \overline{m}\right) = 1. \tag{4.4.2}$$

Proof. Since $V(\cdot)$ and $v(\cdot)$ are conjugate to each other, (4.4.2) is equivalent to (4.4.1).

By Lemma 4.3.3, for $\varepsilon > 0$ we have

$$\hat{\mathbb{V}}\left(\{Z_n - \overline{m} > \varepsilon\} \cup \{Z_n - \underline{m} < -\varepsilon\}\right) \leq 2\exp\left(-\varphi_q\left(\frac{\varepsilon}{\tau_{\varphi_p}(Z_n)}\right)\right), \tag{4.4.3}$$

where $1/p + 1/q = 1$ and we applied that $\varphi_q = \varphi_p^*$. By the condition $\tau_{\varphi_p}(Z_n) \leq c n^{-\alpha}$, for all sufficiently large n we have $\frac{\varepsilon}{\tau_{\varphi_p}(Z_n)} > 1$, so using the definition of φ_q, we obtain that the right-hand side of inequality (4.4.3) is majorized by $C_0 \exp(-K n^{q\alpha})$, where $C_0 = 2\exp\left(\frac{1}{q} - \frac{1}{2}\right)$, and $K = \frac{1}{q}\left(\frac{\varepsilon}{c}\right)^q$.

We now show that the series

$$\sum_{n=1}^{\infty} \hat{\mathbb{V}}\left(\{Z_n - \overline{m} > \varepsilon\} \cup \{Z_n - \underline{m} < -\varepsilon\}\right)$$

is convergent. It suffices to prove the convergence of

$$\sum_{n=1}^{\infty} \exp\bigl(-Kn^{\beta}\bigr), \qquad \text{where } \ \beta = q\alpha > 0.$$

The function $f(x) = \exp(-Kx^{\beta})$ is positive and eventually decreasing. By the integral test,

$$\sum_{n=1}^{\infty} \exp(-Kn^{\beta}) \leq \int_0^{\infty} \exp(-Kx^{\beta})\, dx.$$

Performing the substitution $t = Kx^{\beta}$, we obtain

$$\int_0^{\infty} \exp(-Kx^{\beta})\, dx = \frac{1}{\beta} K^{-1/\beta} \int_0^{\infty} t^{1/\beta - 1} e^{-t}\, dt$$
$$= \frac{1}{\beta} K^{-1/\beta} \Gamma\left(\frac{1}{\beta}\right) < \infty.$$

Now, by the Borel-Cantelli lemma, it follows that the $\hat{\mathbb{V}}$ measure of event that infinitely many of the events

$$\left(\{Z_n - \overline{m} > \varepsilon\} \cup \{Z_n - \underline{m} < -\varepsilon\}\right)$$

occur is 0. As it is true for any $\varepsilon > 0$, so (4.4.1) is true. □

4.5 Strong law of large numbers for independent sub-Gaussian variables

Our aim is to obtain a strong law of large numbers for independent identically distributed sub-Gaussian random variables. However, our Theorem 4.5.1 is true for more general situation. In Example 4.5.2, we construct a plausible model fitting to Theorem 4.5.1.

We shall use the independence notion given in [9]. The sequence of random variables $\xi_1, \xi_2, \ldots$ is called independent if for each $n = 1, 2, \ldots$ and each non-negative measurable functions $f_1, f_2, \ldots$ we have

$$\hat{\mathbb{E}}(f_1(\xi_1) f_2(\xi_2) \cdots f_n(\xi_n)) = \hat{\mathbb{E}}(f_1(\xi_1)) \hat{\mathbb{E}}(f_2(\xi_2)) \cdots \hat{\mathbb{E}}(f_n(\xi_n)).$$

If $\xi_1, \xi_2, \ldots$ are independent, then they satisfy the following property

$$\hat{\mathbb{E}} \prod_{i=1}^{k} \exp\left(\lambda(\xi_i - m)\right) \leq \prod_{i=1}^{k} \hat{\mathbb{E}} \exp\left(\lambda(\xi_i - m)\right) \tag{4.5.1}$$

for any real λ, m, and positive integer k. Now, let $\xi_1, \xi_2, \ldots$ be random variables satisfying the sub-Gaussian property (that is they are φ_2-sub-Gaussian): for fixed constants $\sigma > 0$, $\overline{m}$, and $\underline{m}$

$$\hat{\mathbb{E}}(e^{\lambda(\xi_i - \overline{m})}) \leq e^{(\sigma^2 \lambda^2/2)}, \quad \text{for} \quad \lambda > 0, \tag{4.5.2}$$

and

$$\hat{\mathbb{E}}(e^{\lambda(\xi_i - \underline{m})}) \leq e^{(\sigma^2 \lambda^2/2)}, \quad \text{for} \quad \lambda < 0. \tag{4.5.3}$$

Now, let $S_n = \xi_1 + \cdots + \xi_n$ for any positive integer n. We are ready to prove a strong law of large numbers.

Theorem 4.5.1 (Masasila & Fazekas [49]). *Let* $\{\xi_n, n \geq 1\}$ *be random variables satisfying* (4.5.1), (4.5.2)*, and* (4.5.3). *With* S_n *defined above,*

$$\hat{\mathbb{V}}\left(\left\{\liminf_{n\to\infty} \frac{S_n}{n} < \underline{m}\right\} \cup \left\{\limsup_{n\to\infty} \frac{S_n}{n} > \overline{m}\right\}\right) = 0, \tag{4.5.4}$$

or equivalently,

$$v\left(\underline{m} \leq \liminf_{n\to\infty} \frac{S_n}{n} \leq \limsup_{n\to\infty} \frac{S_n}{n} \leq \overline{m}\right) = 1.$$

Proof. We shall apply Theorem 4.4.1 with $Z_n = \frac{S_n}{n}$. Using the the sub-Gaussian property of ξ_i and (4.5.1),

$$\begin{aligned}
\hat{\mathbb{E}} \exp(\lambda(Z_n - \overline{m})) &= \hat{\mathbb{E}} \exp\left(\tfrac{\lambda}{n} \textstyle\sum_{i=1}^{n} (\xi_i - \overline{m})\right) \\
&= \hat{\mathbb{E}} \textstyle\prod_{i=1}^{n} \exp\left(\tfrac{\lambda}{n}(\xi_i - \overline{m})\right) \\
&\leq \textstyle\prod_{i=1}^{n} \hat{\mathbb{E}} \exp\left(\tfrac{\lambda}{n}(\xi_i - \overline{m})\right) \\
&\leq \textstyle\prod_{i=1}^{n} \exp\left(\left(\tfrac{\lambda}{n}\right)^2 \left(\tfrac{\sigma^2}{2}\right)\right) \\
&= \exp\left(\tfrac{\lambda^2}{n} \left(\tfrac{\sigma^2}{2}\right)\right)
\end{aligned}$$

for $\lambda > 0$. Similarly,

$$\begin{aligned}
\hat{\mathbb{E}} \exp(\lambda(Z_n - \underline{m})) &= \hat{\mathbb{E}} \textstyle\prod_{i=1}^{n} \exp\left(\tfrac{\lambda}{n}(\xi_i - \underline{m})\right) \\
&\leq \textstyle\prod_{i=1}^{n} \hat{\mathbb{E}} \exp\left(\tfrac{\lambda}{n}(\xi_i - \underline{m})\right) \\
&\leq \textstyle\prod_{i=1}^{n} \exp\left(\left(\tfrac{\lambda}{n}\right)^2 \left(\tfrac{\sigma^2}{2}\right)\right) \\
&= \exp\left(\tfrac{\lambda^2}{n} \left(\tfrac{\sigma^2}{2}\right)\right)
\end{aligned}$$

for $\lambda < 0$.

So in Theorem 4.4.1, $\tau_{\varphi_2}(Z_n) \leq \sigma n^{-(1/2)}$ and it implies the result. $\square$

We see, that Theorem 4.5.1 is valid for independent identically distributed sub-Gaussian random variables.

In the following example, we shall use the well-known fact that for a random variable X having normal distribution with expectation m and variance $\sigma^2 > 0$

$$\mathbb{E}e^{\lambda X} = e^{\frac{\lambda^2\sigma^2}{2} + \lambda m}.$$

Example 4.5.2. Let Ω be the real line and let $\mathcal{F}$ be the family of its Borel sets. Let M be an arbitrary non-empty bounded set of real numbers, let $\underline{m} = \inf(M)$ and $\overline{m} = \sup(M)$. Let $\sigma > 0$ be fixed. Let P_m denote the normal distribution with expectation m and variance σ^2. Then $(\Omega, \mathcal{F}, P_m)$ is a usual probability space for any $m \in M$. Let $\mathbb{E}_m X = \int_\Omega X dP_m$ be the usual expectation of the random variable X. Let ξ be the identity map: $\xi(\omega) = \omega$ for any $\omega \in \Omega$ (here Ω is the real line). Then ξ has a normal distribution with mean m and variance σ^2.

Now, define the upper expectation as $\hat{\mathbb{E}}X = \sup\{\mathbb{E}_m X : m \in M\}$ and the upper probability as $\hat{\mathbb{V}}(A) = \sup\{P_m(A) : m \in M\}$. Then $\hat{\mathbb{E}}\xi = \overline{m}$ and for the lower expectation $\mathcal{E}\xi = \underline{m}$.

For $\lambda > 0$, from equation

$$\mathbb{E}_m e^{\lambda\xi - \overline{m}} = e^{\frac{\lambda^2\sigma^2}{2} + \lambda(m - \overline{m})}$$

we see that $\hat{\mathbb{E}}e^{\lambda\xi - \overline{m}} = e^{\frac{\lambda^2\sigma^2}{2}}$, so condition (4.5.2) is satisfied, we have equality there, and $\overline{m}$ is the optimal constant in that condition. Similarly, for $\lambda < 0$, we can see that $\hat{\mathbb{E}}e^{\lambda X - \underline{m}} = e^{\frac{\lambda^2\sigma^2}{2}}$ so condition (4.5.3) is satisfied, we have equality there, and $\underline{m}$ is the optimal constant in that condition.

If $\xi_1, \xi_2, \ldots$ are independent random variables, each of them has the same distribution as ξ, $S_n = \xi_1 + \cdots + \xi_n$, then the strong law of large numbers is satisfied, i.e., (4.5.4) holds.

However, we prefer the following explicit construction of the sequence $\xi_1, \xi_2, \ldots$, for which (4.5.1) is satisfied. We shall use the original idea of [36] to construct negatively dependent random variables. Consider copies $(\Omega^{(i)}, \mathcal{F}^{(i)}, P_m^{(i)})$, $i = 1, 2, \ldots$ of the probability space $(\Omega, \mathcal{F}, P_m)$. Then construct their product probability space $(\Omega^{(\infty)}, \mathcal{F}^{(\infty)}, P_m^{(\infty)}) = \prod_{i=1}^\infty (\Omega^{(i)}, \mathcal{F}^{(i)}, P_m^{(i)})$. Then let ξ_i be the ith coordinate random variable, i.e., $\xi_i(x_1, x_2, \ldots) = x_i$ for any i. Then, under $P_m^{(\infty)}$, the random variables $\xi_1, \xi_2, \ldots$ are independent and identically distributed, each having distribution P_m, that is, a normal distribution with expectation m and variance σ^2. But we need upper probability and upper expectation. So let $\hat{\mathbb{V}}^{(\infty)}(A) = \sup_m P_m^{(\infty)}(A)$ for any event A and $\hat{\mathbb{E}}^{(\infty)}(X) = \sup_m E_m^{(\infty)}(X)$ for an appropriate random variable X. Let $f_1, f_2, \ldots, f_n$ be non-negative measurable functions. Then

$$\hat{\mathbb{E}}^{(\infty)}(f_1(\xi_1)f_2(\xi_2)\cdots f_n(\xi_n)) = \sup_m E_m^{(\infty)}(f_1(\xi_1)f_2(\xi_2)\cdots f_n(\xi_n))$$

$$
\begin{aligned}
&= \sup_m E_m(f_1(\xi_1)) E_m(f_2(\xi_2)) \cdots E_m(f_n(\xi_n)) \\
&\leq \sup_m E_m(f_1(\xi_1)) \cdot \sup_m E_m(f_2(\xi_2)) \cdots \sup_m E_m(f_n(\xi_n)) \\
&= \sup_m E_m^{(\infty)}(f_1(\xi_1)) \cdot \sup_m E_m^{(\infty)}(f_2(\xi_2)) \cdots \sup_m E_m^{(\infty)}(f_n(\xi_n)) \\
&= \hat{\mathbb{E}}^{(\infty)}(f_1(\xi_1)) \cdot \hat{\mathbb{E}}^{(\infty)}(f_2(\xi_2)) \cdots \hat{\mathbb{E}}^{(\infty)}(f_n(\xi_n)).
\end{aligned}
$$

So (4.5.1) is satisfied. By Theorem 4.5.1, the relation (4.5.4) is true, so the strong law is satisfied.

Summary

In this chapter, we summarize the main results, contributions, and possible future research directions of this dissertation. Building on the findings presented in the previous chapters, we highlight the methodologies employed and discuss the significance of the principal results obtained. For consistency and ease of reference, we retain the original numbering of theorems and mathematical expressions when restating results from earlier chapters.

The dissertation focuses on several versions of the Strong Law of Large Numbers (SLLN) in three major directions: conditional probability frameworks, quantitative rates of convergence, and nonlinear (subadditive) settings. These directions provide a broader perspective on classical SLLN results and contribute to a deeper understanding of convergence behavior under more general probabilistic structures.

We also present and discuss some of the key lemmas, propositions, and theorems that form the core of this research. These results are based on four published papers: Fazekas & Masasila [24], Fazekas & Masasila [25], Masasila & Fazekas [50], and Masasila & Fazekas [49]. Together, these works develop new theoretical insights and extend existing results related to the Strong Law of Large Numbers.

The classical SLLN guarantees almost sure convergence of sample averages under independence and identical distributions. However, in many modern applications, data exhibit various forms of dependence as observed relative to partial information, or arise in multi-indexed structures where convergence rates are also of interest. Moreover, in robust probability models, uncertainty is represented via families of probability measures rather than a single probability measure, leading naturally to nonlinear expectations and capacities. These considerations motivate the extensions developed in this dissertation.

One of the methodological foundation of this thesis is the general framework introduced by Fazekas and Klesov [23], which shows that Kolmogorov-type maximal inequalities imply Hájek–Rényi-type inequalities, and these in turn imply almost sure convergence. A key advantage of this approach is that it does not rely on specific dependence structures but only on suitable maximal inequalities. Another is proof mining approach introduced by Neri [54]. These approaches are extended to conditional settings, quantitative convergence for random fields, and nonlinear expectation frameworks.

Chapter 1: A General Approach to Conditional Strong Laws of Large Numbers

In Chapter 1, we develop a general framework for conditional strong laws of large numbers on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ equipped with a sub-σ-algebra $\mathcal{F}$. The main objective of this chapter is to develop a general inequality-based method for establishing conditional strong laws. The approach extends the classical method of Fazekas and Klesov to the conditional setting. The key observation is that a conditional Kolmogorov-type maximal inequality implies a conditional Hájek–Rényi inequality, which in turn leads to a conditional SLLN.

Let $S_n = \sum_{k=1}^n X_k$ be the partial sums of arbitrary random variables. We prove that if a conditional Kolmogorov-type maximal inequality of the form

$$\mathbb{E}\left(\max_{1\le k\le n} |S_k|^r \,\middle|\, \mathcal{F}\right) \le \sum_{k=1}^n \alpha_k$$

holds for $\mathcal{F}$-measurable nonnegative random variables α_k, then a conditional Hájek–Rényi-type inequality follows:

$$\mathbb{E}\left(\max_{1\le k\le n} \left|\frac{S_k}{\beta_k}\right|^r \,\middle|\, \mathcal{F}\right) \le C\sum_{k=1}^n \frac{\alpha_k}{\beta_k^r},$$

where (β_k) is an increasing $\mathcal{F}$-measurable normalizing sequence.

Under suitable summability conditions on α_k/β_k^r, we obtain the conditional strong law

$$\frac{S_n}{b_n} \longrightarrow 0 \quad \text{almost surely.}$$

Main results

Theorem 1.2.1. *Let $\{X_k, 1 \le k \le n\}$ be a sequence of random variables, let $S_k = X_1 + \cdots + X_k$. Let $\mathcal{F}$ be a σ-subalgebra, $\alpha_1, \dots, \alpha_n$ be nonnegative $\mathcal{F}$-measurable random variables, $r > 0$ real number. Assume that the general conditional Kolmogorov's type inequality is true, that is*

$$\mathbb{E}\left(\left[\max_{1\le l\le m} |S_l|\right]^r |\mathcal{F}\right) \le \sum_{l=1}^m \alpha_l \qquad \text{for all} \quad 1 \le m \le n. \tag{1.2.1}$$

Then the conditional Hájek-Rényi inequality is true, that is

$$\mathbb{E}\left(\left[\max_{1\le l\le n} \left|\frac{S_l}{\beta_l}\right|\right]^r |\mathcal{F}\right) \le 4\sum_{l=1}^n \frac{\alpha_l}{\beta_l^r} \tag{1.2.2}$$

for $\mathcal{F}$-measurable random variables $\beta_1 \leq \beta_2 \leq \ldots \leq \beta_n$ with $\beta_1 \geq \beta_0$, where β_0 is a positive constant.

Theorem 1.2.2. *Let $\{X_n, n \geq 1\}$ be a sequence of random variables, $S_n = X_1 + \cdots + X_n$ for any n. Let $b_0 \leq b_1 \leq b_2 \leq \ldots$ be $\mathcal{F}$-measurable random variables with $b_n \longrightarrow \infty$ a.s., where b_0 is a positive constant. Let $\alpha_1, \alpha_2, \ldots$ be nonnegative $\mathcal{F}$-measurable random variables. Let $r > 0$ be a fixed number. Assume that for any $n \geq 1$*

$$\mathbb{E}\left(\left[\max_{1\leq l\leq n} |S_l|\right]^r |\mathcal{F}\right) \leq \sum_{l=1}^{n} \alpha_l. \tag{1.2.3}$$

If $\sum_{l=1}^{\infty} \frac{\alpha_l}{b_l^r} < \infty$ a.s., then

$$\lim_{n\longrightarrow\infty} \frac{S_n}{b_n} = 0 \qquad a.s. \tag{1.2.4}$$

Parallel results are also established using conditional probability inequalities instead of conditional expectations. These results provide a unified and abstract principle for proving conditional SLLNs without assuming conditional independence or mixing.

Several classical conditional strong laws are recovered as applications, including results for $\mathcal{F}$-independent sequences and conditionally negatively dependent random variables. Thus, Chapter 1 generalizes many existing results and provides a simple and flexible proof technique.

Chapter 2: Quantitative Strong Laws for Double-Indexed Random Variables

Chapter 2 is devoted to quantitative strong laws of large numbers for random variables indexed by two parameters, i.e. random fields $\{X_{i,j}\}_{i,j\geq 1}$. In such settings, convergence must be studied as $(i, j) \to \infty$ in $\mathbb{N}^2$, typically, conditions stronger than the ones for one-dimensional sequences are required. The main goal is to determine the rate at which the sample averages converge to their expected values. Unlike classical SLLN results, which guarantee convergence but do not quantify the speed of convergence, quantitative strong laws provide explicit bounds on the probability of deviations.

We introduce a geometric description of convergence based on families of curves in $\mathbb{R}^2_+$ and show that convergence along suitably chosen subsequences implies convergence over the entire lattice. This approach allows us to translate one-dimensional maximal inequalities into multi-index convergence results.

For normalized partial sums

$$S_{m,n} = \sum_{i=1}^{m} \sum_{j=1}^{n} X_{i,j},$$

we derive explicit probability bounds of the form

$$\mathbb{P}\left(\sup_{(k,l)\succeq(m,n)}\left|\frac{S_{k,l}}{kl}\right|>\varepsilon\right)\le\frac{C}{mn\,\varepsilon^p},$$

which yield quantitative convergence rates.

Main results

Let $\boldsymbol{n}=(n_1,n_2)\in\mathbb{N}^2$ denote the indices and let $|\boldsymbol{n}|=n_1\cdot n_2$ if $\boldsymbol{n}=(n_1,n_2)$.

Proposition 2.3.1. *Let $\{\xi_{\boldsymbol{n}},\boldsymbol{n}\in\mathbb{N}^2\}$, be non-negative random variables, and let $S_{\boldsymbol{n}}=\sum_{\boldsymbol{k}\le\boldsymbol{n}}\xi_{\boldsymbol{k}}$, $\mathbb{E}\xi_{\boldsymbol{n}}=\mu$ for all $\boldsymbol{n}\in\mathbb{N}^2$. If for each $\alpha>1$,*

$$\sum_{\boldsymbol{p}\in\mathbb{N}_0^2,\boldsymbol{k}^{\pm}(\boldsymbol{p})\neq\boldsymbol{0}}\mathbb{P}\left(\left|\frac{S_{\boldsymbol{k}^{\pm}(\boldsymbol{p})}}{|\boldsymbol{k}^{\pm}(\boldsymbol{p})|}-\mu\right|>\varepsilon\right)<\infty,\quad \textit{for all}\quad \varepsilon>0,$$

then

$$\frac{S_{\boldsymbol{n}}}{|\boldsymbol{n}|}\to\mu\quad \textit{a.s. strongly as}\quad \boldsymbol{n}\to\infty. \tag{2.3.2}$$

Proposition 2.3.2. *Let $\{\xi_{\boldsymbol{n}},\boldsymbol{n}\in\mathbb{N}^2\}$, be non-negative random variables, and let $S_{\boldsymbol{n}}=\sum_{\boldsymbol{k}\le\boldsymbol{n}}\xi_{\boldsymbol{k}}$, $\mathbb{E}\xi_{\boldsymbol{n}}=\mu$ for all $\boldsymbol{n}$. Assume that for any $\varepsilon>0$ and any $\alpha>1$*

$$\sum_{\boldsymbol{p}\ge\boldsymbol{l},\boldsymbol{k}^{\pm}(\boldsymbol{p})\neq\boldsymbol{0}}\mathbb{P}\left(\left|\frac{S_{\boldsymbol{k}^{\pm}(\boldsymbol{p})}}{|\boldsymbol{k}^{\pm}(\boldsymbol{p})|}-\mu\right|>\varepsilon\right)\le\lambda\quad \textit{if}\quad \boldsymbol{l}\ge\boldsymbol{\Lambda}_{\varepsilon,\alpha}(\lambda), \tag{2.3.7}$$

where $\{\boldsymbol{\Lambda}_{\varepsilon,\alpha}(\lambda)\ :\ \lambda>0\}$ is a rate. Then

$$\mathbb{P}\left(\sup_{\boldsymbol{m}\ge\boldsymbol{n}}\left|\frac{S_{\boldsymbol{m}}}{|\boldsymbol{m}|}-\mu\right|>\varepsilon\right)\le\lambda\quad \textit{if}\quad \boldsymbol{n}\ge\boldsymbol{\Phi}_{\varepsilon,\boldsymbol{\Lambda},\alpha}(\lambda),$$

where $\boldsymbol{\Phi}_{\varepsilon,\boldsymbol{\Lambda},\alpha}(\lambda)=\alpha^{\lceil\boldsymbol{\Lambda}_{\frac{\varepsilon}{2\alpha^2},\alpha}(\lambda)\rceil}$ is the rate of convergence of $\frac{S_{\boldsymbol{m}}}{|\boldsymbol{m}|}\to\mu$ a.s., and $\alpha^2=\frac{\varepsilon}{2\mu}+1$. Here $\lceil x\rceil$ denotes the smallest integer being not smaller than x and $\alpha^{\boldsymbol{\Gamma}}=(\alpha^{\Gamma_1},\alpha^{\Gamma_2})$ if $\boldsymbol{\Gamma}=(\Gamma_1,\Gamma_2)$.

Therefore, Chapter 2 provides one of the first systematic treatments of quantitative SLLNs for double-indexed arrays under weak dependence conditions, extending classical results of Hsu–Robbins, Baum–Katz, and others to multi-parameter settings.

Chapter 3: Strong Laws of Large Numbers for General Random Variables under Conditional Sub-additive Expectation and Capacity

In Chapter 3, we move beyond classical probability measures and study strong laws under conditional sub-additive expectations and capacities. We extend the study of SLLN into the framework of conditional sub-additive expectations and conditional capacities (non-additive probabilities). Adopting an axiomatic approach, we define necessary properties for conditional sub-additive operators (expectation and probability) without relying on constructive frameworks like G-expectation. As in Chapter 1, the methodology in this chapter follows the general approach established by Fazekas and Klesov in [23], demonstrating that a conditional Kolmogorov-type maximal inequality implies conditional Hájek–Rényi-type inequality, which subsequently implies SLLN. We introduce Finiteness and Recursivity axioms to handle the translation between unconditional and conditional sub-additive measures. Finally, we define the notion of conditionally negatively dependent random variables in non-linear setting and prove that such variables satisfy the necessary maximal inequalities to obey SLLN. The findings generalize previous results by removing restrictions on dependence structures in the general theorems and proving specific results for negatively dependent variables.

The motivation for this framework arises naturally in problems involving model uncertainty, robust statistics, and mathematical finance. In such situations, the expectation of a random variable is not uniquely determined, and instead one considers the supremum and infimum expectations over a family of probability measures. As a result, limit theorems such as the strong law of large numbers take a different form: instead of convergence to a single constant, the limiting behavior is described by an interval determined by the lower and upper expectations.

Main results

The first two theorems ensure that the conditional Kolmogorov inequality for sub-additive expectation implies the conditional Hájek-Rényi inequality both for sub-additive expectation and capacities.

Theorem 3.4.1. *Let $\{X_k, 1 \leq k \leq n\}$ be random variables belonging to the space $\mathcal{H}$. Assume that the conditional expectation operator $\hat{\mathbb{E}}[\,\cdot\,|\mathcal{F}]$ on space $\mathcal{H}$ satisfies the monotonicity, sub-additivity and positive homogeneity axioms of Definition 3.3.1. Let $\alpha_1, \ldots, \alpha_n$ be non-negative $\mathcal{F}$-measurable random variables, and $r > 0$ be real number. Assume that the general conditional Kolmogorov-type inequality is true, that is,*

$$\hat{\mathbb{E}}\left[\left(\max_{1\leq l\leq m}|S_l|\right)^r \Big|\mathcal{F}\right] \leq \sum_{l=1}^{m} \alpha_l \qquad \textit{for all} \quad 1 \leq m \leq n. \tag{3.4.1}$$

Then the conditional Hájek-Rényi inequality is true, that is,

$$\hat{\mathbb{E}}\left[\left(\max_{1\le l\le n}\left|\frac{S_l}{\beta_l}\right|\right)^r\Big|\mathcal{F}\right]\le 4\sum_{l=1}^{n}\frac{\alpha_l}{\beta_l^r} \tag{3.4.2}$$

for $\mathcal{F}$-measurable random variables $\beta_1 \le \beta_2 \le \ldots \le \beta_n$ with $\beta_1 \ge \beta_0$, where β_0 is a positive constant.

Theorem 3.4.3. *Let $\{X_k, 1 \le k \le n\}$ be random variables, $S_k = X_1 + \cdots + X_k$. Let $\hat{\mathbb{V}}[\,\cdot\,|\mathcal{F}]$ be a conditional sub-additive probability satisfying the axioms normalization, monotonicity, and sub-additivity of Definition 3.3.3. Let r be a positive real number. Let $\beta_1 \le \beta_2 \le \cdots \le \beta_n$ be $\mathcal{F}$-measurable, $\alpha_1, \ldots, \alpha_n$ non-negative $\mathcal{F}$-measurable random variables. Assume that $\beta_1 \ge \beta_0 > 0$, where β_0 is non random. If*

$$\hat{\mathbb{V}}\left[\max_{1\le l\le m}|S_l|\ge\varepsilon\big|\mathcal{F}\right]\le\frac{1}{\varepsilon^r}\sum_{l=1}^{m}\alpha_l \qquad \textit{for all}\quad 1\le m\le n \tag{3.4.3}$$

and for all $\varepsilon > 0$, then

$$\hat{\mathbb{V}}\left[\max_{1\le l\le n}\left|\frac{S_l}{\beta_l}\right|\ge\varepsilon\big|\mathcal{F}\right]\le\frac{4}{\varepsilon^r}\sum_{k=1}^{n}\frac{\alpha_k}{\beta_k^r} \tag{3.4.4}$$

for all $\varepsilon > 0$.

Next, we present strong laws of large numbers in terms of conditional sub-additive expectations and capacities.

Theorem 3.5.1. *Let $\{X_n, n \ge 1\}$ be random variables, $S_n = X_1 + \cdots + X_n$ for any n. Let $b_1, b_2, \ldots$ be q.s. finite, $\mathcal{F}$-measurable random variables with $b_0 \le b_1 \le b_2 \le \ldots$ q.s., $b_n \to \infty$ q.s., where b_0 is a positive constant. Let $\alpha_1, \alpha_2, \ldots$ be non-negative $\mathcal{F}$-measurable random variables. Assume that for the conditional expectation $\hat{\mathbb{E}}[\,\cdot\,|\mathcal{F}]$ the axioms in Definition 3.3.1 are satisfied, where all relations among random variables are understood in the $\hat{\mathbb{V}}$-quasi-sure sense. Assume further that the finiteness axiom in Definition 3.3.6 holds. Let $r > 0$ be a fixed number and suppose that, for any $n \ge 1$*

$$\hat{\mathbb{E}}\left[\left(\max_{1\le l\le n}|S_l|\right)^r\Big|\mathcal{F}\right]\le\sum_{l=1}^{n}\alpha_l \quad \textit{quasi-surely.} \tag{3.5.1}$$

If $\sum_{l=1}^{\infty}\frac{\alpha_l}{b_l^r}<\infty$ quasi-surely, then

$$\lim_{n\longrightarrow\infty}\frac{S_n}{b_n}=0 \quad \textit{quasi-surely} \tag{3.5.2}$$

and $\frac{S_n}{b_n} = \mathrm{O}\left(\frac{\beta_n}{b_n}\right)$, where β_n is defined by (3.5.3).

Theorem 3.5.2. *Let $\{X_n, n \geq 1\}$ be random variables, $S_n = X_1 + \cdots + X_n$ for any n. Let $b_1, b_2, \ldots$ be q.s. finite, $\mathcal{F}$-measurable random variables with $b_0 \leq b_1 \leq b_2 \leq \ldots$ q.s., $b_n \to \infty$ q.s., where b_0 is a positive constant. Let $\alpha_1, \alpha_2, \ldots$ be non-negative $\mathcal{F}$-measurable random variables. Assume that the conditional sub-additive probability $\hat{\mathbb{V}}[\cdot|\mathcal{F}]$ satisfies the axioms in Definition 3.3.3, where all relations among random variables are understood in the $\hat{\mathbb{V}}$-quasi-sure sense. Assume further that the recursivity axiom in Definition 3.3.9 holds. Let $r > 0$ be a fixed number and suppose that, for all $n \geq 1$ and all $\varepsilon > 0$,*

$$\hat{\mathbb{V}}\left[\max_{1\leq l\leq n} |S_l| \geq \varepsilon \middle| \mathcal{F}\right] \leq \frac{1}{\varepsilon^r}\sum_{l=1}^{n} \alpha_l \quad \textit{quasi-surely}\,. \tag{3.5.4}$$

If $\sum_{l=1}^{\infty} \frac{\alpha_l}{b_l^r} < \infty$ quasi-surely, then

$$\lim_{n\to\infty} \frac{S_n}{b_n} = 0 \quad \textit{quasi-surely} \tag{3.5.5}$$

with the convergence rate $\frac{S_n}{b_n} = \mathrm{O}\left(\frac{\beta_n}{b_n}\right)$ quasi-surely.

Chapter 4: Strong Law of Large Numbers for φ-Sub-Gaussian Random Variables under Sub-Linear Expectation Spaces

In Chapter 4, we study the strong law of large numbers (SLLN) in the framework of sub-linear expectation spaces for sequences of φ-sub-Gaussian random variables. The results extend classical exponential concentration methods to the non-additive setting, where uncertainty is represented by a family of probability measures rather than a single probability measure. Consequently, the classical expectation is replaced by the upper expectation $\hat{E}$ and the lower expectation E, and the corresponding probability measure is replaced by the pair of capacities $\hat{V}$ and v.

The chapter introduces a notion of φ-sub-Gaussian random variables under sub-linear expectation spaces. In classical probability theory, a sub-Gaussian random variable is characterized by exponential moment bounds centered around its mean. However, under sub-linear expectations the mean may not be uniquely defined, since the upper and lower expectations need not coincide. To address this issue, we introduced two one-sided exponential moment conditions centered at two different parameters $\underline{m}$ and $\overline{m}$, which correspond to the lower and upper expectations of the random variable.

The main result of Chapter 4 establishes a general SLLN for sequences of φ-sub-Gaussian random variables. This result demonstrates that classical exponential concentration methods remain effective in sub-linear expectation spaces when properly adapted to account for mean uncertainty. The use of φ-sub-Gaussian conditions provides a flexible framework for establishing strong laws under non-additive probabilities, and the general theorem obtained in the chapter may serve as a foundation for further developments in the study of limit theorems under model uncertainty.

Main results

Theorem 4.4.1. *For some $p > 1$, let $\{Z_n, n \geq 1\}$, be a sequence of φ_p-sub-Gaussian random variables with parameters $\overline{m}$ and $\underline{m}$. Let $\tau_\varphi(Z_n)$ be defined according to Definition 4.3.4. If there exist positive numbers c and α such that for every natural number n, the condition $\tau_{\varphi_p}(Z_n) \leq c\, n^{-\alpha}$ is satisfied, then*

$$\hat{\mathbb{V}}\left(\left\{\liminf_{n\to\infty} Z_n < \underline{m}\right\} \cup \left\{\limsup_{n\to\infty} Z_n > \overline{m}\right\}\right) = 0 \tag{4.4.1}$$

and

$$v\left(\underline{m} \leq \liminf_{n\to\infty} Z_n \leq \limsup_{n\to\infty} Z_n \leq \overline{m}\right) = 1. \tag{4.4.2}$$

The main result above is used to obtain a strong law of large numbers for independent identically distributed sub-Gaussian random variables in the following way.

The sequence of random variables $\xi_1, \xi_2, \ldots$ is called independent if for each $n = 1, 2, \ldots$ and each non-negative measurable functions $f_1, f_2, \ldots$ we have

$$\hat{\mathbb{E}}(f_1(\xi_1)f_2(\xi_2)\cdots f_n(\xi_n)) = \hat{\mathbb{E}}(f_1(\xi_1))\hat{\mathbb{E}}(f_2(\xi_2))\cdots\hat{\mathbb{E}}(f_n(\xi_n)).$$

If $\xi_1, \xi_2, \ldots$ are independent, then they satisfy the following property

$$\hat{\mathbb{E}}\prod_{i=1}^{k}\exp\left(\lambda(\xi_i - m)\right) \leq \prod_{i=1}^{k}\hat{\mathbb{E}}\exp\left(\lambda(\xi_i - m)\right) \tag{4.5.1}$$

for any real λ, m, and positive integer k.

Now, let $\xi_1, \xi_2, \ldots$ be random variables satisfying the sub-Gaussian property (that is they are φ_2-sub-Gaussian): for fixed constants $\sigma > 0$, $\overline{m}$, and $\underline{m}$

$$\hat{\mathbb{E}}(e^{\lambda(\xi_i - \overline{m})}) \leq e^{(\sigma^2\lambda^2/2)}, \quad \text{for} \quad \lambda > 0, \tag{4.5.2}$$

and

$$\hat{\mathbb{E}}(e^{\lambda(\xi_i - \underline{m})}) \leq e^{(\sigma^2\lambda^2/2)}, \quad \text{for} \quad \lambda < 0. \tag{4.5.3}$$

Let $S_n = \xi_1 + \cdots + \xi_n$ for any positive integer n.

Theorem 4.5.1. *Let $\{\xi_n, n \geq 1\}$ be random variables satisfying (4.5.1), (4.5.2), and (4.5.3). With S_n defined above,*

$$\hat{\mathbb{V}}\left(\left\{\liminf_{n\to\infty}\frac{S_n}{n} < \underline{m}\right\} \cup \left\{\limsup_{n\to\infty}\frac{S_n}{n} > \overline{m}\right\}\right) = 0, \tag{4.5.4}$$

or equivalently,

$$v\Big(\underline{m} \le \liminf_{n\to\infty} \frac{S_n}{n} \le \limsup_{n\to\infty} \frac{S_n}{n} \le \overline{m}\Big) = 1.$$

Overall Contributions

The results developed in this dissertation contribute to the theory of strong laws of large numbers by extending classical probabilistic convergence principles to several modern stochastic settings. The main contributions can be grouped into four directions that address conditional probability structures, multi-indexed random variables, nonlinear expectations, and concentration-type assumptions. These contributions collectively broaden the applicability of the strong law of large numbers and provide methodological tools for studying convergence in complex stochastic systems.

The dissertation investigates conditional, quantitative, and nonlinear generalizations of the strong law of large numbers (SLLN), building upon classical probability theory while introducing new methods and frameworks.

The first contribution concerns **conditional strong laws of large numbers**. Traditional forms of the SLLN typically assume independence or specific dependence structures. However, in many practical stochastic models, additional information is available through conditioning on a σ-algebra representing background knowledge or observed data. Chapter 1 develops a general framework showing that a *conditional Kolmogorov-type maximal inequality implies a conditional Hájek–Rényi inequality*, which in turn yields a conditional strong law of large numbers.

This result is important because it removes the need for specialized dependence assumptions such as conditional independence, conditional mixing, or conditional association. Instead, once an appropriate maximal inequality is verified, the convergence result follows automatically. Consequently, several previously known conditional SLLNs become immediate corollaries of the developed framework, leading to simpler proofs and broader applicability.

The second contribution focuses on **quantitative strong laws for multi-indexed random variables**, particularly random fields indexed by two parameters. Classical strong laws guarantee almost sure convergence but provide little information about the *rate of convergence*. Chapter 2 addresses this limitation by establishing explicit probability bounds that quantify the speed of convergence of normalized double sums.

The approach introduces a geometric representation of convergence through families of curves in the positive quadrant of $\mathbb{R}^2$. By controlling convergence along carefully chosen subsequences, it becomes possible to transfer convergence rates to the entire index lattice. This methodology allows the derivation of quantitative SLLNs for pairwise independent and quasi-uncorrelated random fields, providing one of the first systematic treatments of convergence rates for double-index arrays under weak dependence conditions.

The third contribution extends strong law results to **non-additive probabilistic frameworks**. Classical probability theory assumes a single additive probability measure, but many modern applications involve model uncertainty, ambiguity, or multiple prior distributions. Chapter 3 develops a general theory of strong laws under conditional sub-additive expectations and capacities.

Within this framework, new conditional maximal inequalities are established, which lead to strong laws formulated in terms of capacities and non-additive expectations. Importantly, these results do not require classical independence assumptions. By extending the maximal inequality approach developed earlier in the dissertation to nonlinear expectation spaces, the chapter provides a unified methodology that connects classical probability theory with capacity-based models of uncertainty.

The fourth contribution investigates **strong laws under sublinear expectations with sub-Gaussian assumptions**. While earlier chapters focus on general frameworks and inequality-based methods, Chapter 4 studies the role of exponential-type tail behavior in ensuring strong convergence. The chapter establishes strong laws for φ-sub-Gaussian random variables under sublinear expectation spaces.

These results demonstrate that sub-Gaussian concentration properties can replace classical moment conditions in guaranteeing convergence. As a result, strong laws remain valid even in nonlinear probabilistic environments where traditional assumptions may not hold.

Further Research Directions

The results obtained in this dissertation open several avenues for further investigation. While the main results establish strong laws under conditional frameworks, multi-index structures, and nonlinear expectations, many related questions remain open. Future work could extend these results in several directions, as outlined below.

Future Directions Related to Chapter 1

The framework developed in Chapter 1 shows that conditional strong laws follow from conditional maximal inequalities. One natural direction for further research is the extension of this method to other versions of SLLNs.

Future Directions Related to Chapter 2

The quantitative results for double-index random fields obtained in Chapter 2 suggest several potential generalizations. One possible extension is the study of higher-dimensional random fields, where indices lie in $\mathbb{N}^r$ for $r > 2$. Establishing quantitative convergence rates in such settings would further generalize the results and may require new geometric methods for describing convergence regions.

Another direction involves exploring more general dependence structures. The current results focus primarily on pairwise independence and quasi-uncorrelated variables. Future research could examine quantitative strong laws under mixing conditions, association, or martingale-type dependencies in multi-indexed settings.

Future Directions Related to Chapter 3

The nonlinear framework developed in Chapter 3 raises several open questions regarding strong laws under general capacities and sub-additive expectations. One direction for future work is the potential extension that involves exploring ergodic-type results under capacities, which would connect the theory of non-additive probabilities with ergodic theory and dynamical systems.

Furthermore, the relationship between capacity-based convergence and classical probabilistic convergence could be studied in more detail. Understanding how quasi-sure convergence relates to almost sure convergence in different probabilistic models may provide deeper insights into uncertainty modeling.

Future Directions Related to Chapter 4

The strong laws established for φ-sub-Gaussian random variables under sublinear expectations suggest several additional research directions. One possibility is to investigate other concentration classes of random variables, such as sub-exponential or heavy-tailed distributions, within sublinear expectation frameworks.

Another direction involves extending these results to dependent sequences under sublinear expectations, where the concept of independence differs from the classical definition. Developing strong laws under generalized dependence structures in nonlinear expectation spaces remains an important open problem.

Finally, future work could explore applications of sublinear strong laws in risk theory and robust statistics, where model uncertainty and ambiguity play central roles.

Bibliography

[1] A. Acosta. Inequalities for b-valued random vectors with applications to the strong law of large numbers. *The Annals of Probability*, 9(1):157–161, 1981.

[2] R. G. Antonini, T. Ngamkham, and A. Volodin. On the asymptotic behavior of the sequence and series of running maxima from a real random sequence. *Statistics & Probability Letters*, 83(2):534–542, 2013.

[3] A. Araujo and E. Giné. *The Central Limit Theorem for Real and Banach Valued Random Variables*. Wiley, New York, 1980.

[4] L. E. Baum and M. Katz. Convergence rates in the law of large numbers. *Transactions of the American Mathematical Society*, 120:108–123, 1965.

[5] G. D. Birkhoff. Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences*, 17(2):656–660, 1931.

[6] H. D. Brunk. The strong law of large numbers. *Duke Mathematical Journal*, 15:181–195, 1948.

[7] V. Buldygin and Y. Kozachenko. *Metric Characterization of Random Variables and Random Processes*. American Mathematical Society, Providence, RI, 2000.

[8] X. Chen. The strong law of large numbers under an upper probability. *Appl. Math.*, 3(12):2056–2062, 2012.

[9] Z. Chen, P. Wu, and B. Li. A strong law of large numbers for non-additive probabilities. *Int. J. Approx. Reason.*, 54(3):365–377, 2013.

[10] J. Y. Choi, H. Y. Seo, and J.I. Baek. On the Hàjek-Rènyi-type inequality for conditionally associated random variables. *Communications for Statistical Applications and Methods*, 18(6):799–808, 2011.

[11] G. Choquet. Theory of capacities. *Ann. Inst. Fourier*, 5:131–295, 1954.

[12] Y. S. Chow. Some convergence theorems for independent random variables. *The Annals of Mathematical Statistics*, 37(6):1482–1493, 1966.

[13] T. C. Christofides. Maximal inequalities for demimartingales and a strong law of large numbers. *Statistics & Probability Letters*, 50(4):357–363, 2000.

[14] K. L. Chung. *A Course in Probability Theory*. Academic Press, New York, 2 edition, 1968.

[15] S. Cohen. Quasi-sure analysis, aggregation and dual representations of sublinear expectations in general spaces. *Electron. J. Probab.*, 17(62):1–15, 2012.

[16] S. Csörgő, K. Tandori, and V. Totik. On the strong law of large numbers for pairwise independent random variables. *Acta Math. Hung.*, 42:319–330, 1983.

[17] F. Delbaen. Coherent risk measures on general probability spaces. In K. Sandmann and P. J. Schoenbucher, editors, *Advances in Finance and Stochastics*, pages 1–37. Springer, Berlin, 2002.

[18] P. Erdős. On a theorem of Hsu and Robbins. *Annals of Mathematical Statistics*, 20:286–291, 1949.

[19] N. Etemadi. An elementary proof of the strong law of large numbers. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 55:119–122, 1981.

[20] N. Etemadi and M. Kaminski. Strong law of large numbers for 2-exchangeable random variables. *Statistics & probability letters*, 28(3):245–250, 1996.

[21] I. Fazekas. Convergence rates in the Marcinkiewicz strong law of large numbers for banach space valued random variables with multidimensional indices. *Publ. Math. Debrecen*, 32(3-4):203–209, 1985.

[22] I. Fazekas. On a general approach to the strong laws of large numbers. *J. Math. Sci.*, 200:411–423, 2014.

[23] I. Fazekas and O. Klesov. A general approach to the strong law of large numbers. *Theory Probab. Appl.*, 45(3):436–449, 2001.

[24] I. Fazekas and N. H. Masasila. A general approach to conditional strong laws of large numbers. *Ann. Univ. Mariae Curie-Sklodowska Sect. A*, 78(1):27–35, 2024.

[25] I. Fazekas and N.H. Masasila. Quantitative strong laws of large numbers for random variables with double indices. *Annales Mathematicae et Informaticae*, 62:55–65, 2025.

[26] H. Föllmer and A. Schied. *Stochastic Finance. An Introduction in Discrete Time*. de Gruyter, Berlin-Boston, 2016.

[27] A. Gut. Marcinkiewicz laws and convergence rates in the law of large numbers for random variables with multidimensional indices. *The Annals of Probability*, pages 469–482, 1978.

[28] A. Gut. Strong laws for independent identically distributed random variables indexed by a sector. *The Annals of Probability*, 11(3):569–577, 1983.

[29] A. Gut. *Probability: A Graduate Course*. Springer, 2012.

[30] J. Hájek and A. Rényi. Generalization of an inequality of Kolmogorov. *Acta Math. Acad. Sci. Hungar.*, 6:281–283, 1955.

[31] P. Hartman and A. Wintner. On the law of the iterated logarithm. *American Journal of Mathematics*, 63:169–176, 1941.

[32] P. L Hsu and H. Robbins. Complete convergence and the law of large numbers. *Proceedings of the National Academy of Sciences USA*, 33:25–31, 1947.

[33] M. Hu and S. Peng. Extended conditional G-expectations and related stopping times. *Probab. Uncertain. Quant. Risk*, 6:4, 2021.

[34] S. Hu and M. Hu. A general approach rate to the strong law of large numbers. *Statist. Probab. Lett.*, 76(8):843–851, 2006.

[35] W. Huang and P. Wu. Strong laws of large numbers for general random variables in sublinear expectation spaces. *J. Inequal. Appl.*, 2019:143, 2019.

[36] P. J. Huber and V. Strassen. Minimax tests and the Neyman–Pearson lemma for capacities. *The Annals of Statistics*, 1(2):251–263, 1973.

[37] M. Janisch. Kolmogorov's strong law of large numbers holds for pairwise uncorrelated random variables. *Theory of Probability & Its Applications*, 66(2):263–275, 2021.

[38] J. P. Kahane. Local properties of functions in terms of random fourier series. *Studia Mathematica*, 19(1):1–25, 1960.

[39] O. Klesov. *Limit Theorems for Multi-Indexed Sums of Random Variables*. Springer, Heidelberg, 2014.

[40] K. Knopp. *Theory and Application of Infinite Series*. Blackie & Son, London–Glasgow, 1951.

[41] A. N. Kolmogorov. Über das gesetz des iterierten logarithmus. *Mathematische Annalen*, 101:126–135, 1929.

[42] A. Kuczmaszewska. On the strong law of large numbers for ϕ-mixing and ρ-mixing random variables. *Acta Mathematica Hungarica*, 132(1-2):174–189, 2011.

[43] M. Ledoux and M. Talagrand. *Probability in Banach Spaces: Isoperimetry and Processes*. Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer, Berlin, 1991.

[44] M. Ledoux and M. Talagrand. *Probability in Banach Spaces: isoperimetry and processes*. Springer Science & Business Media, 2013.

[45] F. Maccheroni and M. Marinacci. A strong law of large numbers for capacities. *Ann. Probab.*, 33(3):1171–1178, 2005.

[46] D. Majerek, W. Nowak, and W. Zięba. Conditional strong law of large number. *Int. J. Pure Appl. Math*, 20(2):143–156, 2005.

[47] J. Marcinkiewicz and A. Zygmund. Remarque sur la loi du logarithme itéré. *Fundamenta Mathematicae*, 4:82–105, 1937.

[48] M. Marinacci. Limit laws for non-additive probabilities and their frequentist interpretation. *Journal of Economic Theory*, 84(2):145–195, 1999.

[49] N. H. Masasila and I. Fazekas. Strong law of large numbers for φ-sub-Gaussian random variables under sub-linear expectation spaces, 2026. arXiv:2602.18175 [math.PR].

[50] N. H. Masasila and I. Fazekas. Strong laws of large numbers for general random variables under conditional sub-additive expectation and capacity. *Mathematics*, 14(5):775, 2026.

[51] P. Matula. A note on the almost sure convergence of sums of negatively dependent random variables. *Statist. Probab. Lett.*, 15:209–213, 1992.

[52] P. Matuła and M. Seweryn. On etemadi's subsequences and the strong law of large numbers for random fields. *Rocky Mountain Journal of Mathematics*, 45(5):1555–1563, 2015.

[53] T. F. Móri. On the strong law of large numbers for weighted sums. *Statistics & Probability Letters*, 18:215–220, 1993.

[54] M. Neri. Quantitative strong laws of large numbers. *Electron. J. Probab.*, 30:1–22, 2025.

[55] R. F. Patterson and R. L. Taylor. Strong laws of large numbers for negatively dependent random elements. *Nonlinear Anal.*, 30(7):4229–4235, 1997.

[56] S. Peng. G-expectation, G-Brownian motion and related stochastic calculus of Ito type. In *Stochastic Analysis and Applications: The Abel Symposium 2005*, pages 541–567. Springer, Berlin, 2007.

[57] S. Peng. *Nonlinear Expectations and Stochastic Calculus under Uncertainty*. Springer, Berlin, 2010.

[58] S. Peng. *Nonlinear Expectations and Stochastic Calculus under Uncertainty: With Robust CLT and G-Brownian Motion*. Springer Nature, Singapore, 2019.

[59] G. Pisier. *Probabilistic Methods in the Geometry of Banach Spaces*. Springer, Berlin, 1986.

[60] P. Pollard. Probability tools, tricks, and miracles. http://www.stat.yale.edu/~pollard/Books/Pttm/Woes30jun23.pdf, 2023. Available online.

[61] Yu. V. Prokhorov. On the strong law of large numbers. *Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya*, 14:523–536, 1950.

[62] B. L. Prakasa Rao. Conditional independence, conditional mixing and conditional association. *Annals of the Institute of Statistical Mathematics*, 61(2):441–460, 2009.

[63] A. Rényi. *Wahrscheinlichkeitsrechnung. Mit einem Anhang über Informationstheorie*. VEB Deutscher Verlag der Wissenschaften, 1962.

[64] P. Révész. *The Laws of Large Numbers*. Probability and Mathematical Statistics. Academic Press, New York, 1968.

[65] A. Rényi. On a new axiomatic theory of probability. *Acta Mathematica Academiae Scientiarum Hungaricae*, 6:285–335, 1955.

[66] H. Y. Seo and J. I. Baek. On Hájek–Rényi-type inequality for conditionally negatively associated random variables and its applications. *Journal of applied mathematics & informatics*, 30(3):623–633, 2012.

[67] A. N. Shiryaev. *Probability*, volume 95 of *Graduate Texts in Mathematics*. Springer, New York, 2 edition, 1996.

[68] R. T. Smythe. Strong laws of large numbers for r-dimensional arrays of random variables. *Ann. Probab.*, 1(1):164–170, 1973.

[69] H. M. Soner, N. Touzi, and J. Zhang. Martingale representation theorem for the G-expectation. *Stochastic Process. Appl.*, 121(2):265–287, 2011.

[70] V. Strassen. A converse to the law of the iterated logarithm. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, (4):265–268, 1966.

[71] R. L. Taylor and T. C. Hu. Sub-Gaussian techniques in proving strong laws of large numbers. *The American Mathematical Monthly*, 94(3):295–299, 1987.

[72] R. L. Taylor, R. F. Patterson, and A. Bozorgnia. A strong law of large numbers for arrays of rowwise negatively dependent random variables. *Stochastic Anal. Appl.*, 20(3):643–656, 2002.

[73] T. Tómács and Z. Líbor. A Hájek–Rényi type inequality and its applications. *Annales Mathematicae et Informaticae*, 33:141–149, 2006.

[74] M. J. Wichura. Some strassen-type laws of the iterated logarithm for multiparameter stochastic processes with independent increments. *The Annals of Probability*, 1(2):272–296, 1973.

[75] D. M. Yuan, J. An, and X. J. Wu. Conditional limit theorems for conditionally negatively associated random variables. *Monatsh. Math.*, 161(4):449–473, 2010.

[76] K. Zajkowski. On norms in some class of exponential type Orlicz spaces of random variables. *Positivity*, 24:1231–1240, 2020.

[77] K. Zajkowski. On the strong law of large numbers for φ-sub-Gaussian random variables. *Ukrainian Mathematical Journal*, 73(3):506–512, 2021.

[78] J. Zhang, Y. Tang, and J. Xiong. Conditional strong law of large numbers under G-expectations. *Symmetry*, 16(3):272, 2024.

[79] L. Zhang. Rosenthal's inequalities for independent and negatively dependent random variables under sub-linear expectations with applications. *Sci. China Math.*, 59(4):751–768, 2016.

Research Conferences and Workshop Participation

1. *Quantitative strong laws of large numbers for random variables with double indices,* The 8th Mediterranean International Conference of Pure and Applied Mathematics and Related Areas (MICOPAM2025), September 8-12,2025, Osijek, Crotia.

2. 27th Conference of Young Statistician Meeting (YSM 2023), September 29 - October 1, 2023, Osijek, Croatia.

3. *Smart Public Space in Prosument Energy Transition* International Interdisciplinary Summer School, June 14-23, 2023, Gliwice, Poland.

Publications related to dissertation

1. Fazekas, I., Masasila, N.H. A general approach to conditional strong laws of large numbers. *Ann. Univ. Mariae Curie-Sklodowska Sect. A* **2024**, *78*(1), 27–35.

2. Fazekas, I., Masasila, N.H. Quantitative strong laws of large numbers for random variables with double indices. *Ann. Math. Inform.* **2025**, *62*, 55–65.

3. Fazekas, I., Masasila, N.H. Explicit rate of convergence on strong laws of large numbers for random variables with double indices. In *Proceedings Book of the 8th Mediterranean International Conference of Pure & Applied Mathematics and Related Areas (MICOPAM 2025)*; Alkan, M., Kucukoglu, I., Öneş, O., Pokaz, D., Ribičić Penava, M., Simsek, Y., Eds.; Osijek, Croatia, 2025; ISBN 978-953-46915-0-2.

4. Masasila, N.H., Fazekas, I. Strong laws of large numbers for general random variables under conditional sub-additive expectation and capacity. *Mathematics*, **2026**, *14*(5):775

5. Masasila, N. H., Fazekas, I. Strong law of large numbers for φ-sub-gaussian random variables under sub-linear expectation spaces. *arXiv*, **2026**, arXiv:2602.18175 [math.PR].

Other publications

1. Masasila, N.H., Ngeleja, R.C., Kigodi, O.J. Mathematical analysis of the role of information on the dynamics of typhoid fever. *Open Access Libr. J.* **2025**, *12*, e10109.

2. Kigodi, O.J., Masasila, N.H. Thermodynamic irreversibility of steady viscous Couette flow with convective cooling and temperature-dependent viscosity. *Heat Trans.* **2025**, *1*(1), 1–14.

3. Kigodi, O. J., Masasila, N. H., Faisal, M., Badruddin, I. A., Zedan A. S. H., Chacha, S. C. Thermal and Entropic Analysis of Viscous Fluid Flow in a Porous Channel With Convective Heat Transfer and Magnetic Field Aspects. *Heat Transf.* **2025**.